Title

A Behavioral Approach to Contemporary Electoral Accountability

Permalink

https://escholarship.org/uc/item/10f102kw

Author

Hill, Adam Douglas

Publication Date

2016

Peer reviewed|Thesis/dissertation

A Behavioral Approach to Contemporary Electoral Accountability

By

Adam Douglas Hill

A dissertation submitted in partial satisfaction of the requirements

for the degree of

Doctor of Philosophy

in

Jurisprudence and Social Policy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christopher Kutz, Chair
Professor Sarah Song
Professor Mark Bevir
Professor Niko Kolodny

Fall 2016

Abstract

A Behavioral Approach to Contemporary Electoral Accountability

by

Adam Douglas Hill

Doctor of Philosophy in Jurisprudence and Social Policy

University of California, Berkeley

Professor Christopher Kutz, Chair

Using techniques from the experimental and computational social sciences, this dissertation identifies three central flaws in one of our most important mechanism designs – elections – and identifies institutional methods to solve these problems.

Mechanism design is the "engineering" side of economic theory. Where game theory "takes the rules of the game as given, and [] makes predictions about the behavior of strategic players," mechanism design "is about the optimal choice of the rules of the game." For mechanism designers, identifying the optimal choice of rules typically depends on predicting how individuals will behave in certain environments. Both mechanism design and game theory, therefore, rely on identifying equilibriums – that is, joint strategic solutions where no one involved has an incentive to deviate from the strategy that they are playing.

Economists argue that electoral accountability can be seen a special version of mechanism design. The logic is this. Electoral institutions establish processes for selecting candidates for public office, in such a way as to incentivize them to act in the public interest. To achieve this result, electoral institutions lead to a cascade of mutually-reinforcing incentives. These institutions (a) incentivize voters to retrospectively evaluate officials' performances such that (b) victorious candidates will tend to act in the public interest. Therefore, electoral accountability is a set of institutions, and the incentives that predictably flow from it. The promise of mechanism designers is that, if the institutions are created correctly, officials will be accountable to the public. And this accountability will occur as a matter of course: mechanically.

In fact, however, this mechanism is broken. This dissertation solves its problems. Chapters 2 and 3 represent three efforts at testing rational choice-based theories of electoral accountability through randomized control trials. These studies take aim at concepts that are crucial to our understanding of a central feature driving retrospective evaluation of policy. Chapter 4 operates at a more theoretical level, but nonetheless retains its behavioral approach. It offers an approach to legal chance that highlights the strategic alternatives facing firms. Chapter 5 offers institutional fixes to the problems identified in Chapters 2, 3, and 4.

CHAPTER ONE: INTRODUCTION

I. *The puzzling place of electoral accountability*

In June 2015, a full 18 months before the 2016 Presidential election, The New York *Times* wrote at length about the political problems facing President Barack Obama. The bulk of these owed to the fact that President Obama "never has to face voters again."[1] Due to term limits, of course, President Obama could not seek a third term in office. Term limits are generally thought of as democracy-enhancing rules: they encourage "citizen candidates," prevent career politicians from becoming de facto dictators, and generally encourage diversity and experimentation.[2] But term limits, as described by the *Times*, had quite a different effect for President Obama's Administration. Rather than affect who ruled in the future, term limits were affecting who ruled in the *present*. Rather than operate as democracy-enhancing rules, they were creating a vacuum at the center of the American state.

The inability to face the voters resulted in an inability to "find[] new sources of political energy," and without political energy only the most necessary of projects could advance. Despite his democratic bona fides – President Obama won two elections – term limits created political headwinds. Without an ability to "face the voters" again, the democratic mandate that President Obama once possessed was diminished.

Electoral accountability, according to this narrative, functions as a type of political asset. The mere opportunity to stand for election is valuable. Political assets of this sort have recently become helpful in the types of bare-knuckle political battles that constitutional theorists like Bruce Ackerman and Alexander Bickel refer to as "everyday politics." Unlike "constitutional politics,"[3] which is, in some important yet hard-to-pin-down sense, the realm of principle, everyday politics is the realm of brute political muscle. The *Times* piece, for example, centers on President Obama's difficulty in passing the highly contentious Trans Pacific Partnership, a multi-lateral trade deal. The TPP debate wasn't about matters of high principle: the debate centered on the bare-knuckles issue of which groups won and which groups lost under the deal.[4] Absent an opportunity to stand before the public, President Obama lost a critical political asset.

Yet the inability to face the voters affected President Obama in the realm of "constitutional politics," too. After Supreme Court Justice Antonin Scalia unexpectedly died in mid-2016, a debate immediately broke out over the issue of whether President Obama had the constitutional authority to appoint a new Justice. Those arguing that the President did not possess such authority made one point: that President Obama was a "lame duck." The argument was striking precisely because electoral accountability appeared to be transcending the realm of

---

[1] John Harwood, *For Obama, a Change of Allies*, N.Y. Times, June 18, 2015, at
http://www.nytimes.com/2015/06/19/us/politics/for-obama-a-change-of-allies.html.
[2] *See*, *e.g.*, Ian Shapiro, THE STATE OF DEMOCRATIC THEORY 60 (2003).
[3] *See* Gerald Gaus, JUSTIFICATORY LIBERALISM 232 (1996).
[4] Later, in the 2016 Democratic primaries, globalization and trade become larger issues. At the time, however, such issues were not front-and-center in Washington.

everyday politics: mainstream commentators were suggesting that President Obama lacked the constitutional authority to appoint a new Justice. Of course, the issue split along partisan lines initially. Yet today even some on the left seem to begrudgingly concede that President Obama's inability to face the voters placed a limitation on the President's scope of legitimate authority. (It didn't help those on the left when video surfaced of Vice President Joe Biden, a Democrat, appearing to have made a similar point.) Thus, the ability to face the voters appears to be more than just a political asset for bare-knuckle brawls, as in the TPP trade case. Rather, the possibility of electoral accountability seems to be functioning as something like an independent source of legal authority, one that increasingly appears to rival the traditional sources in the realm of "constitutional politics."

On this narrative, standing for election is an asset because it allows an incumbent to combat what I will refer to as the substance-based conception of electoral accountability. On this substance-based understanding, electoral accountability is used as a veto on power if the official in question is deemed to somehow not be acting in the interests of the public. Those invoking substance-based electoral accountability seem to imagine how the public would vote, if a referendum were held, on the issue in question. There are only a handful of times when the real-world electoral process trumps the imaginary process: just after an election, just before an election, and during a mid-term election when an incumbent's party prevails.

In other words, the meaning of electoral accountability has shifted almost completely to the more *substance*-based understanding. It feels somehow naive or overly pedantic to point out that by winning the 2012 election, President Obama won the right to rule for four years. It feels hopelessly out of touch to point out that one way of understanding electoral accountability is as a mechanism by which voters hold incumbents liable for their past actions, and incentivize future candidates to act in their interests. The meaning, and the significance, of electoral accountability has shifted away from this more *process*-based understanding.

This dissertation is about electoral accountability. Not so much the conceptual meaning of it, but the more nitty-gritty elements: political blame, accountability in the law-making process, and changing unpopular laws. Nonetheless, we will start here, with the concept of electoral accountability. If we are going to be discussing electoral accountability, we should begin by establishing just what it is we are talking about, if only briefly and in broad strokes.

Substantive electoral accountability operates as a barrier to democratic rule. By elevating the opinion of elites – journalists, politicians, academics – over the opinions of voters, substantive electoral accountability devalues the electoral process. To be sure: substance-based electoral accountability has its merits. In the TPP case, for example, those elites holding up the trade deal very accurately grasped the will of the electorate at that moment. The extent to which the public had turned against globalization became clear during the 2016 Democratic primaries. Nonetheless, the substance-based approach is simply incompatible with the rule of law. It creates confusion about who, exactly, rules us. And it creates a vacuum at the center of government.

Substantive electoral accountability, however, arose because of the failures of the process-based model. This dissertation tackles the primary failures of the process-based model. By pointing out its flaws, it opens avenues to rejuvenate the process-based approach.

After our brief discussion of the concept of electoral accountability, the Introduction will proceed to describe the goals of the dissertation, why these goals matter, and how the dissertation plans on accomplishing the goals.

II. *Two concepts of electoral accountability*

In order to help us get a better handle on substantive electoral accountability, let's consider an analogy.

Imagine that you are at an auction. Fine art, farm equipment, or government-seized property: the items on the block don't matter much. After much yelling and hand waiving, the auction concludes. Then the auctioneer yells, "OK, the auction's over; now it's time to put the goods in the hands of those who value them the most!"

How would you react? Most people, I'm guessing, would be confused. After all, an auction is a process designed precisely to put goods in the hands of those who value them the most.

Next, imagine that you are watching the final game of the NCAA basketball tournament. The buzzer sounds, and victorious team celebrates at half-court. Then the announcers yell, "OK, now it's time to figure out who the best team is!" Again, it would be hard to know how to respond. Single elimination tournaments are designed precisely to determine who the best team is.

Today, a logic very much like that found in the above two examples confuses our democratic process. Yet unlike the cases of the auctioneer or basketball tournament we do not recognize the logic as odd. In fact, just the opposite is true. Across the political spectrum, we treat this logic as a legitimate feature of the democratic process. This flawed logic confuses our understanding one of, if not the, central features of democracy: electoral accountability.

All theories of democracy find a central role for electoral accountability. Elected officials must "face the voters." In so doing, the public judges politicians' performance; a verdict is rendered. But just what role does electoral accountability play in democratic societies? Is electoral accountability a regular and predictable institutional process, which legitimately authorizes victorious candidates to wield the powers of office for a given term? This regular and institutionalized process is that which is envisioned by democratic theory and the rule of law: citizens vote, and officials legitimately wield power for a specified period of time. The alternative vision of electoral accountability minimizes the worth of the process. On this conception, electoral accountability something less regular and predictable: not a process at all, but rather a value, which, *when achieved*, licenses an office holder to wield the powers of office. This latter, substance-based conception of electoral accountability has achieved the status of conventional wisdom.

Despite its prominence, this conventional wisdom suffers from the same illogic as the auction and basketball tournament examples we considered earlier. When politicians, courts, or

citizens call for democratic accountability, the situation is not unlike an auctioneer, after an auction has concluded, saying "OK, now it's time to put the goods in the hands of those who value them most!" Why? The process (the voting) produced a winner, holding certain authority, for a defined period of years. Critics who would deny President Obama the authority he already won are squarely analogous to the auctioneer who would deny the highest bidder his painting.

II.A *The intuition behind the substantive theory of electoral accountability*

Perhaps we should not arrive at this conclusion so quickly. Let's take a step back, and return to the case of the (apparently) confused auctioneer. In doing so, we can gain a better appreciation for the intuition behind the substantive account of electoral accountability.

The most natural response to the confused auctioneer is to explain how auctions work. The auctioneer's remarks suggest that he simply *doesn't understand* what just happened. An auction is a process designed to fetch the highest price for goods. To suggest that those goals need effectuating through some *further* process is to misunderstand what just happened. The same applies to the basketball announcer: a tournament *just is* a process for selecting the best team.

It is possible, however, that the auctioneer and announcer very well understand the point of the auction and tournament. A tournament might be designed to find the best team, but in fact fail to accomplish its goal. In that case, it would make plenty of sense to suggest that the best team might not have been identified. If this is true, then the announcer's comment makes much more sense. In fact, a critique to this effect is often made of the NCAA tournament. The tournament is known as "March Madness" because of its single-elimination format. One slip up, and even the best team is out. Critics argue that, while the "madness" makes the tournament very entertaining, the tournament's single elimination format does reliably produce the best team. Thus, the auctioneer or announcer could be either terribly informed about the processor very well informed about it.

How do we decide if the auctioneer is terribly informed or excellently informed? First, we would ask if there is some external standard by which we can judge the outcomes of the process in question. In the case of the auction, there is: Does the person who will pay the most for a good at a certain time buy the good through the auction? We might face epistemic hurdles in determining whether this is true, but in theory we can know the answer to this question. To stretch John Rawls' terminology, the auction is an example of imperfect procedural justice. There is an independently good, just, or valuable outcome; and the procedure only imperfectly produces that outcome.

The basketball tournament example is a bit more complicated. In sports, we have factors that help us evaluate the quality of teams *other than* the outcome of the tournament. These criteria might include things like the quality of individual players, statistics, and probably the oldest external criteria in sports, the "eye test." Because of the widespread sense that tournaments often do not produce the best team, it is safe to say that tournaments are less perfect examples of procedural justice than auctions are.

Processes tend to produce outcomes less well when the assumptions underlying the procedure change. For example, in-person auctions came to look hopelessly imperfect in the internet age. The goal of auctions is to put goods in the hands of those who value them the most, and limiting the pool of potential buyers by geography appeared to be an unnecessary source of error. Today, most auctions of any significance occur with at least an online component. In such cases, a shift in underlying fundamentals affects the procedure's ability to achieve its purposes.

Even revered, time-honored procedures can begin to look flawed in the face of changing conditions. So, to know if the auctioneer was terribly informed or in fact very well informed about auctions, we need to know how well (or poorly) the auction is achieving its goal. It is quite possible that the auction was so poorly run that the auctioneer rightly surmised that the underlying purpose of the auction had been frustrated. If the imperfections in the procedure are significant, it may fail to achieve its purpose.

In response, an interested party might abandon focus on the procedure, and instead focus on the aims of the procedure. Hence, it might make sense for the auctioneer to try to put the goods in the hands of those who value them the most if there are serious flaws with the auction. Likewise, those who see the NCAA Tournament as failing to produce the best team might create computer simulated tournaments, where each team plays each other team many times. Such unofficial, unsanctioned, and alternative procedures are ways to try to approximate the social goal that the official procedure is failing to realize.

We are currently at such a moment with one of the most important procedures in our society, the democratic procedure, and the process of electoral accountability, in particular. Assumptions underlying the process have shifted, and the procedure is no longer producing the results it aims at. This is why references to "democratic accountability" proliferate. When politicians, courts, or citizens call for democratic accountability, the situation is not unlike an auctioneer, after an auction has concluded, saying "OK, now it's time to put the goods in the hands of those who value them most!" These actors are seeking to approximate a result that is not being achieved through the procedures designed to accomplish it.

The core of this dissertation is about identifying some of the problems that are causing (process-based) electoral accountability to fail to achieve its goals. The hope is that, by identifying these flaws, we can return to a (suitably modified) version of process-based accountability. But before we can delve into the problems we are currently facing, we should first get a handle on how the process of electoral accountability works in theory.

III. *Electoral accountability as mechanism design*

In contrast to the substance-based account of electoral accountability, the process-based account sees electoral accountability as a *mechanism* aimed at reliably producing certain results.

Mechanism design "can be thought of as the 'engineering' side of economic theory," as opposed to the purely explanatory side of economic theory.[5] It is useful to compare mechanism design to the more well-known fields of game theory and decision theory. Where game theory

---

[5] Eric Maskin, *Mechanism Design: How to Implement Social Goals*, 98 AM. ECON. REV. 567, 568 (2008).

"takes the rules of the game as given, and [] makes predictions about the behavior of strategic players," mechanism design "is about the optimal choice of the rules of the game."[6] For mechanism designers, identifying the optimal choice of rules typically depends on predicting how individuals will behave in certain environments. Both mechanism design and game theory, therefore, rely on identifying equilibriums – that is, joint strategic solutions where no one involved has an incentive to deviate from the strategy that they are playing.[7]

The most general way to describe a mechanism is as a way of reliably producing social outcomes. Mechanisms are valuable because they offer social planners assurance (as much as is possible) that certain goals will be realized. For example, the rules of basketball are designed to result in the best team having the most points at the end of a game. "Best" here isn't circular: If a quirk in the rules is tending to produce winners that aren't the best, the rule will be changed. This is how the shot clock was introduced. Weaker teams could hold on to the ball for large chunks of the game, resulting in skewed outcomes. Likewise, an auction is designed to put goods in the hands of those who value them the most. And a faculty search committee is designed to produce the best hire.

Economist Eric Maskin argues that electoral accountability is a special version of this more general way of trying to control the future state of the world. Electoral institutions establish processes for selecting candidates for public office, in such a way as to incentivize them to act in the public interest. To achieve this result, electoral institutions lead to a cascade of mutually-reinforcing incentives. These institutions (a) incentivize voters to retrospectively evaluate officials' performances such that (b) victorious candidates will tend to act in the public interest.[8] Electoral accountability is a set of institutions, and the incentives that predictably flow from it. The promise of mechanism designers is that, if the institutions are created correctly, officials will be accountable to the public. And this accountability will occur as a matter of course: mechanically.

III.A *How to understand appeals to electoral accountability*

We can now see why the process-based account of electoral accountability is incompatible with the substance-based account. If electoral accountability works as advertised (that is, mechanically), then we shouldn't need to make reference to it as a political value. If a process works as advertised, one need not appeal to the values the process is intended to produce. "Mere values" – that is, values without institutional mechanisms to realize them – typically work only when explicitly or implicitly appealed to. Mere values must be addressed, precisely because there is no institutional mechanism to realize them.

Likewise, if electoral accountability works as advertised, why should we need to treat it as a quasi-legal concept, as Republicans seeking to stop the appointment of a new Justice did?

---

[6] Tilman Borgers and Doug Smith, *Robust Mechanism Design and Dominant Strategy Voting Rules*, 9 Theoretical Economics 339, 340 (2014).

[7] Maskin, *supra* note 5, at 571, fn 4.

[8] This cascade is complex, but it is not unique to electoral mechanisms. Eliminating the shot clock in basketball led to changes in game plans, which lead to changes in the selection of players, which resulted in further changes to game plans. The ontology of a mechanism, then, is part institution, and part what Robert Nozick refers to as an "invisible hand" phenomenon.

Sources of legal authority, like political values, must be appealed to. Law requires a process of "naming, blaming, and claiming." If electoral accountability worked as advertised, we would not need to appeal to it as a source of quasi-legal authority. Indeed, it would be redundant to do so.

Both treating electoral accountability as a political value or as a quasi-legal value are ways to trying to give force to the moral value of one person-one vote. In the absence of a well-functioning mechanism for achieving this goal, we reach for politics and law. We see electoral accountability transmuting to, alternatively, a political value (the TPP case) and quasi-constitutional norm (judicial appointments) – as opposed to a mechanism that we can rely on – because it is not functioning as intended. We see appeals to electoral accountability because our current approach is not functioning as intended. Our current electoral mechanisms are premised on what I will call the "traditional" model of accountability. This model has three serious shortcomings.

III.B *Three flaws in the traditional model*

First, the traditional model treats elected representatives as the primary actors in government. In fact, the demise of the non-delegation doctrine, and the rise of intra-government delegation, has increased the number of actors involved in making policy beyond the legislature. In other words, the first flaw is about ***who*** exercises political power.

The second flaw is about ***how*** that power is exercised. The traditional model divides the world into coercive lawmaking, which is characterized as endogenous to politics, and "probabilistic background factors," which are exogenous to politics. In fact, this divide is not nearly so neat. The rise of new regulatory tools, like nudging, has deeply complicated the relationship between endogenous and exogenous factors.

If the first flaw is about who exercises power, and the second is about how power is exercised, the third flaw is about ***when*** power can be exercised. The traditional model assumes that congress-people can change course when voters demand it. But this assumption sits uncomfortably with decades of jurisprudence holding that legal change must be slow and gradual, lest it upset stable background institutions and individuals' reasonable expectations.

The traditional model of electoral accountability has driven half a century of research. It has contributed much to our understanding of how elections produce a correspondence between public outcomes and private preferences. But it has failed to keep up with the development and growth of major political institutions and practices. This failure has led to the groping for new ways of realizing the moral value of one person-one vote. These ad hoc approaches - treating electoral accountability as a political or quasi-legal value - are insufficient. One person-one vote is so significant that it should be mechanized, not dependent on ad hoc appeals. One person-one vote should be assured. This dissertation is the first step along that path.

In order to chart a course back to a world where electoral accountability is assured, we need a broader picture of the traditional model that has failed.

IV. ***The traditional approach to process-based electoral accountability***

Since the path-breaking work of V.O. Key in the mid twentieth century,[9] theories of electoral accountability have sought to explain how social planners can use competitive elections to allow citizens to hold officials accountable.[10] The traditional theory has two main results. First, it explains how private information held by citizens is translated into public policy (more specifically, translated to the production of public goods). Second, it explains the "self-reinforcing" nature of democracy. That is, it explains why those in power chose to hold elections *at all* (as opposed to merely using the power they already possess).

These two results link electoral accountability to the production of public policy. To see this, consider the problem facing a wholly benevolent dictator, who seeks to provide a set of common goods to the public. The first problem the benevolent dictator would face is that of determining just what goods to provide. If the dictator knows the optimal basket of public goods to provide, then the solution to the challenge is simple: the dictator can simply raise funds for the goods and provide them.

It is unlikely, however, that even a wholly benevolent dictator can maximize net social surplus (defined as gross social benefit minus the cost of providing the goods) alone. This is because net social surplus is a product of citizens' preferences. Just as one would not ordinarily expect an auctioneer to know how much individuals will bid, likewise we tend to think that governments will not know, ex ante, citizens' preferences. This theoretical point is often borne out when political parties seek to select candidates, define their platforms, in the event of an electoral defeat, redefine themselves.[11]

The first goal of a theory of electoral accountability, therefore, is to explain how to extract information about citizens' preferences and translate it into policy. The solution is as indirect as it is elegant. It depends on an insight, first formalized as an economic theory by Key, about the way in which rational voters use elected officials' past performance as the basis for future decisions. Key hypothesized that voters will hold officials accountable for policies that do not accord with voters' preferences. This threat of electoral sanction incentivizes officials to act in the public interest. Elections are "a sanctioning device that induces elected officials to do what the voters want," since "[t]he anticipation of not being reelected in the future leads elected officials not to shirk their obligations to the voters in the present."[12] The voter, then, is the "rational god of vengeance and reward."[13]

IV.A *The problem of agency slack*

---

[9] V.O. Key, The Responsible Electorate: Rationality in Presidential Voting, 1936-1960 (1966).

[10] Kenneth Arrow and Anthony Downs wrote alongside Key, and together the three defined the modern "economic" approach to democratic theory.

[11] Bendor et. al, *Theories of Delegation*, 4 ANN. REV. POL. SCI. 235 (2001).

[12] James Fearon, Electoral accountability and the control of politicians: selecting good types versus sanctioning *poor performance*, in Democracy, Accountability, and Representation 56 (1999).

[13] Key, *supra* note 9, at 568.

The more successful the voter in doling out vengeance and reward, the less "agency slack" in the system of representative democracy. Agency slack represents the difference between what voters want, and what elected officials do. The less slack in the system, the more the electoral system promotes the preferences of voters. Hence, "[t]he standard justification for electoral democracy is that competitive elections give leaders an incentive to provide public goods and more generally to align public policy with citizens' preferences."[14]

Promoting the preferences of voters is important because "[v]oters and citizens are the ultimate principals of the policymaking process."[15] "[C]ongressmen are" merely "their agents."[16] Voters delegate power to elected representatives. Delegation is an implicit or explicit contract that conditionally grants "authority from a principal to an agent that empowers the latter to act on behalf of the former."[17] Although a "contract" is in place between principal and agent, the agent is constrained primarily by the sanctioning and reward mechanism. Principals select agents with regard to their governance capacity, not their governance goals.[18] Sanctioning and reward mechanisms are what determines how the agent acts.

Across political, economic, and social life, principals have ways of reducing the "agency costs" that inevitably arise when power is delegated. These methods seek to reduce the costs of shirking, information asymmetries, monitoring costs, and so forth. In the context of democracy, elections are the primary tool for reducing agency slack.[19]

If elections are efficacious, officeholders' decisions are driven by a comparison of the present value benefit of two alternative income streams: the benefits of impressing the voters and (increasing the probability of) remaining in office versus the benefits of shirking and (increasing the probability of) leaving office.[20] Although performance depends jointly on the actions of the officeholder and exogenous factors outside of the officeholder's control, the electorate is able to observe performance indicators, such as the state of the economy. If, as Ronald Reagan asked, voters are "better off [today] than [they] were four years ago," then the officeholder is likely to be reelected.

Traditional conceptions of electoral accountability treat such outcomes as an indication that "elected officials [did] what the voters want."[21] We can assume that agent-representatives' true preferences differ from those of his constituents.[22] That is, because the agent-representatives did what the public wants, they diverged from what *they* wanted to do. Thus, if voters are happy, we can assume that elections are operating to incentivize agent-representatives: elections are

---

[14] James Fearon, *Self-enforcing democracy*, 126 Q. J. ECON. 1661 (2011).

[15] Barry Weingast, The congressional-bureaucratic system: a principal agent perspective (with application to the SEC), 44 PUBLIC CHOICE 147, 151 (1984).

[16] *Id.*

[17] Abbott, et al., Two Logics of Indirect Governance: Delegation and Orchestration, 11 BRITISH J. OF POL. SCI. 1, 3 (2015).

[18] *Id.* at 5. *See also* Healy et al., *Partisan Bias in Blame Attribution: When Does it Occur?*, 1 J. EXPERIMENTAL POL. SCI. 144 (2015).

[19] They are not, however, the only tool. Separation of powers also operates to reduce agency slack, by reducing the probability that factional interests will confound the general will.

[20] R. Barro, The control of politicians: an economic model, 14 PUBLIC CHOICE 19 (1973).

[21] Fearon, *supra* note 12, at 56.

[22] Barro, *supra* note 20, at 25.

obliging agent-representatives to act in certain ways, on pain of sanction. This conclusion, in turn, allows an observer to conclude that a contractual agency relationship exists.

But why should elected officials, once in office, continue to fund and hold elections?[23] Social contract theorists, such as Locke, recognized that the threat of rebellion can be used as a last resort to keep elected officials in line. The problem with this solution is it assumes that the citizenry has access to good information about incumbent performance. In fact, in increasingly complex and technocratic societies, evaluating official performance from without is very difficult. The closest incumbents can come to demonstrating to the public that they are acting in the public's interest is to hold elections. Of course, gross malfeasance might be revealed in any particular case. Nixon's recorded phone calls, for example, offered strong proof of malfeasance. But as an *institutional* solution, we can't rely on evidence of this sort emerging. Holding elections, with easily understood and public rules, offers an institutional solution to the coordination problem of evaluating incumbent performance, and thus incentivizes incumbents to fund and hold future elections.

The continued maintenance of agency relationship, as confirmed by the existence of elections, authorizes otherwise ordinary individuals to govern through coercive measures. We observed earlier that mechanisms must have some goal or value to guide their design process. With its roots in American democracy, traditional conceptions of electoral accountability are tightly aligned with the social contract theory of democratic legitimacy.[24] Hence, it was of utmost importance to explain how certain individuals could come to possess the power to coerce the population.

To put the point differently, on the traditional view of electoral accountability, if we could *not* identify a mechanism through which public preferences were reliably turned into policy, we could conclude that the purported agent-representatives are, in fact, not authorized to govern coercively.

For all the traditional theory has to offer, however, it is broken. It no longer comports with the reality of our democratic system. As a result, the entire idea of process-based electoral accountability is being challenged. And so however elegant the traditional theory, it needs revision.

VI. *Contemporary electoral accountability*

The dissertation offers a *contemporary* approach to (process-based) electoral accountability. Contemporary electoral accountability situates the study of electoral accountability in an environment that highlights important aspects of the contemporary political

---

[23] This prospect seems a bit far fetched, but we might be witnessing a related phenomenon today in the form of gerrymandering. Gerrymandering is, in a sense, a way of avoiding the practical realities of elections without cancelling them.
[24] *See, e.g.,* Fearon, *supra* note 14 (considering Lockean versions of accountability).

landscape. As mentioned earlier, these updates can be thought of as the *who*, the *how*, and the *when* of electoral accountability.

But this is not the only important revision this dissertation makes to the traditional theory of electoral accountability. Where the traditional theory offers a rational choice based account, this dissertation offers a behavioral approach to the study of electoral accountability. Since the specific empirical findings depend on these behavioral methods, this section begins with a brief overview of some of the behavioral methods used here and then moves to a discussion of the specific findings.

VI.A The behavioral approach

Most of the prior work contributing to the traditional conception of electoral accountability uses the rational choice approach. By contrast, this dissertation offers a behavioral approach to contemporary electoral accountability. A behavioral approach draws on the methodology and substantive insights developed by experimental economics and cognitive psychology. Although frequently touted as a competitor to rational choice approaches,[25] behavioral approaches are better described as working in tandem with rational choice approaches. At their core, behavioral approaches test rational choice models by subjecting their assumptions to empirical scrutiny, often in the form of randomized control trials or quasi-experiments. Some behavioral approaches do more: they compare the concepts used by RCT approaches with what are often thicker, richer concepts derived from philosophy or other humanities. If these concepts perform better than the rational choice concepts, then, ideally, they will be incorporated in suitably modified models.

This pattern - integrating behavioral insights into rational choice models - is widespread across the social sciences. Yet it is peculiarly absent in the electoral accountability literature. To be sure, some scholars have sought to incorporate behavioral insights. Most notably, Bendor et al. authored *A Behavioral Theory of Elections* in 2011.[26] Arguing that "a change is overdue," the book length work sought to incorporate insights around issues of bounded rationality to the study of electoral accountability.[27] The key contribution of Bendor et al. was to show, as a formal matter, why voting is not irrational. In working on the paradox of voter turnout, Bendor et al. addressed a key issue that long beguiled the rational choice literature. Notably, however, Bendor et al. confined their contributions to formal modeling. Thus, the pattern that has spread across the social sciences - where rational choice-based theories bring on board insights developed by behaviorists - has failed to fully materialize in the study of electoral accountability.[28]

---

[25] Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14 J. ECON. PERSPECTIVES 137 (2000).
[26] Bendor et al., A Behavioral Theory of Elections (2011).
[27] *Id*. at 3.
[28] The primary finding that results from applying the behavioral methodology to the study of electoral accountability is that new institutions - in particular, the growth of the administrative state - appear to impede electoral accountability significantly. To be clear: as Chapter Two shows, delegation as such impedes accountability. To put the point differently: if citizens could cast votes for bureaucrats, delegation would still tend to insulate principals and leave agents "holding the bag." Likewise, the new tools that governments employ, such as nudging, can impede electoral accountability. The complexity wrought by delegation, and the indirectness wrought by soft paternalism, reduce the quality of individual judgments about blameworthiness. Fortunately, we can design institutions that debias these tendencies.

The behavioral approach used in this dissertation draws on insights from economics, psychology, and philosophy. The two empirical chapters in the dissertation use experimental designs to test hypotheses derived from rational choice models and philosophical insights. These hypotheses are tested using data gathered by Amazon Mechanical Turk. Mechanical Turk allows researchers to test hypotheses on participants worldwide. The behavioral approach allows us to address many important questions that other methods struggle with. For example, many studies of electoral accountability struggle with intractable questions about how informed voters are, or how competent they are. As the below example shows, we are able to elegantly avoid these issues here.

VI.A.1 A normative benchmark

A core struggle facing accounts of electoral accountability is the question of how competent voters are. Using the behavioral approach described above, the dissertation is designed to avoid intractable questions about voter competence. This provides one example of how the behavioral approach improves on and complements the rational choice method.

Whatever one might conclude about voters' beliefs, intelligence or capacity for reasoned deliberation, it is clear that voters' actions establish a normative benchmark. Elections are at the center of an unrivaled normative hegemon: the practice of democracy. While conceptions of democracy are many and contested, a system of free, fair, and open decision-making is at the heart of all of them. Whatever negative conclusions you might draw about voters' ability to shape elections, it is clear that is has been *good enough* to be the crucial moment in an unrivaled institutional practice. Whatever theory best explains why elections are good - whether they respect citizens qua equals or tend to produce the best outcomes - it is clear that voters perform well enough to ground the legitimacy of the practice.

The dissertation seeks to compare the efficacy of new practices with traditional practices. If voters' performance in elections with traditional institutions is good enough to establish a normative benchmark, then we should gauge the efficacy of new institutional arrangements against our benchmark arrangements.

Avoiding intractable issues of voter competence through the benchmark approach is just one of the many advantages of pursuing the study of electoral accountability through behavioral methods. Rather than continue to enumerate these advantages, let us move immediately to the *product* of this method: the empirical findings.

VI.B *The who: incorporating the administrative state into electoral accountability*

The traditional model treats elected representatives as the primary actors in government. In fact, the demise of the non-delegation doctrine, and the rise of intra-government delegation, has increased the number of actors involved in making policy beyond the legislature. As a result, the process of policymaking has become more complex. In the traditional electoral accountability narrative, the primary actors in the system are principal-voters and agent-representatives. By contrast, the administrative state is not a meaningful part of the story. A

handful of pieces *have* sought to incorporate the administrative state; but those accounts stress that "it is not a bureaucratic imperative that drives agency decision-making but rather a congressional-electoral one."[29] On these stories, agencies that fail to adhere to congressional discipline are sanctioned; "political discipline and motivation for regulatory policy choice works through Congress."

This dissertation tells a richer story about agency action, and the consequences for electoral accountability. Looking specifically at how government officials are disciplined by the public, Chapter Two reveals that officials can avoid taking full responsibility for their actions by taking certain types of actions. By avoiding responsibility for their actions, officials avoid electoral consequences for their actions. This avoidance of electoral consequences stands in direct contrast to traditional electoral accountability theories. If governmental actors are able to avoid consequences, the whole idea of the electorate "sanctioning" government officials collapses.

Officials are able to avoid responsibility by delegating (Chapter Two) to agencies. When elected representatives delegate to agencies, agencies take a disproportionate share of the blame. In fact, even if the elected representative is wholly responsible for the policy, and the agency has no autonomy over the policy, third party observers will attribute some non-trivial measure of blame to the agency.

VI.B *The how: incorporating new tools available to policymakers*

The second flaw with the traditional approach is about how power is exercised. The traditional model assumes officials govern through coercive lawmaking. In fact, the rise of new regulatory tools, like nudging, means that officials govern not only through the issuance of norms backed by force. Today, officials govern equally often by changing the context in which individuals make their decisions. For example, speed bumps are often better ways to regulate the speed at which individuals drive than are speed limits. Speed bumps change the context of driving, while speed limits change the norms around driving. These changes in context are designed to make compliance easier.

The traditional electoral accountability story has nothing to say about these types of governance. If there were no relevant differences between using law and using nudges, then this would not be an issue. In fact, however, there is an important difference between the two. As Chapter Three shows, third party observers are much less likely to blame officials who use nudges to regulate than they are to blame officials who use legal norms to regulate. This implies that officials can escape blame through the use of newer means of governance. And as in Chapter Two, the ability to act yet escape blame undermines the logic of the traditional account.

VI.C *The when: incorporating background conditions*

So far we have touched on who exercises power, and how power is exercised; the final empirical chapter in the dissertation looks at when power can be exercised. Under the traditional model, congresspeople can change course when voters demand it. But this assumption sits

---

[29] Weingast, *supra* note 15, at 149.

uncomfortably with decades of jurisprudence holding that legal change must be slow and gradual, lest it upset stable background institutions and individuals' reasonable expectations.

Traditional electoral accountability stories assume that a change in policy seamlessly translates to a change in regulation. Another way to put this is that traditional electoral accountability treats policy change as costless: if individual preferences shift, regulation should follow suit. However, almost every legal philosopher and social scientist in the 20th century argues that laws must be stable in order to be effective. This is primarily due to the importance of maintaining reasonable expectations.

Chapter Four of the dissertation seeks to reconcile these competing visions. It shows how different individuals and corporate entities will respond differently to changes in law, depending on their background plans. It sketches a framework for tallying the costs and benefits of legal change. In so doing, it suggests a way for courts to determine how frequently law can change.

VII. *A new brand of electoral accountability*

It is one thing to provide new empirical evidence supporting the claim that the traditional electoral accountability is not, in fact, compatible with the contemporary political world. This dissertation goes beyond that. In Chapter Five, it returns to the traditional theory of electoral accountability and suggests avenues for revising the theory. Chapter Five, then, offers a path forward for those interested in revising process-based accounts of electoral accountability.

The most promising vision for updating traditional electoral accountability is to focus on the primary policymakers in contemporary democracies: administrative agencies. The conventional wisdom among public law scholars holds that regulatory policies created by administrative agencies "should be maximally responsive to the preferences of a majority of the electorate."[30] The difficulty is that the officials staffing agencies are appointed, not elected. Scholars seeking to legitimate the "headless fourth branch" of government through electoral accountability need a mechanism by which citizen preferences about regulatory policy can discipline agencies. According to the conventional wisdom, "the best way to assure bureaucratic responsiveness to majoritarian preferences is to make agency policy choices as responsive as possible to the preferences of the elected political leadership."[31] This approach was first made prominent by now-Supreme Court Justice Elena Kagan. The central insight, as articulated in *Chevron* v. *Natural Resources Defense Council, Inc.*, is that "[w]hile agencies are not directly accountable to the people, the Chief Executive is ... ." 467 U.S. 837, 865 (1984). One way, therefore, to approach the issue of electoral accountability is through what we might call "derivative liability." On this view, elected officials can be held derivatively liable for the actions of administrative agencies under their control.

At first glance, derivative accountability promises to offer an electoral mechanism through which agencies can be subjected to political discipline. While this view is promising, it suffers from three fatal flaws. First, agencies do so much that it is hard to ask voters to calculate all of their actions and assign them to a single individual. There are 15 executive departments in

---

[30] Matthew Stephenson, *Optimal Political Control of the Bureaucracy*, 107 MICH. L. REV. 53, 56 (2008).
[31] *Id.*

President Obama's Cabinet. Given the empirical evidence provided here, it is simply implausible that voters will be able to accurately assess their performance, aggregate that information, and rationally decide whether to punish the incumbent. Second, the ballot box does not offer the level of specificity voters need to communicate their preferences about the performance of agencies. For example, if Commerce, Defense, and Energy perform well, but Labor, State, and Transportation falter, how does a voter communicate that assessment at the ballot box? Third, agencies are largely non-political, so by design do not respond to voters.

Derivative accountability, then, must be significantly revised if it is to offer a compelling theory of accountability. In Chapter Five, we consider some institutional design fixes that will help overcome the problems identified in Chapters Two and Three. For example, Chapter Five proposes that the Office of Information and Regulatory Affairs, housed in the Office of Management and Budget, produce "responsibility worksheets" that specify which institution controlled what aspect of a new policy and produce "coercion worksheets" that specify the real costs to citizens in terms of the coercion they are expected to face from nudges. These worksheets could serve to de-bias individuals' erroneous perceptions of blame.

VIII. Roadmap

Chapters 2 and 3 represent three efforts at testing rational choice-based theories of electoral accountability through randomized control trials. These studies take aim at concepts that are crucial to our understanding of a central feature driving retrospective evaluation of policy: blame. In study after study, blame has been shown to drive more political action than praise. Chapters 2 and 3 test how our evolving political system impacts blameworthiness judgments. In Chapter 2, I analyze how delegating policymaking authority changes blameworthiness judgments. And in Chapter 3, I analyze how soft paternalism, or nudging, changes assessments of blame.

Chapter 4 operates at a more theoretical level, but nonetheless retains its behavioral approach. Traditional electoral accountability depends on a rigid legal formalism. It assumes that, when laws change, behavior automatically changes. Chapter 4 rejects this formalist thinking. Instead, it offers an approach to law - and, in particular, legal change - that takes a more realistic, cognitive approach. This is a crucial development for studies of electoral accountability, since legal change *the* payoff of political accountability. Understanding when it works – and when it doesn't – is paramount.

Chapter 5 offers institutional fixes to the problems identified in Chapters 2, 3, and 4. The primary result of Chapters 2 and 3 is that delegation and new legal tools, such as soft paternalism, inhibit individuals' ability to hold officials accountable for failed policies. The model presented in Chapter 5 accounts for these problems, and makes institutional recommendations to remedy them.

CHAPTER TWO: DOES DELEGATION UNDERMINE ACCOUNTABILITY?

I. The significance of political blame for electoral accountability

The foundational premise of the traditional model of electoral accountability is that voters will hold elected officials accountable for failed policies. If incumbents do not act in accordance with voters' preferences, then voters hold them accountable at the ballot box. This dynamic is crucial for the traditional model, because it explains how voter preferences influence policy. In sanctioning officials, voters are able to exert some degree of control over their behavior.

In order to hold officials accountable for failed policies, voter must be able to assess officials' performance. There is no other basis to ground voters' judgments. This Chapter shows how a pervasive feature of contemporary politics – delegation to administrative agencies – undermines voters' ability to assess officials' performance. With the demise of the non-delegation doctrine, **who** wields political power in the U.S. (and across the West) has shifted from elected officials to appointed bureaucrats. In this Chapter, I show that this shift, while justified in terms of policymaking efficiency, decreases the ability of voters to assess the performance of officials.

The presence of delegation skews voters' ability to assign blame. For example, in situations where a principal (elected official) delegates to an agent (bureaucrat), voters under-assign blame to the principal. In fact, even where the agent is totally constrained, and possesses no autonomy over the policy, voters still assign blame to the agent for failed policies.

The implications of these findings for the traditional model are significant. If delegation skews voters' beliefs about blame, then voters will be less adept at holding elected officials accountable. In fact, in many cases, their assessments will systematically under-blame elected officials. This, in turn, means that voters' ability to sanction elected officials is seriously diminished by delegation. Delegation, in other words, critically undermines the foundational premise of the traditional model of electoral accountability.

II. Summary of empirical results

To test how delegation affects voter performance at assigning blame, this Chapter reports the findings of three rounds of experiments. The first round of experiments reported here tests the claim that delegation *indirectly* affects voter behavior by raising the costs of information about official performance. Experiments 1a and 1b test whether the observability of certain types of information impacts blame attributions. These experiments provide evidence of whether voters tend to allocate more blame to legislatures as the quantity of observed information increases. The second round tests whether delegation *directly* affects voter behavior. A

significant literature holds that, if voters[1] possess complete information as to the respective contributions of the legislature and agency, they will accurately apportion blame. However, this Chapter shows that delegation reduces the blame principals face even in cases where participants possess complete information. Whereas the first two rounds of experiments test the blame faced by legislatures, the third round of experiments tests how delegation affects the blame faced by agencies.

The results presented here show that in the absence of full information, voters will tend to shift blame away from legislatures. (Another way to think about this finding is that participants resolve ambiguity in favor of principals.) In short, information asymmetries indirectly affect blame. Strikingly, however, participants tend to under-assign blame to legislatures even when they are presented with a complete description of the respective actions of the legislature and agency. In other words, participants' tendency to blame legislatures does not seem to depend entirely on voters being "imperfectly informed" of the respective contributions of the legislature and agency, as a significant prior literature argued.[2] In other words, delegation impacts voter behavior directly, as well as indirectly.

Another way to explain these results is that voters' evaluations of policy are not independent of the institutional structures that produce the policy. Rather, delegations of policy-making authority systematically induce voters to under-assign blame to legislatures. Legislatures face less blame even in cases where they delegate to non-autonomous, "instrument" agencies. In effect, legislatures can delegate the final, public-facing component of a policy-making process to an agency in order to soften the blame they face.

The traditional model is simply incompatible with these empirical findings. It assumes that voters' evaluations of policy *are* independent of the institutional structures that produce the policy. This is more than a theoretical problem with the model. In the real world, institutions that produce policy have changed. That change is captured in the experiments here. The results of these experiments, to the extent that they generalize to the real world, imply that voters are under-blaming incumbents. If voters are under-blaming incumbents, then the sanctioning function of elections is not functioning well.

III. *Framing the research question of this Chapter*

Can principals delegate the authority to take immoral or antisocial actions to their agents in order to evade responsibility for the consequences of the decision? A burgeoning line of research in experimental economics suggests that principals can shift responsibility to their agents, even where the principal exerts significant control over the agent's decision.[3] One of the

---

[1] Throughout the discussion of my empirical findings, I will use the terms "voters," "participants," and "third party observers" more or less interchangeably. To avoid ambiguity, I will make clear if I am using "voters" to refer to real-world electorates.

[2] See, e.g., Matthew C. Stephenson, Legislative Allocation of Delegated Power: Uncertainty, Risk, and the Choice Between Agencies and Courts, 119 HARV. L. REV. 1035, 1044 (2006).

[3] Hamman et al., Self-Interest Through Delegation: An Additional Rationale for the Principal-Agent Relationship, 100 AM. ECON. REV. 1826 (2010); Lucas Coffman, Intermediation Reduces Punishment (and Reward), 3 AM. ECON. J. 77(2011); Bartling and Fischbacher, Shifting the Blame: On Delegation and Responsibility, 79 REV. ECON. ST. 67 (2012).

main results, for example, shows that in the context of a dictator game, principals face less punishment by delegating a decision that produces inequitable payoffs to an agent, even if the agent is powerless to choose the equitable outcome.[4] Thus, on the basis of this literature it appears that principals can, in fact, employ agents to evade responsibility.

However, these dictator games measure immoral or antisocial behavior only indirectly. These studies interpret inequitable allocations in a dictator as immoral or antisocial behavior, and they interpret third-party decisions to deduct money from the agents' allocations as "punishment."[5] Both methodological decisions pose difficulties. These studies provide no evidence that participants in fact construe inequitable payouts as immoral or antisocial. Indeed, Sanjiv Erat suggests that inequitable payouts might be straightforwardly prosocial. While not all of the studies make available their instructions to participants, those that do make the instructions available do not characterize inequitable payouts as immoral or antisocial. Although researchers versed in, for example, public goods literature might be strongly inclined to interpret inequitable results as blameworthy or antisocial,[6] one would need further evidence to conclude that the players themselves take inequitable payouts as blameworthy. In short, the experimental economic games that have thus far driven this burgeoning literature have failed to provide direct measures of players' beliefs that inequitable dictator game distributions are in fact blameworthy.

This Chapter addresses this shortcoming in the literature. It provides the first study of delegation and blame to directly test third-party observers' beliefs about the blameworthiness of the principals' and agents' behaviors. Instead of relying on indirect measures of the immorality or antisocial nature of certain behaviors, the experiments reported here directly measure third-party observers' evaluations of the blameworthiness of the principals' and agents' behavior. Using vignettes, the experiments reported here operationalize the behavior of principals and agents in the context of legislative delegation to administrative agencies. After reading the vignettes, participants in the study rated the blameworthiness of the actors in several treatments, including a treatment in which the principal delegates to a powerless intermediary agent. Even in these cases, where the agent is effectively powerless to change the outcome, participants blame principals significantly less than in cases where the principal brings about the outcome directly.

Accordingly, this Chapter provides the first clean evidence that principals can delegate a blameworthy decision to a powerless intermediary and nonetheless evade responsibility. That is, in addition to producing empirical evidence that directly challenges the foundational premise of the traditional model of electoral accountability, this Chapter also makes strides in the more abstract study of principal-agent relationships.

---

[4] Grossman and Oexl, Delegating to a Powerless Intermediary: Does it Reduce Punishment?, 16 EXPERIMENTAL ECON. 306 (2012).

[5] Erat's own study sought to solve this problem of measuring immoral or antisocial behavior by shifting from a dictator game to a sender-agent-receiver game. In Erat's study, the principal engages in deception. This approach does not solve the problem, however, since deception might well be warranted and in any event we receive no direct evidence of what the participants make of the deception under the circumstances. Erat acknowledges as much: "Whether or not deception is immoral is admittedly a debatable, and much debated, point." Sanjay Erat, *Avoiding Lying: The Case of Delegated Deception*, 93 J. ECON. BEHAVIOR 273, 274 (2013). Yet despite acknowledging that whether deception is immoral is much debated, Erat notes that his paper "assume[s] that people do perceive deception to be an immoral act." *Id*.

[6] Hamman, *supra* note 3.

III.A. Legislative Delegation of Policy-Making Authority

This chapter studies the relationship between delegation and blame in the context of legislative delegation of policy-making authority to administrative agencies. In such situations, we are often confronted with the question of how voters allocate blame when Congress and an agency are jointly involved in creating unpopular policy. Ideally, rational voters with perfect information would attribute blame in proportion to the relative responsibility of each institution.[7] The complexity of contemporary policy making, however, makes certain types of information material to the process of blame attribution highly costly to obtain.[8] Across a wide range of policy areas, including financial regulation,[9] the environment,[10] telecommunications,[11] labor,[12] and international trade,[13] information concerning the relative responsibility of Congress and the agencies in producing policy is effectively unobservable to voters. How do voters allocate blame for unpopular policies in the face of such uncertainty? This article reports experimental evidence designed to help answer this question.

One prominent line of thought, the Clarity Thesis, argues that voters tend to resolve such uncertainty in favor of legislatures.[14] When voters know that an agency was involved in the policy-making process, but do not know the extent of its role, the Clarity Thesis predicts that they will underassign blame to legislatures and overassign blame to agencies. Legislatures that exploit this tendency can pass unpopular polices yet escape blame.[15]

The Clarity Thesis thus highlights one of the core problems voters face as they seek to hold officials accountable: How to apportion comparative blame in the face of incomplete information? Delegation increases "the informational burden on voters,"[16] thus raising the costs of accurately assigning blame to officials for policy failures. Inaccurate assessments of blame for

---

[7] Fred Cutler, *Government Responsibility and Electoral Accountability in Federations*, 34 PUBLIUS: THE J. OF FEDERALISM 19 (2004).

[8] Jacob Gersen, *Unbundled Powers*, 96 VIRGINIA LAW REV. 301 (2010).

[9] Roberta Romano, *Regulating in the Dark*, in REGULATORY BREAKDOWN: THE CRISIS OF CONFIDENCE IN U.S. REGULATION, Coglianese, ed. (2012).

[10] David Schoenbrod, Power Without Responsibility: How Congress Abuses the People Through Delegation (1993).

[11] Douglass Arnold, The Logic of Congressional Action (1990).

[12] Daniel Farber and Philip P. Frickey, LAW AND PUBLIC CHOICE: A CRITICAL INTRODUCTION (1991).

[13] David Epstein and Sharyn O'Halloran, Delegating Powers: A Transaction Cost Politics Approach to Policy Making Under Separate Powers (1999).

[14] Justin Fox and Stuart Jordan, *Delegation and Accountability*, 73 J. OF POLITICS 831 (2011); Giandomenico Majone, *The Regulatory State and its Legitimacy Problems*, 22 WEST EUROPEAN POLITICS 1 (1999); Morris Fiorina, *Legislative Choice of Regulatory Forms: Legal Process or Administrative Process*, 39 PUBLIC CHOICE 33 (1982); Jide Nzelibe and Matthew C. Stephenson, *Complementary Constraints: Separation of Powers, Rational Voting, and Constitutional Design*, 123 HARV. LAW REV. 617 (2010); Gersen, *supra* note 8.

[15] From a principal-agent perspective, voters are the ultimate principals in a democracy. Elected representatives are their agents (Weingast 1984). Administrative agencies are voters' subagents. As a large literature on the social psychology of trust and cooperation shows, individuals tend to be more trusting of those with whom they have closer, more frequent interactions (Tyler 2010). While a large gulf exists between voters and their elected representatives, voters might be more trusting of their representatives, with whom they have close, more frequent contact, and thus resolve factual uncertainty in their favor by shifting blame to agencies. Further, voters might be more trusting of their elected representatives because they exercise more control over them (Ferejohn 1999). Nonetheless, legislatures can leverage this dynamic in order to escape blame for unpopular policies.

[16] Nzelibe and Stephenson, *supra* note 14, at 623.

unpopular policies, in turn, weaken the sanctioning function of retrospective voting.[17] Raising the costs of electoral discipline effectively "insulates [legislators] from political retribution"[18] and overpunishes agencies. Broad delegations of authority—those, for instance, which do not establish detailed standards for administrative action—are thought to offer the most insulation to legislators.[19] Broad delegations can weaken or break the "traceability chain," linking political actors to specific policy choices.[20] While delegation allows legislators to shirk, voters shift their blame to agencies, thereby forcing them to take the political brunt of unpopular decisions. The net effect is that shirk-prone representatives can exploit these tendencies and pursue suboptimal policies with greater frequency.[21] On this view, limiting legislators' ability to "disguise their responsibility"[22] requires the creation of institutional arrangements that allow voters to "observe the decisions made by each institution."[23] Only then will there be "no 'clarity of responsibility' problem."[24]

The results presented here tend to support the Clarity Thesis claim that in the absence of full information, voters will tend to shift blame away from legislatures. Information asymmetries indirectly affect blame. Strikingly, however, participants tend to underassign blame to legislatures even when they are presented with a complete description of the respective actions of the legislature and agency. In other words, participants' tendency to blame legislatures does not seem to depend entirely on voters being "imperfectly informed"[25] of the respective contributions of the legislature and agency, as the Clarity Thesis argues. Delegation appears to impact voter behavior directly, as well as indirectly.

While delegation tends to reduce blame for legislators, its effects on agencies are more variable. Broadly speaking, attributions of agency blame rationally reflect comparative differences in the power and autonomy of agencies. Attributions of agency blame are sensitive to the degree of legislative oversight and control over the agency and the type of policy-making authority that agencies exercise. The less control the legislature exercises over the agency, and the more responsibility for creating the legal norm the agencies possess, the more blame they face. For instance, an agency exercising more expansive authority—exercising the power, that is, to create law or policy—will face more blame than an agency exercising more narrow authority, such as the authority to merely enforce norms created directly by a legislature. While participants' comparative attributions of blame matched the responsibility exercised by the agency, less responsible agencies tend to face more blame than is rationally warranted (i.e., than is predicted by participants' attributions of blame in the case of direct, nondelegated action), while more responsible agencies reliably face less blame than is rationally warranted. While constrained agencies face less blame than autonomous agencies, constrained agencies face more blame than is rationally warranted, and autonomous agencies less blame than is rationally

---

[17] John Ferejohn, *Incumbent Performance and Electoral Control*, 50 PUBLIC CHOICE 5 (1986).
[18] Arnold, *supra* note 11, at 101.
[19] Fiorina, *supra* note 14.
[20] Arnold, *supra* note 11.
[21] Fox and Jordan, *supra* note 14, at 841.
[22] Fiorina, *supra* note 14, at 47.
[23] Nzelibe and Stephenson, *supra* note 14, at 145.
[24] *Id*.
[25] Stephenson, *supra* note 2, at 1044.

warranted. This result suggests that voters might hedge their bets as their knowledge of the secondary agents diminishes.

These results have ramifications for foundational administrative law doctrines. Post-*Chevron* approaches to accountability have sought, in varying degrees, to give "maximum authority to the most politically responsive decision maker," on grounds that doing so "maximizes the responsiveness of policy to majoritarian preferences."[26] Agency heads, of course, are unelected. If, however, some elected office that possesses leverage over agency behavior can be reliably held to account for agency performance, voters might be able to hold agencies derivatively accountable. Agency policy choices that draw the ire of the electorate, on this view, can be expressed through punishment of the President, who is directly accountable to voters. With the demise of other modes of administrative accountability—interest group pluralism, for instance[27]—it is derivative accountability that animates approaches to administrative accountability. The findings reported here, however, suggest that derivative electoral accountability is subject to bias caused by the dynamics of blame associated with delegation. If derivative accountability is thought to operate via the executive branch—that is, if citizens are to blame the president for agency action—the public's evaluation of the policies of agencies may tend, in the case of less powerful agencies, to be inflated, and, in the case of more powerful agencies, to be deflated. In other words, delegation tends to clump attributions of agency blame, which will result in relatively blameless agencies facing undue heat from the public, and relatively blameworthy agencies getting a pass.

If derivative accountability is thought to run through the legislature, the problems are even more straightforward: delegation reduces the blame the principal faces. Delegation tends to artificially inflate the public's opinion of legislative action, relative to a baseline of direct legislative action. Such skewed valuations of agency action sit uneasily alongside the post-*Chevron* "commitment[] to electoral accountability."[28] If, however, we identify the way in which attributions of blame are sensitive to the legal structure governing the relationship between legislature and agency and the type of policy-making authority exercised by the agency, we can begin to consider institutional designs that better accomplish the twin goals of robust agency policy making and democratic accountability. Most optimistically, we can assess the tendencies that push observers to blame more in one direction or another and institutionally harness these tendencies to enhance accountability.

The remainder of this Chapter is organized as follows. Section IV sets out the fundamental features of the Clarity Thesis. Sections V, VI, and VII explain the design of the three experiments reported here and discuss the data. Finally, Section VIII addresses the policy implications of the results and broaches avenues for future research.

IV. The Clarity Thesis

---

[26] Matthew Stephenson, *Optimal Control of the Bureaucracy*, 107 HARV. LAW REV. 53, 56 (2008).

[27] Lisa Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. LAW REV. 461 (2003).

[28] John F. Manning, *Constitutional Structure and Judicial Deference to Agency Interpretations of Agency Rules*, 96 COLUM. LAW REV. 612, 634 (1996).

The Clarity Thesis proceeds in two steps. It first makes a claim about the way in which constitutional structures of government that permit the delegation of policy-making power affect the costs of, and thus voters' access to, information about officials' performance. When institutional arrangements permit delegation, as in the case of separation of powers regimes, and legislators take advantage of that permission, the informational burden on voters increases. Second, the Clarity Thesis makes a claim about how voters perform in low-information environments. When voters are faced with a decision under incomplete information, the clarity thesis predicts that they will tend to under-blame legislatures and over-blame agencies. The result is less electoral accountability. This section explains these two steps in more detail.

IV.A. Delegation's Indirect Impact on Blame

The Clarity Thesis starts with the claim that delegation raises the costs of obtaining information, thus weakening or breaking what Douglas Arnold refers to as the "traceability chain" linking officials to policy outputs.[29] If voters bearing the costs of a policy seek to hold legislators liable, they must perceive the costs, those costs must be associated with a particular law or policy, and the actors behind the policy must be visible. "Weakening the traceability chain is a superb method for protecting legislators from their constituents' wrath for imposing costs on them," Arnold argues.[30]

Constituent wrath, in other words, is lessened when three types of information are unobservable: a policy's political history, its enactment, and its effects. The observability of these three types of information determines a policy's traceability chain. As Nzelibe and Stephenson propose, the information necessary for voters to adequately evaluate politicians is that information concerning "not only the policy choice and the policy outcome, but also the way in which the policy was adopted: which institutions supported it, which opposed it, whether the President acted unilaterally or with congressional support, [and] whether the proposal died due to lack of presidential support or due to congressional opposition."[31] A policy's traceability chain, then, is composed of three distinct stages in the lifecycle of a policy: (1) the political history of a policy's adoption, (2) the official enactment of the law or policy, and (3) the effects of the law or policy.

Delegation, the Clarity Thesis holds, tends to obscure the first two stages. These first two stages are grounded in the quite plausible assumption that voters might, in addition to caring about the effects of policy, be concerned with the way a policy was adopted: the impetus that led officials to adopt it, which institutions supported and opposed it, the reasons why officials supported or opposed the policy, and so forth. Delegation, the claim goes, makes it more costly to obtain this information. It does so for two reasons. First, delegation adds complexity to the process of enacting laws and formulating policy. Second, delegation exacerbates the informational asymmetries inherent in representative government. Elected officials, it is thought, acquire information in the course of their employment that places them at an informational advantage vis-à-vis most voters. When policy-making authority is delegated, voters must work at

---

[29] Arnold, *supra* note 11, at 100.
[30] *Id*.
[31] Nzelibe and Stephenson, *supra* note 14, at 652.

an even greater information disadvantage in assessing the performance of officials.[32] In short, delegation makes it more difficult to figure out which actors supported or opposed any given policy, and their reasons for doing so.

It is worth pointing out how the clarity thesis differs from other claims about voter sophistication. Political scientists have long observed that voters tend to be severely uninformed about many issues. As Elmendorf and Schleicher write: "If there is any well-accepted fact in political science, it is that most voters pay little attention to politics and know little about the basic institutions of government."[33] What additional value, then, does the Clarity Thesis add by pointing out one more source of voter confusion? While "[d]ecades of research show that citizens are often ignorant about politics," more recent work has sought to explain voting behavior as a function of voters' responses to certain "shortcuts" or "cues."[34] Shortcut-based voting does not require intimate knowledge of politics, but, rather, a basic sense of which party was in charge and how their own or society's welfare improved or diminished.[35] As Elmendorf and Schleicher explain, "[s]o long as these voters discern which party is in charge and which is the principal opposition, they can cast a retrospective vote for the governing party or coordinate on an alternative, depending on their sense of local conditions."[36]

The Clarity Thesis challenges the first of these two assumptions. Delegation, the Clarity Thesis argues, decreases voters' ability to assess the source of policy. Such an inability tends to undermine the ordinary ways—shortcuts and cues—in which voters overcome information asymmetries. Delegation has the effect of obscuring the source of policy because it multiplies the number of individuals and institutions involved in the creation of law and policy.[37] Piecing together the history of a policy requires voters to follow the trajectory of a policy across institutions. At minimum, delegation requires voters interested in the history of a policy's adoption to track two institutions: the legislature and the agency to which power was delegated. These two institutions might be responsive to different sets of constituents, have different policy preferences, or understand the relevant facts differently. Such differences might affect the performance of the institutions. Voters faced with assigning responsibility for an unpopular measure must decide whether the legislature drafted an ill-conceived statute or whether the agency poorly implemented a well-designed idea. If an agency poorly implemented a statute, was that implementation scheme due to the agency's misunderstanding of the facts, or was it simply following the lead of the legislature? What is more, a policy's trajectory may not always proceed so linearly. Policies can be adopted in piecemeal fashion, through a series of negotiations between institutions.

---

[32] Gersen, *supra* note 8.

[33] Christopher Elmendorf and David Schleicher, *Districting for a Low-Information Electorate*, 121 YALE L. J. 1846, 1850 (2012).

[34] Oliver James, Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments," 21J. of Pub. Admin. Research & Theory 399, 402 (2010).

[35] Huber et al., Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents, 106 Am. Pol. Sci. Rev. 720 (2012).

[36] Elmendorf and Schleicher, *supra* note 32, at 1853.

[37] Sean Gailmard and John W. Patty, Learning While Governing: Expertise and Accountability in the Executive Branch (2012).

Following a policy's trajectory is particularly difficult because officials typically possess better (although certainly not perfect) information about the state of the world and the consequences of various policy choices than do voters. Democratic government is premised on the fragmentation of authority, which produces fragmentation of knowledge. Even absent delegation, authority is partitioned between at least two entities: the people and those they elect. Each partition of authority—each link in the chain of delegation—creates an information asymmetry. Delegation levers up the asymmetry between the public and officials. It creates new sets of actors with private information, and thus presents voters with second-order informational dilemmas. If an agency poorly implements a policy, for instance, voters must figure out whether the legislature believed that the agency would implement it poorly. If a legislature sends a well-designed regulatory scheme to an agency to implement, and the agency skimps on the resources necessary to implement it, should the voter infer that the legislature knew that the agency would fail to fund the project? Delegation exacerbates the information asymmetries inherent in representative government.

The second type of information important for voters' ability to assign responsibility is information about the policy choice itself. Unlike information about a policy's history, this type of information is public. In the case of legislatures, it is available in the text of the statute and, in the case of agencies, it is available in the Federal Register or through less formal publications, such as the agency's website. Despite the relative ease of accessing this information, the Clarity Thesis observes that delegation does, in fact, increase the quantity of the information that the voter must process in order to accurately attribute responsibility. In particular, voters must determine what the legislature enacted itself, and what it left for the agency to do. If this information becomes costly to obtain, voter performance suffers.

The Clarity Thesis is particularly concerned with delegation's effect on information concerning a policy's enactment or political history. Thus, the first round of experiments varies the observability of these two types of information, and always makes the effects of a policy observable to participants. The Clarity Thesis predicts that participants who cannot observe this information will tend to blame the legislature.

IV.B. Delegation's Direct Impact on Blame

The Clarity Thesis maintains that if voters possess these three types of information necessary to trace a chain of political responsibility—(1) the political history of a policy's adoption, (2) the official enactment of the law or policy, and (3) the effects of the law or policy—then delegation does not impact voter performance. Recent work in experimental economics, however, suggests that delegation affects attributions of blame even in cases where subjects possess full information about the history of an outcome. This literature suggests that the Clarity Thesis explains only part of the relationship between delegation and blame. The remainder of the explanation, this literature suggests, stems from the effect of indirect agency. The second round of experiments reported here tests whether delegation has the type of direct impact on blame suggested by this prior experimental literature. If it does, then the Clarity Thesis will need to be modified accordingly.

Experimental economics studies reveal an asymmetry between individuals' perceptions of direct and indirect action.[38] Norm violators face fewer attributions of blame, and less punishment, when they intermediate their relationship with the victim of the norm violation.[39] Principals that carry out an act via an agent tend to face less blame in the context of economic games, like the dictator and ultimatum games,[40] and in the context of market-driven business transactions.[41] This result holds even where the agent's interests are aligned with the principal's interests.[42] Using vignettes, Bazerman et al. show that firms that outsource to a business partner face less blame than do firms that produce the same outcome directly. In the context of market mechanisms, attributions of blame appear to be sensitive to the process or structure that produces an outcome.[43] Notably, these results do not depend on hiding or otherwise obscuring information from subjects. In the dictator and ultimatum games testing the effects of delegation on blame, subjects observe the incentive structure facing each participant and the actions that they take. Likewise, in the vignette-based experiments, subjects are informed of the details of the principal-agent relationship and the incentives facing each actor. In short, this literature suggests that even when individuals possess all three types of information that the clarity thesis identifies as necessary and sufficient to correctly allocate blame, delegation may nonetheless induce voters to underblame legislatures and overblame agencies.

The second round of experiments tests whether delegation directly impacts blame. The experiments reported here present participants with precisely the three types of information that the clarity thesis predicts are necessary and sufficient to produce accurate assessments of blame. In other words, participants read exactly the sort of information the absence of which is thought to lead to inaccurate assignments of responsibility. If, after presenting participants with this information, we nonetheless find inaccurate assignments of responsibility, one cannot reject the clarity thesis argument as to the indirect effects of delegation. Indeed, the experiments reported in Round 1 support this argument. However, such results would tend to support the claim that delegation might trigger some cognitive bias, which tends to result in participants underblaming the legislature and overblaming the agency.

V. Round 1: Does Delegation Indirectly Impact Legislative Blame?

The first round of experiments reported here tests the claim that delegation indirectly induces voters to overblame legislatures. The Clarity Thesis holds that delegation indirectly affects voter behavior by raising the costs of information, and thus increases the uncertainty as to the relative responsibility of the legislature and agency. Voters, in response to this low-information environment, tend to resolve the ambiguity in favor of the legislature.

---

[38] Coffman, *supra* note 3.

[39] Bazerman, et al., *Dirty Work, Clean Hands: The Moral Psychology of Indirect Agency*, 109 ORG. BEHAVIOR & HUM. DECISION PROCESSES 134 (2009).

[40] Bartling & Fischbacher, *supra* note 3.

[41] Bazerman, *supra* note 38.

[42] Chaim Fershtman and Uri Gneezy, *Strategic Delegation: An Experiment*, 32 RAND J. of ECON. 352 (2001).

[43] Andrew T. Hayashi, Occasionally Libertarian: Experimental Evidence of Self-Serving Omission Bias, 29 J. of Law, Economics, & Organization 711 (2013).

Experiments 1a and 1b test whether providing respondents with information concerning (1) the role of the legislature and agency in the policy's political history, (2) the institution responsible for the policy's formal enactment, and (3) the effects of the policy on the electorate impacts blame attributions. Does varying each of these three types of information change the quantity of blame participants allocate to the legislature?

Experiments 1a and 1b employ vignettes that tell the story of a jurisdiction's response to a serious problem—a natural disaster, for instance—in concise (350 words or less), nontechnical language. To measure the effect of information on the quantity of blame participants allocate to the legislature, each participant was randomly assigned to one of four conditions. To establish a baseline level of blame, one group of participants was assigned to a direct condition. This group read a vignette describing a wholly legislative response, which did not involve an agency, to the problem. The remaining participants were assigned to one of three indirect conditions. In the indirect conditions, the legislature utilizes an agency to carry out the response. The three indirect conditions, however, contain varying quantities of information about the agency's role in the response. In the first indirect condition, participants read a vignette that contains only information about the effect of the response. The respective roles of the legislature and agency are entirely opaque to the participant. In the second indirect condition, participants read a vignette that contains information about the effect of the response and history of the official enactment. In this vignette, the participant is aware not only of the effects, but also of which institution—the legislature or the agency—made the decision to enact the policy. Finally, in the third indirect condition, participants read a vignette that contains information about the effect of the response, the official enactment, and the policy's political history. Only this third indirect condition contains the full quantity of information the clarity thesis argues is necessary and sufficient to accurately assess blame.

V.A. Experiment 1a

Experiment 1a was administered online to 248 participants, recruited and compensated via Amazon's Mechanical Turk web services. All participants completed two demographic questions (41 percent female, M_{age} = 33) and were randomly assigned to one of the four conditions described above: direct legislative action, $indirect_{effects}$ legislative action, $indirect_{eff+enact}$ legislative action, or $indirect_{all}$ legislative action.

To test the clarity thesis claim that delegation indirectly affects blame, Experiment 1a compares the blame scores in each of the three indirect conditions with the baseline blame score established by the direct condition. In the indirect conditions, the legislature employs a "mere tool" or "instrument" agency.[44] It thus seeks to eliminate any effects resulting from factors such as perceived lack of foreknowledge of the agency's intentions or lack of control over the agency. To further emphasize the fact that the agency is a mere tool of the legislature, the vignettes in Experiments 1a and 1b are structured such that the legislature makes a choice that renders the agency able to act in precisely one way. More concretely, in both experiments' indirect conditions, the legislature promises to convey funds to address some social problem, but then reneges on that promise and channels the funds elsewhere. The agency, whose job it was to distribute those funds, is thus left with no option but to fail to address the social problem, since it

---

[44] Christopher Kutz, Complicity: Ethics and Law for a Collective Age 153 (2007).

has no funds with which to work. The resulting vignette is one in which the actions available to the mere instrument agency are completely determined by the legislature. The agency must take the unpopular action. As a result of the legislature's broken promise, "large parts of the city remain devastated" and "[s]ignificant numbers of residents still lack basic services like electricity and hot water." Participants were then asked: "On a scale of 1 (not blameworthy at all)—10 (very blameworthy), how blameworthy do you think the legislature's action was?"

The Clarity Thesis predicts that as participants' information concerning the respective roles of the legislature and agency increases, legislative blame scores will rise. Since participants in the direct condition and the indirect$_{all}$ condition possess complete information, we should see equivalent levels of blame in these two conditions. Participants assigned to the indirect$_{effects}$ condition read a vignette in which the legislature delegates responsibility for the disaster response to an agency.

The legislature then delegated responsibility for managing the clean-up efforts to SQL, an official government agency with expertise in handling emergencies. The legislature felt comfortable delegating responsibility for this important task to SQL because it exerts high levels of control and oversight over the agency. SQL's primary job was distributing the money that the legislature promised to the city. SQL, however, never distributed the money. As a result, large parts of the city remain devastated. Significant numbers of residents still lack basic services like electricity and hot water.

Participants assigned to this condition have access to information about the institutions involved in the policy, and the effects of that policy, but no information concerning how or why the policy was established. In other words, the indirect$_{effects}$ condition omits information concerning the two types of information the clarity thesis deems crucial to accurately allocating blame.

By comparison, participants assigned to the indirect$_{eff+enact}$ condition read a vignette identical to the indirect$_{effects}$ condition, except that it contained information concerning which institution—the legislature or the agency—had, in fact, enacted the measure. In the indirect$_{eff+enact}$ condition, participants read that "[a]fter consulting with SQL, however, the legislature enacted a bill re-directing the funds." Thus, in the indirect$_{eff+enact}$ condition, participants possess one of the two types of information the clarity thesis deems crucial to accurately allocate blame.

Finally, in the indirect$_{all}$ condition, participants read a vignette that contained all the information in the indirect$_{eff+enact}$ condition, plus information concerning the political history of the policy. The political history is identical to the history set out in the direct condition: the legislature believed it was in its electoral interests to distribute the clean-up funds elsewhere, so it abandoned the clean-up efforts. Recall that the legislature "exerts high levels of control and oversight over the agency" and, since the legislature made the decision to redirect the funds away from the agency, the agency could not have conveyed the funds to the city. The effect, then, is that in the indirect$_{all}$ condition the agency is a mere instrument of the legislature.

As shown in Figure 1, as the quantity of information increases, so does the accuracy of the blame allocated to the legislature, as compared with the baseline blame score established in the direct condition ($M_{Direct} = 8.929$, SD = 1.455). Participants who read the indirect$_{effects}$ condition ($M_{effects} = 5.86$, SD = 2.75) rated the legislature as significantly less blameworthy than

participants who read the indirect$_{\text{eff+enact}}$ condition (M$_{\text{eff+enact}}$ = 7.72, SD = 2.24). Adding the third type of information in the indirect$_{\text{all}}$ condition (M$_{\text{all}}$ = 8.26, SD = 2.51) further raises blame scores. Analyzing Experiment 1a using a Welch two-sample *t* test reveals significant differences between the direct condition and the indirect$_{\text{effects}}$ condition (t = 7.759, p = < 0.001), as well as between the indirect$_{\text{effects}}$ condition and the indirect$_{\text{eff+enact}}$ condition (t = 4.139 p = < 0.001).

Thus, consistent with the Clarity Thesis, these results support the claim that participants' accuracy improves as more information is revealed. Low-information environments tend to advantage legislatures in cases where the legislature deployed a mere instrument agency to deflect blame.

V.B. Experiment 1b

Experiment 1b employs the same structure as Experiment 1a. The legislature promises to commit funds to a social problem, but then reneges on that reason for selfish reasons. The agency is thus left with no option but to fail to address the social problem, since it has no funds with which to work. Experiment 1b was administered online to 212 individuals, recruited and compensated through Amazon's Mechanical Turk. After completing two demographic questions (45 percent female, M$_{\text{age}}$ = 32), all subjects read a scenario describing how "[l]ast year on election day, a malicious computer virus sabotaged the voting machines in jurisdiction B. As a result, the legislature promised to overhaul the voting machines by next year's elections." As in Experiment 1a, participants assigned to the direct condition read a scenario describing a wholly legislative response.

After considering the matter more carefully, however, the legislature realized that overhauling the voting machines could open the door to overhauling other aspects of the electoral process. This prospect worried the incumbent legislators, since they knew how to win in the current electoral environment. Thus, rather than overhaul the voting machines, the legislature decided to use the money for other purposes, including studies showing how well other aspects of the electoral system worked. The legislature enacted a bill diverting the voting machine money to these other purposes.

As a result, the voting machines remain susceptible to computer viruses. The compromised integrity of the voting machines has induced many to lose faith in the quality of the elections in jurisdiction B.

The remaining participants were randomly assigned to one of three indirect conditions, corresponding to the three indirect conditions in Experiment 1a. In the indirect$_{\text{effects}}$ condition, participants read a vignette that stated that the legislature "delegated responsibility for managing the overhaul to PPW, an official government agency with expertise in elections." Because the indirect$_{\text{effects}}$ condition hides information concerning the policy's enactment and political history, participants in this condition know that "PPW, however, never purchased or installed the new machines," but do not know why this is the case. In the indirect$_{\text{eff+enact}}$ condition, participants read a vignette that explains that "the legislature enacted a bill diverting the funds." Finally, in the indirect$_{\text{all}}$ condition, participants read a vignette that explains the political history leading to the legislature's decision, as set out in the baseline condition.

As in Experiment 1a, using a Welch two-sample t test, participants who read the direct scenario (M_{Direct} = 8.96, SD = 1.73) found the legislature significantly more blameworthy (t = 6.02, p = < 0.001) than those who read the indirect$_{effects}$ condition (M_{Direct} = 6.4, SD = 2.54). Blame continued to rise as participants read vignettes with more information (Indirect$_{eff+enact}$ = 7.74, SD = 2.24; Indirect$_{all}$ = 8.6, SD = 2.13).

Figure 1: Impact of increasing information on attributions of blame.
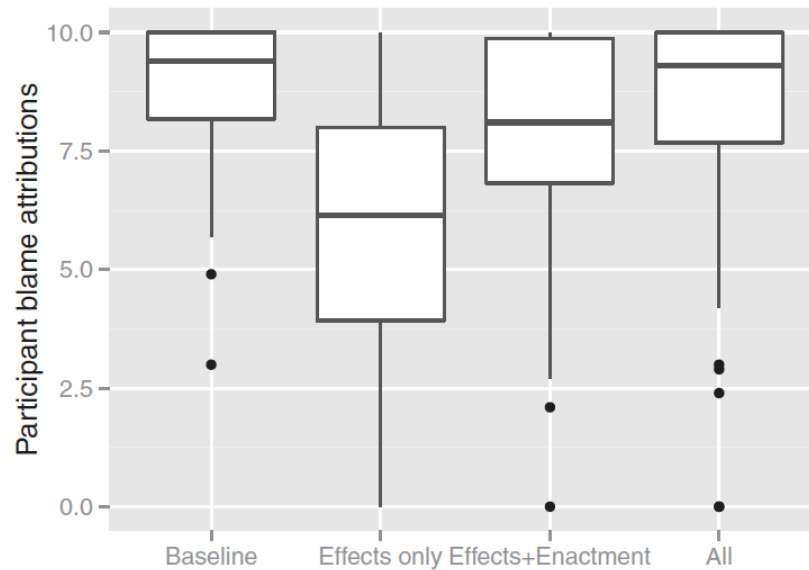
NOTE: Attributions of blame increase as participants gain more information about the history of the policy-making process.

V.C. Round 1 Summary

The experiments reported here suggest that delegation indirectly affects blame, as predicted by the Clarity Thesis. The experiments in Round 1 test whether participants evaluating the indirect actions of a legislature more accurately assess blame when they possess more information concerning a policy's enactment and political history. As predicted by the Clarity Thesis, participants tended to resolve informational ambiguities in favor of the legislature. Delegation, then, would appear to benefit the legislature in two respects. First, delegation tends to benefit legislatures to the extent that they are primarily responsible for unpopular policy. Even where legislatures do not seek to shift the blame to the agency, voters will tend to treat unobservable information showing the legislature should shoulder responsibility as information implicating the agency. One need not attribute any failure of good faith by voters in such a situation; rather, this tendency to resolve informational ambiguities in favor of the legislature might result from a tendency by voters (i.e., the political principals in a democracy) to favor an agent that they selected, rather than agents that they did not select. Second, where legislatures do seek to shift the blame for an unpopular policy, they can accomplish this feat, at least in part, simply by introducing an agent into the process. If delegating to an agent tends to lower the

visibility of the respective causal roles, a legislature seeking to shift the blame need not publicly exaggerate the agency's role; voters will do that.

VI. Round 2: Does Delegation Directly Impact Legislative Blame?

In Round 1, participants performed best when they possessed information concerning the respective roles of the legislature and agency. The stronger the traceability chain, the better the performance. Even with the advantage of the strongest traceability chain, however (i.e., in the indirect$_{all}$ condition), participants on average rated the legislature acting through a mere instrument agency as less blameworthy than a legislature acting directly. These data fit uncomfortably with the Clarity Thesis. While the results in Round 1 support the Clarity Thesis' claim that voters tend to resolve informational ambiguities in favor of the legislature, those results do not support the claim that full information results in accurate assessments of blame. This position aligns with prior work that shows that blame evaluations are a function of the degree of control and foresight an individual possesses (Cushman et al. 2011). Nonetheless, even with full information, participants on average rated the legislature acting through a mere instrument agency as less blameworthy than a legislature acting directly. Even where the legislature possessed a high degree of control over the agency, and foresaw that the agency would comply with its preferred policy position, participants nonetheless assigned it less blame than they assigned to a legislature enacting the same policy directly.

The second round of experiments tests whether, consistent with the results on delegation and blame in the experimental economics and moral psychology literatures, delegation reduces attributions of blame even in those cases where subjects possess information about the way in which a policy was adopted. To do so, it employs vignettes that set out, in concise (350 words or less), plain language the history of a policy jointly produced by a legislature and administrative agency. Each vignette tells the story of some problem the voters of a jurisdiction faced—increased pharmaceutical costs, for instance—and the way the legislature and agency addressed that problem. In so doing, the narrative explains "the way in which the policy was adopted" and the policy choices made by the legislature and agency (Knutson 2010). Because the vignettes are clear, accessible, and present precisely the sort of facts that the clarity thesis argues are determinative of voter behavior, these experiments provide clean tests of the clarity thesis claim that delegation primarily affects blame by creating informational burdens.

Experiments 2a and 2b each modify the structure of the experiments in Round 1 in order to further test whether delegation directly impacts blame. Experiment 2a seeks to nudge participants toward more "reason-based choice" by presenting them with both direct and indirect scenarios.[45] Whereas Experiments 1a and 1b employ "separate evaluation"—a between-subjects design in which participants assessing the blameworthiness of indirect legislative action do not also assess the blameworthiness of direct legislative action Experiment 2a employs "direct evaluation," in which participants read both indirect and direct vignettes. By presenting participants with both types of vignettes, Experiment 2a might nudge participants toward treating the direct and indirect conditions more similarly. Experiment 2b addresses a potential objection, namely, that even the indirect$_{all}$ conditions still lack information relevant to tracing a chain of

---

[45] Christopher Hsee and Jiao Zhang, *General Evaluability Theory*, 5 PERSPECTIVES ON PSYCHOLOGICAL SCIENCE 343 (2010).

responsibility for the policy. Thus, Experiment 2b provides even more information to participants than is contained in the indirect$_{all}$ conditions in Experiments 1a and 1b.

If, after modifying the structure of the experiments in Round 2, participants treat direct and indirect legislative action as equivalently blameworthy, then one might explain the difference between the blame scores in Round 1's indirect$_{all}$ conditions and Round 1's direct conditions as an artifact of the experimental design. If, however, voters find that the indirect action type is less blameworthy even with complete information as to the history of the policy's adoption, then it would support the claim that delegation, in addition to creating informational burdens, directly affects voter behavior.

A. Experiment 2a

Experiment 2a employs a version of what Hsee and Zhang refer to as "joint evaluation."[46] In experiments employing joint evaluation, "multiple alternatives are simultaneously available for consideration [by participants], while under separate evaluation only one alternative is considered."[47] The finding underlying this design is that joint evaluation generally "allows for the consideration of attributes that are difficult to assess in isolation and is therefore more likely to result in reason-based choice . . ."[48] Thus, Experiment 2a employs a within-subjects design, where all participants first read the direct condition, then read the indirect condition.

In the direct condition, participants were asked to attribute blame to the legislature. In the indirect condition, participants were asked to assess the blameworthiness of both the legislature and the agency. This design might elicit a more reasoned comparative assessment of the legislature versus the mere instrument agent than the separate evaluation featured in Round 1.

Experiment 2a was administered online to 142 individuals, recruited and compensated through Amazon's Mechanical Turk. After completing two demographic questions (47 percent female, $M_{age} = 36.9$), all subjects read a scenario describing how "hackers unleashed a malicious virus on the computerized trading systems used by large companies in several jurisdictions," causing great harm to their respective financial systems. "In response, the legislature of each jurisdiction passed a bill requiring computerized trading systems to comply with a new set of security measures," which were "very expensive to implement." To encourage rapid adoption, legislatures promised to reimburse companies that implemented the new system within six months. Each jurisdiction, however, "chose different ways of enforcing the law." All subjects were presented with accounts of how the enforcement process proceeded in two of these jurisdictions. In the direct condition (State A), the legislature reneged on its promise to reimburse the companies and, as a result, "many companies went out of business, and those that had not yet implemented the costly measure decided not to do so." All subjects were then asked: "On a scale of 1 (not blameworthy at all)—10 (very blameworthy), how blameworthy do you think the legislature's action was?" In the indirect condition, the legislature "delegated the task of enforcing the law to SECRA, which is an official administrative agency, created by the

---

[46] *Id.*

[47] Bazerman, et al., *supra* note 40, at 2.

[48] *Id.*

legislature to serve as its agent." Subjects were then told that the legislature "exercises a relatively high degree of control over" the agency. In the indirect condition, "the legislature did not make the funds available to SECRA," leaving the agency "unable to reimburse the companies for the upgrades they made." As a result, the same consequences followed. All subjects were then asked: "On a scale of 1 (not blameworthy at all)—10 (very blameworthy), how blameworthy do you think SECRA's action was? How blameworthy was the legislature's delegation to SECRA?"

Experiment 2 was analyzed using a Welch two-sample paired t test. Participants rated the direct scenario (M_{Direct} = 8.51, SD = 1.81) significantly more blameworthy than the indirect scenario ($M_{Indirect}$ = 6.70), SD = 3.37). The results are significant (t = 6.45, p = < 0.001). Thus, as Figure 2 shows, even where the agency's hands are tied by the legislature's policy decisions, and where participants are able to jointly evaluate the direct and indirect acts, delegation appears to directly impact attributions of blame.

Together, Experiments 1a, 1b, and 2a challenge the clarity thesis claim that information is the exclusive way in which delegation impacts voter behavior. While delegation might obscure information voters would find relevant, policies that result from delegated arrangements appear to diminish the blame legislatures face even when voters are presented with the history of a policy's adoption and the relevant choices made by the legislature and agency. Although counterintuitive, these results accord with prior studies that reveal that delegation diminishes blame in economic games and business transactions.

Figure 2: Comparing blame scores where legislature acts directly versus delegates to powerless intermediary.
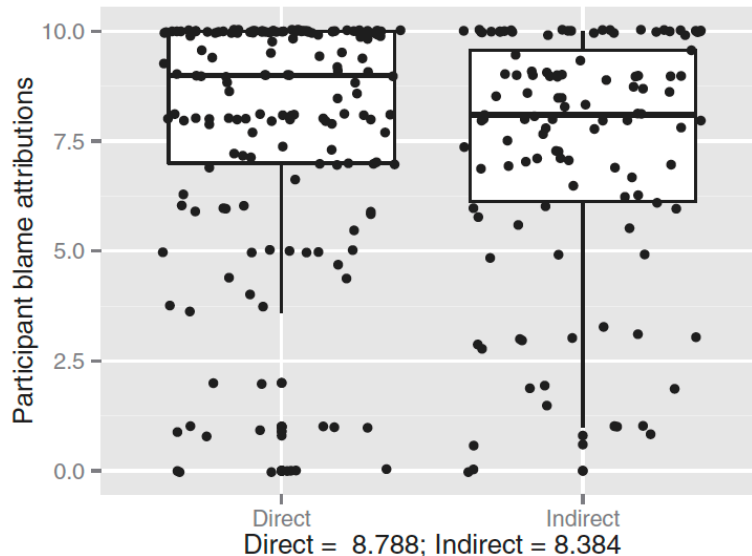
NOTE: Participants rate legislatures that act directly as more blameworthy than legislatures that act indirectly through a powerless agency.

VI.B. Experiment 2b

One might object to the claim that delegation directly influences blame by arguing that the vignettes in Experiments 1a, 1b, and 2a omit material information about the degree of control and foresight that the legislature possesses over the agency. Such an omission might induce participants to resolve informational ambiguities in favor of the legislature. As a check against this concern, Experiment 2b layers more information on top of the information contained in the indirect$_{all}$ hurricane disaster vignette employed in Experiment 1a.

Experiment 2b's indirect$_{plus}$ condition adds information concerning the nature of the legislature's power over the agency, as well as information concerning the agency's role in the unpopular policy decision. Using a new round of 85 Mechanical Turk participants, Experiment 2b compares participants' blame scores in the indirect$_{all}$ condition with the indirect$_{plus}$ condition. The indirect$_{plus}$ vignette reads in relevant part as follows, with the added information italicized.

Last month, a terrible hurricane ravaged a large city on the western coast of jurisdiction B. In the immediate aftermath of the storm, the legislature promised the city to pay for all of its clean-up costs.

The legislature then delegated responsibility for managing the clean-up efforts to SQL, an official government agency with expertise in handling emergencies. SQL's primary job was distributing the money that the legislature promised to the city. The legislature felt comfortable delegating responsibility for this important task to SQL because it exerts high levels of control and oversight over the agency. *For instance, the legislature can remove the head of SQL at will. Further, unwritten conventions empowered the legislature to exert control over the agency's use of discretion.*

After considering the matter more carefully, however, the legislature realized that the residents of the city were unlikely to ever make substantial donations to their future re-election campaigns. "Wouldn't it be better," the legislators reasoned, "to direct the funds meant for the disaster relief to other voters more likely to help us in the future?" *SQL told the legislature that doing so would significantly hamper the city's recovery, but it was ineffective in persuading the legislature.*

Thus, the legislature enacted a bill re-directing the funds. As a result, large parts of the city remain devastated. Significant numbers of residents still lack basic services like electricity and hot water.

Participants receiving the indirect$_{plus}$ condition read a vignette that contains all the information contained in the indirect$_{all}$ vignette, plus more information concerning the legislature's control over the agency and the agency's role in formulating the unpopular policy. The added information further demonstrates the agency's lack of control, and portrays the agency as opposing the unpopular policy.

Despite the added information, participants reading the indirect$_{plus}$ vignette did not find the legislature more blameworthy. In fact, participants reading the indirect$_{plus}$ vignette found the legislature, on average, slightly less blameworthy (indirect$_{plus}$ = 8.6, indirect$_{all}$ = 8.20). The difference, however, is not statistically significant (t = 1.25, p = 0.21). Experiment 2b, then, suggests that adding more information does not alter participants' attributions of blame.

In sum, Experiments 1a, 1b, and 2a suggest that policies that result from delegated arrangements appear to diminish the blame legislatures face even when voters are presented with the history of a policy's adoption and the relevant choices made by the legislature and agency. While counterintuitive, these results accord with prior studies that "revealed a moral preference for indirect agency" under certain conditions.[49] Experiment 2b addresses one possible objection to these results, namely, that the vignettes lack crucial information concerning the extent of legislative control and foresight. In response to this objection, Experiment 2b layers more information concerning legislative control and foresight on the hurricane response vignette. Blame scores, however, do not rise when this new information is included. Consistent with prior research on cue taking and information overload, this result suggests that the quantity of information contained in the indirect$_{all}$ condition is sufficient for participants to trace a chain of responsibility for the policy.

VII. Round 3: Does Delegation Impact Agency Blame?

The first two rounds of experiments suggest that delegation both directly and indirectly impacts attributions of legislative blame. Proponents of the clarity thesis sometimes argue that this tendency to soften legislative blame is an unalloyed good for legislators. Legislatures might well care about how much collateral damage agencies will face, however. Even if legislatures are interested in the fates of agencies only in order to take the blame for unpopular policies, they still need legislatures to possess some popular legitimacy. Less cynically, legislatures rely on agencies to accomplish things, even if they need them to take the fall in any particular instance. Agencies are, over the long run, valuable partners. If agencies are to manage blame, they need to have an idea of how blame avoidance will impact their partners' fate. Thus, we need to know how a blame-shifting strategy affects agencies.

The results considered thus far suggest that the Clarity Thesis is correct in arguing that delegation reduces attributions of legislative blame, but incorrect in suggesting that the exclusive or primary causal mechanism underlying the effect is informational. Rather, legislative blame declines even where information is complete. But what of agency blame? Fiorina argues that delegation allows legislators to exploit informational asymmetries in order to "disguise their responsibility for the consequences of the decisions ultimately made" and shift blame to agencies.[50] More recently, Justin Fox and Stuart Jordan have set out the necessary preconditions for blame shifting, one of which is that "politicians must have more information than voters about the actions the bureaucracy is likely to take with any authority delegated."[51] When these preconditions exist, legislators can strategically delegate to agencies in order to shift blame for unpopular policies. In sum, misattributions of agency blame, according to the Clarity Thesis, tend to result in overblaming agencies. Agency blame, on this view, is a one-way ratchet.

---

[49] Bazerman, et al., *supra* note 40, at 5.
[50] Fiorina, *supra* note 14, at 47.
[51] Fox and Jordan, *supra* note 11, at 834.

The third round of experiments tests whether misattributions of agency blame might occur absent incomplete information, and whether the misattributions always involve overblaming agencies. Further, it tests these claims under a variety of institutional arrangements. The Clarity Thesis typically fails to specify two crucial features of the legislature-agency relationship: (1) how much oversight or control the legislature possesses over the agency, and (2) how much policy-making authority the agency possesses. Round 3 varies these conditions in order to gauge their influence on blame. Doing so allows us to address questions such as whether voters are more punitive when legislatures delegate more power to agencies. More generally, introducing these two conditions helps us draw a larger picture of the way diverse institutional arrangements affect attributions of blame.

Figure 3: Agency blame: difference between direct and indirect conditions.
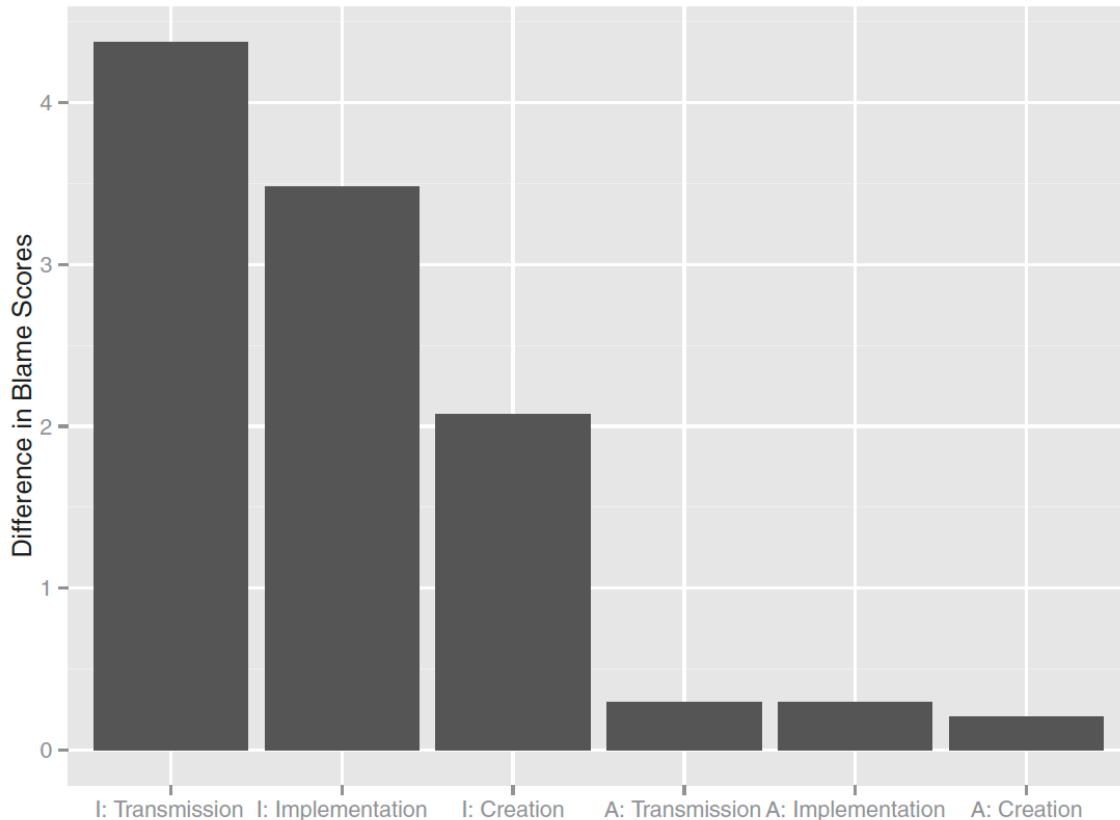
NOTE: Y-axis represents the difference (in all cases the agency receives less blame) between direct legislative condition and indirect agency condition. The first three differences (preceded by an "I") represent cases of limited agency authority, while the latter three cases (preceded by an "A") represent more expansive agency authority. Larger numbers represent lesser agency blame.

VII.A. Legislative Control Over Agencies

Political scientists have long debated the question of how legislatures best constrain agencies.[52] One, if not the primary, way in which legislatures exert control over agencies is through oversight.[53] Schwartz and McCubbins distinguish, classically, between two modes of agency oversight: police patrols and fire alarms. Legislatures engage in police patrol oversight when they actively monitor agency activity. Active monitoring, however, is costly, and so they may become involved only when the agency draws the ire of some third-party constituency. Only then does the legislature seek to put out the agency's fire.

Experiment 3 incorporates a mix of police-patrol and fire-alarm style oversight. Unlike the experiments in Rounds 1 and 2, which tested only instrument agencies, Round 3 includes autonomous agencies. The vignettes state whether the legislature possesses a high or low degree of oversight over the agency. They then illustrate this oversight with an example. Thus, autonomous agencies are defined by two features in this experiment, corresponding to the two features that define instrument agencies. Legislatures exercise comparatively less control and oversight over autonomous agencies and, in order to illustrate this autonomy to participants, autonomous agencies act against the preferences of the legislature that delegated power to them. Instrument agencies, by contrast, are monitored and comply with legislative preferences.

VII.B. Policy-Making Authority

The Clarity Thesis addresses the role of the breadth of the agency's authority only insofar as it is argued that broad delegations of authority provide the most insulation for legislators. Fiorina, for instance, argues that legislators can provide agencies with broad grants of authority without explicit *Schecter Poultry*-style standards in order to shift responsibility for unpopular decisions to agencies. Schoenbroad argues that congressional legislation that features "attractive abstractions," such as instructions to agencies that they maintain "orderly markets," help legislators avoid blame. Fox and Jordan argue that legislators can avoid blame by turning over policy decisions wholesale to agencies that they believe share their preferences.

Experiment 3 tests the idea that broad delegation leads to more blame under conditions of full information. All else equal, one might hypothesize that the more policymaking authority an agency possesses, the more blame voters will be willing to assign it. To test this hypothesis, Experiment 3 was structured around three types of vignettes, which correspond to three types of policy-making authority an agency might possess.[54] In the transmission type, the legislature simply delegates the power to enforce or apply norms. Such authority might be seen as "a relatively lower-order activit[y]."[55] In this sense, enforcement authority exists at the bottom of the policy-making hierarchy. In the implementation type, the agency possesses the authority to "translate" norms or standards into regulated practices (Freeman 2000). Finally, at the top of the policy-making hierarchy is the power to create norms. In the creation vignette, the agency possesses the legislative-like authority to author norms. In

---

[52] See, e.g., Arthur Lupia and Mathew D. McCubbins, Learning from Oversight: Fire Alarms and Police *Patrols Reconsidered*, 10 J. of LAW, ECONOMICS, & ORGANIZATION 96 (1994).

[53] Rachel E. Barkow, Insulating Agencies: Avoiding Capture Through Institutional Design, 89 TEXAS L. REV. 15 (2010).

[54] Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543 (2000).

[55] *Id*. at 573.

comparing attributions of blame for these three types of policy-making authority, Experiment 3 tests whether agencies with more power to control the content of a policy receive more blame.

VII.C. Procedure and Results

The study was administered to 312 subjects (42 percent female, Mage = 31.8), recruited and compensated through Amazon's Mechanical Turk. Participants read three scenarios, corresponding to the three types of policy-making authority just mentioned. For each scenario, all participants read a direct condition (legislative action), which established a baseline level of blame, and an indirect condition. Although all participants read the same direct condition, the indirect conditions varied the degree of control the legislature possessed over the agency.

In each scenario, a multijurisdictional problem confronts the legislatures of several states. Participants read how the legislature of State A responded to the problem directly (the baseline scenario), and then read how the legislature of State B responded to the problem indirectly by delegating some authority to an agency. In the direct scenario, the legislature responds to the impetus by enacting a law. In the indirect scenarios, the legislature responds to the impetus by enacting a law and delegating either transmission, implementation, or creation authority to an agency. The acts taken, and the consequences that follow, are the same in the direct and indirect scenarios.

In the transmission scenario, which is substantially similar to the vignette employed in Experiment 1, the impetus for political action is a breakdown of an important brokerage service, which threatens stock markets. The brokerages can be updated with costly repairs, for which the legislature promises to reimburse the companies. In fact, the (legislature or agency) reneges on its promise, and a financial fallout ensues. In the implementation scenario, a pro-environment ideological shift results in heightened pollution regulations, which creates a demand for hands-on regulatory implementation. However, the (legislature or agency) failed to meet with all but the most profitable companies and, as a result, drafted cost-insensitive, ineffective regulations. In the creation scenario, an opportunity to improve coordination in the shipping industry could lower pharmaceutical prices if uniform standards are established. Despite this opportunity, the (legislature or agency) adhered to an ideology of experimentation over uniformity, and prices remained high.

As in Experiment 2, in the instrument condition the legislature exerts high levels of control over the agency and, in the course of the vignette, the agency in fact complies with policy preferences of the legislature, which is intended to reinforce the legislature's control. In the autonomous condition, on the other hand, the legislature is described as possessing relatively little control over the agency and, in the course of the vignette, the agency violates the legislature's policy preferences.

The degree of legislative control over the agency ends up making a significant difference in the blame participants parceled out. In each of the three scenarios, when legislatures delegated the task to an instrument agency, participants blamed the instrument significantly less (transmission: $t = 10.13$, $p = < 0.001$; implementation: $t = 8.78$, $p = < 0.001$; creation: $t = 7.14$, p

$= < 0.001$). As shown in Figure 3, participants found the acts of autonomous agencies to be much more blameworthy than they did instrument agencies.

The divergence in the blame faced by instrument agencies and autonomous agencies suggests that delegation does not uniformly affect agency blame. Much as participants found autonomous agencies more blameworthy than instrument agencies, so, too, did participants find those agencies that were delegated more extensive policy-making authority more blameworthy. As policy-making authority increases, the difference between the baseline scenario and the indirect scenario decreases. In the transmission scenario, the difference between the baseline case and the indirect conditions (the average of the instrument agency blame score and the autonomous agency score) is 2.335. In the implementation scenario, the difference drops to 1.885. In the creation scenario, the difference falls to 1.140. If one defines the error score as the difference between the paired baseline blame score and the agency blame score, then one observes that the error score monotonically decreases, as shown in Figure 4. The declining difference between the baseline scenario blame score and the indirect blame score reflect increased attributions of responsibility to those agencies possessing more policy-making authority. This increased attribution of responsibility is, this Chapter hypothesizes, a result of the increased role the agency possesses in determining the content of the resultant policy.

As a robustness check, Experiment 3 was also conducted using subjects drawn from a population of students and staff at the Experimental Social Science Laboratory at UC-Berkeley. Ninety-six subjects participated (63 percent female, Mage = 26.5). Participants were solicited by laboratory staff members via email and public events calendars. These results follow the same pattern as the results obtained through the Mechanical Turk subject pool and so are not fully reported here. The degree of legislative control over the agency significantly impacts the quantity of blame the agencies face. In the three direct scenarios, in which the legislature enacts policy without delegating, participants gauged the legislative blame to be 8.43 in the transmission scenario, 7.94 in the implementation scenario, and 7.58 in the creation scenario. As in the Mechanical Turk subject pool, where legislatures delegated the task to an instrument agency, participants blamed the instrument significantly less (in each of the three cases, $p = < 0.05$). However, participants found autonomous agencies to be much more blameworthy.

Figure 4: Differences between legislative and agency blame scores across agency authority scenarios.
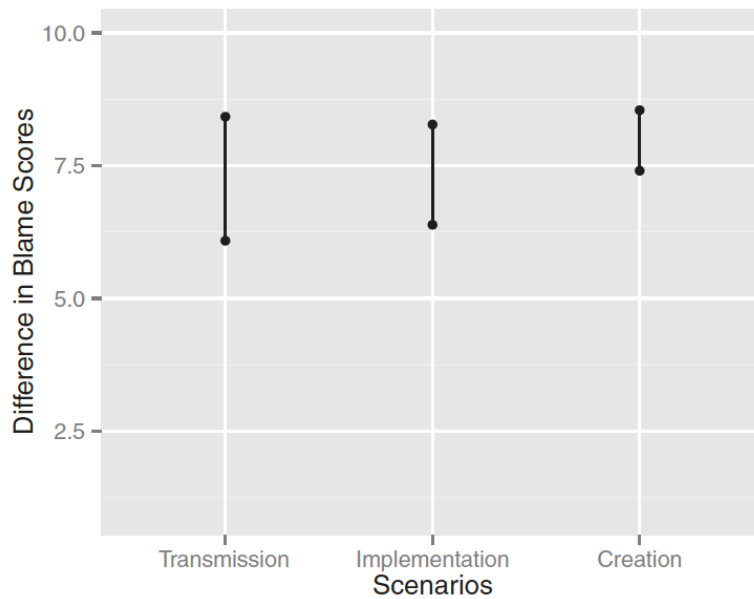
NOTE: Agency blame scores are the average of instrument and autonomous agency blame scores in each scenario.

VIII. Conclusion

The present study suggests that evaluations of policy are not independent of the structures that produce them. Rather, delegations of policy-making authority systematically induce voters to underassign blame to legislatures. Legislatures face less blame even in cases where they delegate to nonautonomous, "instrument" agencies. In effect, legislatures can delegate the final, public-facing component of a policy-making process to an agency in order to soften the blame they face. This finding coheres with a body of economic research that shows that attributions of blame diminish when a principal employs an agent.

Political accountability, however, "hinges on the principals' [in this case, the voters'] ability to tailor sanctions and rewards to choices made by their agents," which, in turn, "depends . . . on the principals' access to a clear picture of those choices and their relationship to policy outcomes."[56] Generations of scholars have assumed that voters' ability to sanction incumbent officials does not depend on changes in the basic institutional design of the political system. The findings reported here, however, show that our ability to assess blameworthiness diminishes when power is delegated. These findings undermine the foundational premise of the traditional theory of electoral accountability.

---

[56] Ethan Bueno Mesquita and Dimitri Landa, *An Equilibrium Theory of Clarity of Responsibility*, unpublished manuscript on file with author.

CHAPTER THREE: DOES SOFT PATERNALISM UNDERMINE
ACCOUNTABILITY?

I. The significance of new regulatory tools for electoral accountability

This Chapter reveals the second major flaw in the traditional theory of electoral accountability: the consequences of its dated conception of *how* power is exercised. The traditional model assumes legislatures use laws to change behavior. Increasingly, however, regulators are using other tools to change behavior, including soft paternalism.[1] Unlike hard paternalism, soft paternalism works by leveraging background forces to exploit biases in individuals' reasoning. All of this is well-known, and does not by itself pose anything more than a cosmetic threat to the traditional model.

The results in this Chapter, however, show that this dated conception of how power is exercised undermines the fundamental logic of the model. As discussed in Chapters One and Two, the traditional model operates on a "sanctioning" theory of voting.[2] In order for the sanctioning theory to work, the electorate must accurately assess the performance of officials. The results here suggest that individuals under-blame regulators for failed nudges. Taking the level of blame issued by individuals for failed laws as the baseline, we see that officials can escape blame by regulating via nudge. As a result, when voters seek to "sanction" officials, they will under-sanction certain officials. On the traditional model, sanctioning via vote is the mechanism by which individual preferences are conveyed to leaders. Thus, under-sanctioning impedes the flow of preferences from individuals to leaders, and ultimately reduces the ability of officials to allocate resources in the right places.

II. Summary of empirical results

This Chapter reports results that show that even if nudges are entirely transparent, individuals nonetheless under-blame regulators when nudges fail. In order to determine how soft paternalism affects perceptions of blame, the three experiments reported here randomly assign participants to a soft paternalism condition and a hard paternalism condition. The First Round of experiments compare how individuals assign blame when two policies, one soft paternalistic and the other hard paternalistic, go wrong. Even though the consequences of the failed policies are the same, individuals tend to blame the nudges less.

After determining that soft paternalism produces under-blaming, this Chapter generates data on potential causes of this under-blaming. The first hypothesis the Chapter tests (in the Second Round of experiments) about why nudges generate less blame is that when regulators act

---

[1] Throughout this Chapter, I will use the terms "soft paternalism" and "nudge" interchangeably. In theory, nudges are just one category of soft paternalism. In practice, the two are used interchangeably.

[2] As discussed in Chapter Two, the traditional model is compatible with a "selection" theory of voting. But even in a selection theory of voting, voters must have a sense of the blameworthiness of officials.

by altering the "choice environment," individuals to not perceive the state as *acting*. The second experiment does not provide support for this hypothesis, however. It appears as though individuals understand that the state is acting when it alters the choice environment. The second hypothesis the Chapter tests (in the third experiment) about why nudges generate less blame is that individuals to not perceive nudges as coercive. The experiments here show that individuals find nudges to be less coercive than laws, even when the consequences for violation are the same.

III. Framing the research question of this Chapter

Critics frequently argue that nudges are more covert, less transparent, and more difficult to monitor than traditional regulatory tools. Edward Glaeser, for example, argues that "[p]ublic monitoring of soft paternalism is much more difficult than public monitoring of hard paternalism."[3] As one of the leading proponents of soft paternalism, Cass Sunstein, acknowledges, while "[m]andates and commands are highly visible,"[4] soft paternalism, "and some nudges in particular[,] may be invisible."[5] "The public can observe the size of sin taxes and voters can tell that certain activities have been outlawed,"[6] but nudges often "appear to be invisible."[7]

To many, this lack of transparency is feature of soft paternalism, not a bug. Soft paternalism leverages individual cognitive biases to produce outcomes that approximate those achieved by a completely informed, rational decision-maker.[8] Many of such interventions operate by altering the salience of some feature of a decision-making environment. For example, if foods in a cafeteria line placed at eye-level are more likely to be purchased than foods placed at waist level, soft paternalism recommends placing healthy foods where they are most easily seen. However, individuals "often correct against bias when they become aware of it."[9] Thus, the fact that the regulatory intervention calls for altering the "choice environment" (in this case, the cafeteria line) so as to diminish the salience of the unhealthy food is typically not rendered salient to those in line. "Covertness," therefore, "is a feature of particular nudges," including "those that affect unconscious processing."[10] In short, "[s]ome of [soft] paternalism's most celebrated techniques may not work very well were they known to those on the receiving end."[11]

Soft paternalism's lack of transparency, critics argue, undermines its very promise. Soft paternalism is soft because it imposes trivial costs on those who seek to avoid the regulatory scheme. It offers "an option to resist."[12] Yet if the intervention is unperceived, as a practical matter resistance is unlikely. Further, this lack of transparency creates accountability problems.

---

[3] Edward L. Glaeser, *Paternalism and Psychology*, 73 U. CHI. L. REV. 133, 151 (2006).
[4] Cass R. Sunstein, Why Nudge?: The Politics of Libertarian Paternalism 143 (2014).
[5] *Id*. at 145.
[6] Glaeser, *supra* note 3, at 151.
[7] Cass R. Sunstein, *Nudging and Choice Architecture: Ethical Considerations*, at 3, available at http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2551264.
[8] Sunstein, *supra* note 4.
[9] Ryan Calo, *Code, Nudge, or Notice?*, 99 IOWA L. REV. 775, 795 (2014).
[10] Sunstein, *supra* note 4, at 3.
[11] Calo, *supra* note 9, at 795.
[12] *Id*.

Glaeser argues that the lack of transparency associated with soft paternalism makes it "intrinsically difficult to control and ... at least on these grounds, more subject to abuse than hard paternalism."[13] Ryan Calo argues that soft paternalism may "reduce in citizens even the habit or instinct of thinking and choosing for themselves."[14] As a result, regulators may over time come to control more areas of life while facing less and less resistance, effectively "sidestepping political and judicial process."[15]

In response to this challenge, proponents of nudging argue that invisibility for any given individual in a particular choice environment is compatible with "careful public scrutiny" of the nudge.[16] That is, invisibility for individuals acting within a choice environment is compatible with high salience for the public scrutinizing the intervention at a distance. As Sunstein points out, graphic images on cigarette packing, and prominently displayed fuel economy labels, for example, "received considerable public scrutiny and attention."[17] Of course, regulators don't control the news cycle or the public's attention; they can't guarantee that all nudges will receive attention. However, Sunstein argues that regulators can follow John Rawls' "publicity principle," which "bans government from selecting a policy that it would not be able or willing to defend publicly to its own citizens."[18] So, for instance, "if government officials use cleverly worded signs to help reduce litter ... they should be happy to reveal both their methods and their motives."[19] Against charges that nudges lack transparency and produce unaccountable regulators, proponents argue that nudges should be "visible, scrutinized, and monitored" by the public at large.[20]

Sunstein and Thaler's proposed solution to problem of holding regulators employing *soft* paternalism accountable is modeled on the same institutional design used to hold officials accountable for *hard* paternalism. This institutional design is nothing if not familiar.[21] The solution is the traditional model of electoral accountability, which we have already seen in Chapters One and Two. For the purposes of this Chapter, the salient aspects of this mode of accountability are as follows. Retrospective oversight assumes that (1) public visibility of regulation allows for (2) public scrutiny and blame and (3) political monitoring and sanction. If the public deems regulations unacceptable, "[p]olitical safeguards are triggered ... the government must defend itself publicly[,] ... [a]nd if the public defense is perceived as weak, the proposed action may well crumble."[22]

This solution to soft paternalism's accountability problem – importing the traditional model of electoral accountability – rests on an important assumption: that retrospective oversight works as well with nudges as it does with hard paternalism. More technically, it assumes that

---

[13] Glaeser, *supra* note 3, at 151.
[14] Calo, *supra* note 9, at 795.
[15] Id.
[16] Cass Sunstein, *Behavioral Economics and Paternalism*, 122 YALE L. J. 1826, 1893 (2013)
[17] *Id*.
[18] Richard Thaler and Cass Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness 244 (2008).
[19] *Id*.
[20] Sunstein, *supra* note 16, at 1892.
[21] John Ferejohn, *Incumbent Performance and Electoral Control*, 50 PUBLIC CHOICE (1986).
[22] Sunstein, *supra* note 16, at 1980.

publicizing nudges renders individuals at least as capable of scrutinizing nudges as they are capable of scrutinizing traditional command and control legislation.

Indeed, even critics of soft paternalism concede this point: they assume that, so long as nudges are transparent, regulators can be held accountable via democratic oversight. In the current debate, the point of contention is whether nudges are, in fact, transparent, and, if not, how they can be made so. But both sides agree that if nudges *are* sufficiently visible, individuals will be at least as capable of scrutinizing nudges as they are capable of scrutinizing traditional command and control legislation. This Chapter tests that crucial assumption.

Section II of the Chapter situates nudges within the broader context of regulatory "mechanism choice." It highlights that fact that nudges must balance efficiency with accountability. Because accountability is an important regulatory value – considered, among other things, in judicial analyses and regulatory impact analyses – it is natural to suggest that regulators trade some efficiency for accountability. Since soft paternalism is a relatively new regulatory tool, regulators and advocates have focused primarily on establishing the efficiency of nudges and thought less of how we hold regulators accountable for "nudges gone wrong."

To identify and (begin to remedy) this deficiency, Section III provides the first study of soft paternalism to directly compare how third-party observers politically evaluate hard and soft paternalistic regulatory interventions. The results reported here suggest that choice of regulatory instrument – law or nudge – makes a significant difference in the political responses of third party observers. Individuals perform less well scrutinizing nudges than they do scrutinizing laws, even when both interventions are equally visible. Holding constant all dimensions of the situation other than the choice of regulatory instrument – including magnitude of cost resulting from the unsuccessful policy – participants consistently blamed regulators less for nudges that fail to achieve their desired policy goals than for laws that fail to achieve their policy goals.

Thus, the response of Sunstein and Thaler to critics appears inadequate. The accountability architecture designed to hold regulators responsible in the context of command and control regulation performs less well in the context of nudges. It produces less blame and is more permissive. Even in plain view, nudges are more difficult to monitor than laws.

After diagnosing this accountability problem, Section IV of the paper turns to identifying potential causes of the problem. Section IV generates data about the causes of the divergence in blame allocations. Here, the paper undertakes what might be described as two diagnostic tests that produce data that will help us to understand why individuals might perform less well when evaluating nudges. These diagnostics operationalize two theories from the literature on nudges. In particular, it tests (a) whether individuals fail to perceive or conceptualize the creation of choice architectures as state action at all and (b) whether individuals conceptualize nudges as less coercive than law, even when the material costs of the two are equivalent.

III.A. *Efficiency versus accountability*

"All models of administration," argued then Professor Elena Kagan, "must address two core issues: how to make administration accountable to the public and how to make

administration efficient or otherwise effective."[23] The question of how to increase the efficiency and effectiveness of regulation involves a question of mechanism choice - which tool available to regulators best addresses the problem. By contrast, the question of accountability involves a question of which "actors have the right to hold other actors to a set of standards, to judge whether they have fulfilled their responsibilities in light of these standards, and to impose sanctions if they determine that these responsibilities have not been met."[24] Historically, "these two goals [efficiency and accountability] often conflict with each other."[25]

Regulation seeks to solve problems created by market failures; in particular, problems of externalities, market power, and asymmetric information.[26] Regulators try to solve these problems through regulatory tools such as taxes, liability rules, and mandates.[27] At the root of the conflict between effectiveness and accountability is the fact that effectiveness is often best promoted by delegating significant political or legal power to regulators.[28] Successful implementation of regulatory tools often "requires delegations that provide significant discretion to agency officials."[29] In turn, the provision of broad discretion creates the opportunity for "government failure." Significant discretion exposes the public to significant downside risk. Discretion "can ... impose its own problems, including compliance costs, inhibition of innovation, ancillary risks, and rent-seeking."[30] Further, as Todd Zwyicki argues, "[b]ureaucrats are themselves susceptible to biases."[31] Thus, government failure is the counterpart to market failure. As regulation is the solution to market failures, accountability is the solution to government failure.[32]

To this standard template, behavioral economics adds the idea of the behavioral market failure. These offer a "supplement to the standard account of market failures" and enlarge "the class of justifications for government regulation."[33] Where neoclassical approaches to regulation identify the market's propensities to err, behavioral economics identifies "the human propensity to err."[34] As in the case of traditional market failures, nudges typically require delegation and discretion, which, in turn raises the risk of government failure. Nudges, like more traditional types of regulation, must address the tension between efficiency and accountability.

---

[23] Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245, 2331 (2001).

[24] Ruth Grant and Robert Keohane, *Accountability and Abuses of Power in World Politics*, 99 AM. POL. SCI. REV. 29 (2005).

[25] Kagan, *supra* note 21, at 2331.

[26] Jonathan B. Wiener and Alberto Alemanno, Comparing regulatory oversight bodies across the Atlantic: the Office of Information and Regulatory Affairs in the US and the Impact Assessment Board in the EU, in COMPARATIVE ADMINISTRATIVE LAW 310 (Rose-Ackerman, ed.) (2011).

[27] Ryan Bubb and Richard Pildes, *Behavioral Economics Trims its Sails and Why*, 127 HARV. L. REV. 1593, 1602 (2014).

[28] Richard Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1676-79 (1975).s spectrum, the two should be deemed equally coercive. er kind) imposed on choosers by choice architects.", with points marked i

[29] Kagan, *supra* note 21, at 2331.

[30] Wiener, *supra* note 26, at 310.

[31] Todd Zwyicki, *Behavior, Paternalism, and Policy: Evaluating Consumer Financial Protection*, at *32 (2014), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2408083.

[32] Wiener, *supra* note 24, at 311.

[33] *Id*.

[34] *Id*.

The challenge, then, facing regulators is not merely to select the most efficient and effective tool. Rather, the challenge is to select that tool which will minimize the sum of market failures and government failures. Up to this point, proponents of soft paternalism have generally focused on establishing that nudges serve as an effective tool to combat traditional and behavioral market failures. Proponents have not, however, established that nudges minimize the sum of government failures. In particular, proponents of soft paternalism have not established that nudges minimize what is arguably the most crucial type of government failure: accountability failures.

In short, while efficiency and results matter greatly for regulation, accountability matters, too. Just because a new tool (like nudges) is more efficient does not mean that it can avoid how it impacts accountability.

III.B. Accountability architectures

There are two main oversight frameworks designed to avoid accountability failures. Technocratic oversight involves tools like impact assessments, cost-benefit analysis, cost-effectiveness analysis, and risk-risk tradeoff analysis.[35] These tools seek to reduce bureaucratic bias and information asymmetries between agencies and regulatory oversight bodies.[36] For example, Sunstein argues that "rules that embody soft paternalism should be subject to public scrutiny in advance, often through notice-and-comment rulemaking."[37] Under this method, agency officials would evaluate nudges in order to determine if they are justified. Democratic oversight, by contrast, aims to create accountability through the ballot box.[38] Soft paternalists advocate a standard similar to Rawls' publicity condition, which holds that laws are publicly acknowledged and individuals are aware, or can become aware, of the law.[39]

Proponents of soft paternalism consider both of these frameworks as ways of achieving regulatory accountability. Although the approaches differ in many ways, both approaches require individuals – be they voters or technocrats – to engage in retrospective evaluation of nudges. In other words, both of these accountability architectures depend on individuals accurately perceive the blameworthiness of regulators for failed nudges. Hence the importance of this study: if individuals are misperceiving blameworthiness for nudges, then a central input that drives their judgments about accountability will be biased.

As discussed in Chapters One and Two, on the traditional electoral accountability model, voters sanction officials for poor performance. Elections serve as "a sanctioning device that induces elected officials to do what the voters want," since "[t]he anticipation of not being reelected in the future leads elected officials not to shirk their obligations to the voters in the present."[40] Voters' actions are a function of "their evaluations of officeholders on their actual performance in office rather than on hypothetical promises they might make during a

---

[35] Wiener, *supra* note 26.

[36] Michael Livermore, *Cost-Benefit Analysis and Agency Independence*, U. CHI. L. REV. 609, 616 (2014).

[37] Cass Sunstein, *Behavioral Economics and Paternalism*, 122 YALE L. J. 1826, fn. 2 (2013).

[38] Wiener, *supra* note 26, at 310.

[39] Daniel Hausman, *To Nudge or Not to Nudge*, J. POL. PHIL. 123, 132 (2010).

[40] James Fearon, Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance 61 (1999).

campaign."[41] If we assume that voters treat elections thusly, the politicians will be incentivized to exercise what control it does have over agencies to enact popular policies.

Also consistent with the traditional model of electoral accountability is the "selection" approach. The selection approach is a forward-looking exercise, rather than a backwards-looking endeavor. "[V]oters can use the policy choices of politicians to draw inferences about their underlying characteristics."[42] The selection model critiques the sanctions model on grounds that election promises are cheap talk, but actions reveal information about a candidate's true type. Voters then use those underlying characteristics to predict how well the politicians will perform in a future period. The traditional model of electoral accountability fits most naturally with the sanctions model, but it is not incompatible with the selection model.

While there are several viable approaches to creating accountability institutions for regulators, they all depend on voters assessing the performance of politicians. Soft paternalism is an increasingly popular regulatory tool. Thus, individuals' skill at assessing the blameworthiness of regulators for failed nudges is a critical part of accountability architectures. Our research question for this Chapter, then, is whether individuals are as adept at assessing blame when soft paternalism fails as they are at assessing blame when hard paternalism fails.

IV. Round One: Does Soft Paternalism Produce Less Blame than Hard Paternalism?

The first round of experiments reported here tests the claim that individuals blame regulators less for failed instances of soft paternalism than they do for failed instances of hard paternalism. This thesis is vitally important to proponents of soft paternalism, for it allows them to respond to critics who argue that soft paternalism's covertness undermines accountability. Sunstein and Thaler argue that third party observers will hold regulators *equally accountable* for failed nudges. Proponents of soft paternalism argue that while some nudges are covert for individuals operating within the choice environment, they can be made entirely transparent to to elites who investigate them, so that enacting officials can be held accountable.

The experiment in Round One test that claim: whether varying the type of regulatory instrument policymakers use to solve a problem affects individuals' perceptions of blame. The experiments in Round One employ vignettes that tell the story of a jurisdiction's response to some problem – credit card debt, for instance – in concise, nontechnical language. To measure the effect of regulatory instrument on the quantity of blame participants allocate to regulators, each participant was randomly assigned to one of two conditions. To establish a baseline level of blame, one group of participants was assigned to the "law" condition. This group read a vignette describing a wholly *hard paternalist* response, using the traditional instruments of law, to the problem. The remaining participants were assigned to the "nudge" condition. In the nudge conditions, the policymaker tries to solve the problem by altering the choice environment, rather than by law.

Experiment 1 was administered online to 140 participants, recruited and compensated through Mechanical Turk. All participants completed two demographic questions (48 percent

---

[41] Ferejohn, *supra* note 21, at 6.
[42] Justin Fox and Stuart Jordan, *Delegation and Accountability*, 73 J. OF POLITICS 831 (2011).

female, M_{age} = 29) and were randomly assigned to one of the two conditions described above: law or nudge. To test the claim that third party observers will hold regulators as blameworthy for failed nudges as for failed laws, Experiment 1a compares the blame scores of the control and treatment groups.

In order to compare blame scores of nudges with laws, we need some way to make the harm resulting from the failed policies comparable. In Chapter Two, the harm was defined as exactly the same (e.g., the costs of a failed environmental policy). We were able to define the harm thusly because the tools the legislature and agencies used to regulate were the same. That strategy does not work here, however. In Chapter Three, we want to compare the blameworthiness wrought by the failure of two different types of policymaking tools: laws and nudges.

At first glance, one might think that it is impossible to compare the results of laws and nudges, since nudges seek to impose no material costs on those who "opt-out" of a default choice architecture. Hence the *softness* of the paternalism. However, regulation can impose non-material yet significant costs. Sunstein acknowledges this point explicitly in his 2014 book, *Why Nudge? The Politics of Libertarian Paternalism*, concluding that "psychic costs, no less than material costs, can alter behavior; indeed psychic costs might have the same effect as large material ones."[43] In other words, nudges do, in fact, impose costs.

Since both laws and nudge impose costs, we can try to equalize those two sets of costs. If the two sets of costs are equivalent, participants in the survey should treat the enacting officials as equally blameworthy. If we observe differences between the two sets of blame scores, we can attribute those differences to the choice of regulatory instrument (law or nudge), since those are the only relevant differences between the two vignettes.

Experiment 1 posed three scenarios to participants. The scenarios are drawn directly from Sunstein's own examples in his 2014 paper on the ethics of nudging.[44] This ensures that the examples are representative of the core instances of soft paternalism.

IV.A Experimental scenarios and results

In the first scenario, a university seeks to limit the harm to the environment resulting from printing too many sheets of paper. In the nudge condition, the university created a choice architecture that made double-sided printing the default. It imposed a fee of $.20 per page on students wishing to override the default setting and print on single-sided paper. The vignette stipulated that this fee would cost students about $100 over the course of the semester. Though the university had good intentions (to help the environment), "[t]he new policy imposed the largest costs on poorer students at the University." Students "who chose not to override the default choice felt embarrassed to circulate double-sided class handouts and turn in double-sided term papers," and as a result "[t]hose poorer students who did pay the fee had to work more hours at their part-time jobs to account for the new expense." In the law condition, "[s]tudents

---

[43] Cass Sunstein, WHY NUDGE? THE POLITICS OF LIBERTARIAN PATERNALISM 86 (2014). Further, Sunstein writes, "Imposition of affective costs is paralleled by the creation of affective benefits …." *Id.* at 86.
[44] Sunstein, *supra* note 5.

were instructed that if they illegally altered the computer's settings and overrode the default, they would be subject to a fee" of $100 per semester. The same consequences resulted as in the nudge condition. After reading the vignette, participants in both conditions were asked: "On a scale of 1 (not blameworthy at all)—10 (very blameworthy), how blameworthy do you think the legislature's action was?"

In the second scenario, regulators seek to nudge individuals away from credit card debt, since "many citizens were carrying very high balances and paying very high interest rates on their credit cards." In the nudge condition, a new policy "required credit card companies to inform customers, in a clear and conspicuous way, of the costs of paying the minimum amount every month, as opposed to the costs of paying the full amount in a specified period." Afterwards, consumers were provided with educational materials about the true costs of credit card debt, aimed at encouraging consumers to pay more of their bill each month. After this nudge, consumers could choose to pay only the minimum amount, should they chose to. The policy succeeded in inducing consumers to pay more each month. However, it "had the side effect of leaving many poorer consumers without enough money at the end of each month," which led to them "borrowing money at even higher interest rates." In the law condition, regulators legally mandated that consumers pay more each month. The side effects were the same.

In the third scenario, regulators seek to respond to decreased air quality issues by limiting the amount of pollution. In the nudge condition, regulators nudge consumers toward buying cars with a revolutionary fuel economy mechanism. It turned out, however, to be faulty, and deaths from car accidents tripled. In the law condition, regulators mandated consumers purchase cars with the new mechanism. As in the previous conditions, the same results followed in both scenarios.

Proponents of soft paternalism would predict that participants gauge the two sets of officials (those enacting policies via nudges and those enacting policies via law) as equally blameworthy. On this view, we should see equivalent levels of blame in the two conditions. In fact, however, participants tended to score the nudge condition as *less* blameworthy. Participants who read the nudge condition rated the regulator as significantly less blameworthy than participants who read the law condition in the second and third experiment. In the first experiment, participants also scored the law regulators as more blameworthy, but the difference was not statistically significance ($p = 0.061$). The results were analyzed using a Welch two-sample t-test, which is one of the most conservative measures of statistical significance.

Chart 1: Summary Statistics for Round One

	Mean	**Standard deviation**	**p-value**
Experiment 1	Nudge=4.43, Law=5.30	N=2.71, L=2.76	0.06137
Experiment 2	N=5.15, L=7.15	N=2.60, L=2.45	< 0.001
Experiment 3	N=6.21, L=7.33	N=2.83, L=2.45	0.01292

Thus, contrary to what proponents of soft paternalism argue, nudges appear to provide shelter to regulators from blame. Even if nudges are made entirely transparent, individuals nonetheless find failed nudges less blameworthy than failed laws, even when the consequences

are the same. Given these results, Rawls' publicity principle is not sufficient to ensure equal accountability between laws and nudges. Thus, the traditional account of electoral accountability needs to be updated to reflect the use of a variety of regulatory tools by officials, not just hard laws.

V. Round Two: Why Does Soft Paternalism Produce Less Blame?

This Round of experiments tests two hypotheses that might explain why nudges produce less blame than laws. If we are able to identify why individuals process nudges as they do, we might be able to craft institutional designs that avoids or minimizes this bias.

V.A. The "no state action" hypothesis

The first hypothesis is that nudges attract less blame because individuals do not perceive regulators as *doing anything* when they alter a choice environment. The intuition here is that observers may not process rearranging an environment as an action. And if the government is not acting (nor is it failing to act), how can it be held blameworthy? Regulators may simply seem *uninvolved* in the way, say, that a cafeteria line is laid out.[45] As Sunstein argues, choice architectures are inevitable.[46] Some item will be first in a cafeteria line, no matter what. In this way, newer forms of regulation work on a different canvas than older forms, like hard law, do. The presence of hard law is not inevitable. It may seem to individuals that regulators are simply not in the business of arranging choice architectures, in the way that they are in the business of passing laws. Basic civics teaches us how a bill becomes a law; it does not teach us how a cafeteria line is arranged.

Thus, if individuals view the state as simply uninvolved in the business of architecting choices, they may see less reason to hold regulators blameworthy. To test that hypothesis, I gave the same vignettes described above to a new group of participants. Except rather than ask participants about the regulators' blameworthiness, the survey asked participants two questions (*a*) "On a scale of 1 (not proactive at all) ---- 10 (very proactive), rate how proactive the University was in addressing the problem" and (*b*) "On a scale of 1 (false) ---- 10 (true), assess the truth of the following statement: 'The University's plan changed the costs of printing.'" All participants were assigned to the control condition (law) or the experimental condition (nudge).

The results of this experiment do not provide evidence for the "no state action" hypothesis. For each of the three vignettes, there was no statistical difference between proactivity measures for the law condition and the nudge condition. It appears as though individuals are able to perceive that regulators act when they alter a choice architecture.

V.B. The "coercion" hypothesis

---

[45] In other realms, we typically think of an act or omission to be necessary to trigger moral or legal liability. See Jonathan Haidt and Jonathan Baron, *Social roles and the moral judgement of acts and omissions*, 26 EURO. J. SOCIAL PSYCH. 201 (1996). Likewise, criminal culpability requires an act or omission. I am unaware of any sort of moral, social, or legal liability that might result from what we might call *rightful noninvolvement*.

[46] Sunstein, *supra* note 5, at 11.

The second hypothesis considered in Round Two is the idea that nudges are inherently less coercive that laws. There is a fair bit of theoretical confusion in the soft paternalism literature on what, exactly, constitutes coercion and, relatedly, what distinguishes "soft" from "hard" coercion. We need not wade into that debate here, but we will need to clarify a couple of points before moving to the results of the experiment.

First, Grant Lamond supplies a definition of coercion that fairly captures the lay sense of coercion used in the experiment here: "any use of pressure that is sufficient in the circumstances to make [most reasonable people] do what he would not otherwise do, and is deliberately imposed for that purpose."[47] One might argue about whether nudges deploying rational persuasion constitute coercion, but for our purposes it is enough to know that coercion is something like an intentional use of pressure (of one sort of another). The experiment does not depend on the participants identifying edge cases.

Some argue that nudges are the kinds of things that are, by their nature, non-coercive or exert very little coercion.[48] Sunstein considers this view, noting that "it might be best to understand paternalistic interventions in terms of a continuum from hardest to softest, with points marked in accordance with the magnitude of the costs (of whatever kind) imposed on choosers by choice architects."[49] And, indeed, the most canonical of nudges, such as the cafeteria line example, impose very low magnitude costs. However, Sunstein ends up rejecting this view. He argues that it is more accurate to say that the categorical distinction between hard and soft paternalism is that the latter imposes only immaterial costs (on end-users), where hard paternalism imposes material costs (on end-users).[50] Soft paternalism, Sunstein argues, can exert very strong psychic pressure, which could have very high resulting costs. And hard paternalism can exert the most minimal of pressure, such as a small fine. In other words, the essence of the coercion used by soft paternalism is psychic, and the essence of the coercion used by hard paternalism is material.

If that is true, does it mean that we simply can't compare the coerciveness of soft and hard paternalism? How do we compare psychic harms with material costs? Are the two types of coercion incomparable? One way of finding out would be to ask participants to compare the two, and see what type of results we get. After all, what we are interested in is how ordinary individuals perceive the comparison. An alternative to this strategy is to ask participants to judge the coerciveness of the *consequences* of failed nudges and laws, not the nudges and laws themselves. Doing so establishes more of an equivalency between the two things being compared. The experiment here equalizes the costs of failed nudges and failed laws. Both sets of consequences are material.

Here, we use the same experiment we have used previously. Instead of asking participants about blameworthiness or coercion, however, this experiment asks, "On a scale of 1 (not coercive at all) ---- 10 (very coercive), rate how coercive the [regulators'] response was."

---

[47] Grand Lamond, *Coercion and the Nature of Law*, 7 LEGAL THEORY 35 (2001). Under Lamond's definition, coercion might be non-material. An unjust law, for example, might be intellectually or emotionally coercive.
[48] See, e.g., David Adam Friedman, Public Health Regulation and the Limits of Paternalism, 46 CONN. L. REV. 1687, 1693 (2014).
[49] *See, e.g.,* Sunstein, *supra* note 2, at 1859.
[50] *Id.* at 1860.

We assigned 86 participants to either the control (law) or treatment (nudge) condition. In each of the three scenarios, participants who read the nudge condition rated the nudge as significantly less coercive than the law, even though the two had the exact same consequences.

Chart 2: Summary Statistics for Coercion Hypothesis

	Mean	**Standard deviation**	**p-value**
Experiment 1	Nudge=4.74, Law=6.87	N=2.72, L=2.59	< 0.001
Experiment 2	N=4.24, L=6.55	N=3.11, L=2.97	< 0.001
Experiment 3	N=3.49, L=6.34	N=2.58, L=2.59	< 0.001

The evidence presented here is not definitive, but it suggests that individuals might perceive nudges as less coercive than law, even when their consequences are the same. To be sure, there very well may be other things going on in these experiments that produce these results. It might be the case, for example, that there is some expressive message produced by hard paternalism that leads either to a feeling that they are more coercive. Despite the fact that we do not know for sure what is leading individuals to find nudges less blameworthy, it is important to continue to pursue lines of inquiry such as this. If we have a sense of what is driving individuals to underblame regulators for failed nudges, we can perhaps use information such as that presented in Round Two to help de-bias our institutional designs.

VII. Conclusion

Contrary to what both proponents of soft paternalism argue, nudges do appear to undermine accountability. Proponents argue that, although any given instance of soft paternalism might be "invisible" to the end-user, soft paternalism does not undermine accountability because journalists and other interested parties can publicize their use. Sunstein argues that such publicity, which satisfies the Rawlsian standard, ensures that nudges are adequately evaluated by the public. The experiments reported here strongly suggest that this is not true. Individuals blame regulators less for failed nudges than for failed laws. Since more and more regulation occurs through non-traditional means, such as nudges, the traditional account of electoral accountability needs to be revised in order to avoid under-blaming certain groups of regulators.

CHAPTER FOUR: DOES STABILITY UNDERMINE ACCOUNTABILITY?

I. The significance of stability to electoral accountability

Much of the discussion around the traditional model of electoral accountability focuses on voters holding officials accountable. As important as that piece of the theory, however, is the process whereby officials enact change. The question this Chapter raises is **when** officials can make legal changes. The traditional approach to electoral accountability treats policy change as costless. It assumes that, if individual preferences change, regulation should change, too – immediately.

The theoretical literature on legal change, however, argues that legal change must be *rate-limited*. That is, the legal theory literature argues that the rule of law demands that the rate of legal change must be limited. If this literature is correct, then – *contra* the traditional theory of electoral accountability – policy change is not costless and democratic change is limited by some level of acceptable change.[1] If this legal theory literature is correct, then it appears as though the need for legal stability undermines political accountability. More provocatively, it appears as though the rule of law is incompatible with political accountability.

Chapter Four of the dissertation charts an intermediate course between these two extreme views. The traditional model assumes that change is costless and can occur as frequently as necessary. By contrast, the leading legal theories argue that change is so costly that democratically made policy can only change every so often. Chapter Four charts an intermediate course by, first, uncovering what interests are at stake when we talk about legal change and, second, by working out how those interests are affected by change. It sketches a framework for tallying the costs and benefits of legal change.

Thus, this Chapter shows that the traditional model is wrong to assume that change is costless. But it also shows that change is not *as* costly as previous theories have thought. Here, we not only identify a crucial flaw in the traditional model, but offer a path forward for future models. This intermediate course allows us to reconcile the rule of law with political accountability – an important feat, which the traditional model is unable to accomplish.

Finally, a note on method. As discussed in the Introduction, behavioral approaches are frequently seen as competitors to rational choice approaches. In contrast, this dissertation treats rational choice perspectives and behavioral as partners. Behavioral studies tend to work best when there is a solid foundation of theory underlying them. Behavioral approaches that test rational choice models by subjecting their assumptions to empirical scrutiny are more conceptually precise than studies that rely on folk wisdom and (often unstated) assumptions.[2] In the prior two Chapters, there were substantial bodies of theoretical work that my behavioral studies could draw on. While there is not a body of theoretical work on stability, but it does not squarely address the issues around electoral accountability. This Chapter seeks to remedy that gap.

II. Framing the research question

Our lives proceed amid dense webs of regulation. We intend, endeavor, and cooperate alongside regulations designed both to govern us *and* to possess the capacity to flexibly respond to rapidly changing circumstances. Although governance by administrative regulation is now largely uncontroversial, *flexible* regulatory governance is anything but. Legal philosophers have long argued that the rule of law is fundamentally incompatible with the instability generated by flexible regulatory regimes. According to these theories, flexible regulatory regimes

---

[1] Frustratingly, the legal theory literature explicitly rejects the idea that we can specify how much change is acceptable.

[2] Moreover, from a Bayesian perspective, one of goal of interdisciplinary research is trying to unite and link otherwise discrete bodies of scholarship. Viewing rational choice and behavioral approaches as partners accomplishes that goal.

produce norms that lack the capacity to guide conduct, and thus the capacity govern. This view poses a very real problem for the traditional theory of electoral accountability, which is premised on the idea that laws can and should change as frequently as the public demands.

In order to reconcile these two positions, this Chapter develops a framework for analyzing the ways in which instability resulting from responsive regulation can impair the ability of law to guide, and thus regulate. Law is a tool used by officials to regulate. Just as a knife must be sharp to cut, law must guide conduct in order to regulate. Increasingly, officials seek to use law to regulate a rapidly changing world. The flexibility required to regulate such a world leads to instability in law, and instability in the law, it is widely believed, impairs law's ability to guide conduct. In short, according to conventional wisdom, instability and the rule of law are incompatible. As McCubbins, Rodriguez, and Weingast argue, "there are inevitable tradeoffs between the values of stability and consistency on the one hand and adaptation on the other."[3]

In fact, the ways in which instability impairs law's ability to guide conduct is complex and multi-faceted. The framework developed by this Chapter allows us to move past the simple claim that law is incompatible with instability, and to analyze, instead, the way in which officials can assess and distribute the costs to legality born out of the goal of responsive and flexible regulation.

III.A. How frequently do regulations change?

Surprisingly, we do not have good measures of how frequently regulations change. One might have imagined that those who argue that regulations must be rate-limited would have gathered this data, given how important they argue it is to the rule of law. Although the legal theorists making these arguments are hesitant to define what, exactly, makes a body of law unstable, it is instructive to consider the levels of volatility present in the American regulatory system.

In order to establish some sense of how frequently regulations change, this Chapter reports first of its kind data. Using data gathered from the U.S. Code of Federal Regulations, we can see how frequently regulations change over the period from 1986-2000.

The CFR data contains year-over-year statements of the rules in the CFR. It does not itself report change. The key to analyzing this data is to find a replicable pattern that captures the *differences* between yearly reports. After finding such a pattern, we can glean data from the code that shed some light on how frequently regulations are revised, which means a partial amendment to the text, or amendment, which means a wholesale revision of the text.[4]

The Code contains approximately 200,989 discrete provisions, and contains 137,651 amendments and revisions in the period from 1986-2000. Title 40 of the Code, which houses environmental regulations, saw over 10,000 revisions and amendments to discrete regulations. Title 48, which governs federal acquisitions systems, saw over 12,000 revisions and amendments. Title 7, which has approximately 17,000 discrete regulations governing agriculture, was revised or amended over 13,000 times. Over that period, some Titles averaged over one change per year. Title 19, which regulates customs duties, contains approximately 2,527 discrete regulations, and saw 2,828 changes. While many provisions did not change at all during this period, some changed rather frequently. E.g., 49 CFR section 1002.2, which governs filing fees paid to a transportation board, was revised 16 times.

With a sense of how frequently regulations change, let us now turn to the theoretical literature on stability and the rule of law.

III. Law and guidance

---

[3] Mathew D. McCubbins, Daniel B. Rodriguez, and Barry R. Weingast, *The Rule of Law Unplugged*, UNIVERSITY OF TEXAS RESEARCH PAPER SERIES 158 (2010).

[4] A description of the method by which the data was gathered, the Python script used to gather the data, and final dataset are available at https://github.com/hillad/stability.

Legal positivism is, in the first instance, a thesis about the conceptual foundations of legal validity.[5] It seeks to parse the *social* mechanisms that underlie the *legal* decisions that, together, determine the substance and govern the processes of a given legal system.[6] Natural law approaches, by contrast, seek to parse not only the social but also the *moral* mechanisms that underlie legal determinations of validity.[7] For both positivists and natural lawyers, proposition *P* accurately states the law governing some topic in a given jurisdiction, if and only if *P* meets the applicable criteria of legal validity, *C*, in that jurisdiction.[8] The fundamental difference between the two approaches is that positivists maintain that *C* obtains in only virtue of social mechanisms, while natural lawyers remain open to the possibility that *C* obtains in virtue of moral mechanisms.

For positivists, the social mechanisms at the foundations of the law are explicated through *descriptive* facts: claims about the behaviors or mental states of judges, legislators, executives, or voters.[9] So, for instance, positivists may point to facts about the behavior of courts to explain the content of *C*, but not to the merits of *C* itself. Positivists, in other words, reject grounding their explanations of the criteria of legal validity in *value* facts: normative or evaluative claims.[10] These are taken to bear only an arbitrary relationship with legal norms.[11] Natural lawyers, by contrast, might well explain the content of *C* through reference to value facts. Perhaps, for instance, *C* states the criteria of legal validity because it is demanded by democratic values. Explanations of these sorts are not open to positivists, for whom descriptive facts possess explanatory primacy. In giving descriptive facts explanatory primacy, positivists are committed to the claim that there is no law, properly so called, absent that which is created through social mechanisms.

One of the most important of these social mechanisms is the process by which legal officials posit a law, and individuals respond. Say that a legislature duly enacts proposition *P*. What follows? What is the characteristically *legal* way in which individuals are influenced by *P*? The way in which one answers that question determines her conception of *legal guidance*.

A conception of legal guidance explains what attributes law must possess, such that it has the capacity to guide or direct individual conduct. Although positivism is, in the first instance, a thesis about legal validity, it is not only that. It is also a descriptive theory that seeks to explain how law, "as something that must be *posited* through some social act or activity, either by enactment, decision, or practice,"[12] possesses the capacity to guide or direct individual conduct.

---

[5] *See*, *e.g.*, Mark Greenberg, *How Facts Make Law*, 10 LEGAL THEORY 157, 198 (2004); John Gardner, *Legal Positivism: 5 ½ Myths*, 46 AM. J. JURIS. 199 (2001); Christopher Kutz, *The Judicial Community*, 11 PHIL. ISSUES 442 (2001); David Enoch, *Reason-Giving and the Law*, in Leslie Green and Brian Leiter, eds., 1 OXFORD STUDIES IN PHILOSOPHY OF LAW 1-15 (2011).

[6] *See* Gerald J. Postema. *Coordination and Convention at the Foundations of Law*, 11 J. LEGAL STUDIES 165, 167 (1982) ("Thus positivists in the tradition stemming from Bentham locate the [criteria of validity] in matters of social fact, thereby rejecting the view that the validity of a law is a function of its truth or moral soundness."); *see also* Andrei Marmor, *The Rule of Law and its Limits*, LAW AND PHILOSOPHY, 23(1)(2004): 5.

[7] *See* Kutz, *supra* note 1, at 444-45. John Finnis, *On the Incoherence of Legal Positivism*, 75 NOTRE DAME L. REV. 1597-98 (2000).

[8] Kutz, *supra* note 1, at 443.

[9] Greenberg, *supra* note 1, at 157. For the classic statement, *see* H.L.A. Hart, *Positivism and the Separation of Law and Morals* 71 HARV. L. REV. 71(4)(1958): 593-629. As Ronald Dworkin put it, for positivists, legal validity is a matter "not [of the] content [of norms] but with their *pedigree* or the manner in which they were adopted or developed." Ronald Dworkin, TAKING RIGHTS SERIOUSLY (Cambridge: Harvard University Press, 1978): 17.

[10] Greenberg, *supra* note 1, at 157. There is, of course, the matter of hard and soft positivism. Soft positivists allow that value facts can possess derivative significance in determining the content of the law. Hard positivists argue that value facts can never do even that. The debate over whether value facts can affect the content of laws, not whether they can affect the criteria of legal validity.

[11] *Id*. at 157-59. Positivism's insistence on the primacy of descriptive facts gives rise to a thorny question: How can facts about what we happen to have been doing in the past give rise to genuine reasons about what we should be doing in the future? Or, put differently, how can it be that what we happen to do around here can provide reasons to keep doing it? *See*, *e.g.*, Kutz, *supra* note 1, at 446.

[12] Scott Shapiro, *Law, Morality, and the Guidance of Legal Conduct*, 6 LEGAL THEORY 127 (2000) (emphasis in original).

This question is particularly pressing for legal positivists. If a natural lawyer claims that *value* facts possess explanatory primacy in an account of legal validity, such that, at least in some cases, when *P* requires one to ϕ, one *thereby* has a genuine or unqualified reason to ϕ, then the task of explaining the connection between law and practical reasoning is, arguably, less urgent. For natural lawyers, the connection between the content of *P* and individual beliefs and action is more straightforward. In short, if *P* is underwritten by some moral logic, it *prima facie* possesses the capacity to guide conduct.[13] Positivists, however, have nothing of the sort upon which to rely. They must, instead, explain how it is that laws, which are not necessarily underwritten by any moral logic, possess the capacity to guide individual behavior.

The centrality of legal guidance has not eluded positivists. Scholars working in the field have developed a comprehensive conception of legal guidance and set out the social and legal conditions under which it obtains. The conventional wisdom conceptualizes legal guidance as that state of affairs in which an individual "is able to learn of his obligations or rights from [the relevant legal norm] without engaging in deliberation."[14] This conception of legal guidance treats guidance as a purely *epistemic* phenomenon. Law guides when it possesses the capacity to *convey information* about the content of the law to individuals. In order to possess this capacity, positivists hold, "there are certain conditions that the law has to meet."[15] A wide consensus holds that eight attributes "state necessary conditions"[16] for the existence of legal guidance and thus legality. These attributes are, by now, familiar. Laws must be general; they must be promulgated; laws must not succumb to retroactivity nor contradiction; they must be clear, and possible to follow; there must be congruence between the rules as written and as applied; and, finally, laws must be stable. That legal norms must possess each of these attributes in order to guide – in the sense of conveying to individuals their legal rights and obligations – is conventional wisdom among legal positivists.[17]

This Chapter argues that, under the standard conception of epistemic guidance, stability is not, in fact, required to guide conduct. Yet, legal theorists have long linked stability with guidance. In an effort to rationalize this long tradition of linking stability with guidance, this Chapter presents an alternative account of legal guidance. Rather than conceptualize guidance as purely epistemic, the Chapter presents a *motivational* account of guidance. Whereas an account of epistemic guidance explains merely how law conveys its requirements, an account of motivational guidance explains how law could be *the reason why* individuals comply. Such an account, the Chapter argues, is valuable not only because it best explains legal theorists' long-standing tendency to link stability with guidance, but also because it offers an account of the way in which the state and its citizens cooperate to overcome the harm to reliance interests possibly brought about by unstable law. The account of motivational guidance presented here models the state and its citizens as playing an assurance game, and argues that legal change, even change that is responsive to widely-held problems, will tend to only occur when both sets of parties are sufficiently assured that the other will cooperate. We begin our argument, however, with a much more standard idea: the way in which law, as a descriptive matter, is able to shape individual deliberation.

III.A. Epistemic Legal Guidance: Mediating between Law and Deliberation

When the U.S. federal government enacts new criminal laws, or seeks to change health care policy, or changes the rules governing securities laws, it seeks to regulate the behavior of hundreds of millions of individuals.[18] Regulation at this scale poses a problem: how is the state to *inform* this number of individuals of their

---

[13] There are a number of ways in which we might draw a link between natural law and the capacity to guide behavior. For instance, one might maintain that the capacity to guide conduct is a necessary feature of *moral* norms. If all moral norms possess the capacity to guide, then any legal norm underwritten by a moral norm will possess the capacity by extension.

[14] *See* Shapiro, *supra* note 8, at 153.

[15] Marmor, *supra* note 2, at 5.

[16] Scott Shapiro, LEGALITY (Cambridge: Harvard University Press): 395. Earlier, Shapiro observes that "regimes that flout these principles are simply not engaged in the basic activity of law." *Id*. at 394.

[17] *See generally* H.L.A. Hart, THE CONCEPT OF LAW (2)(Oxford: Oxford University Press, 1997); Joseph Raz, *The Rule of Law and its Virtue*, in THE AUTHORITY OF LAW (1979); Jeremy Waldron, *The Concept and the Rule of Law*, HARV. J. L. & PUB. POL'Y. 30(2006):15-30; John Gardner, *Hart on Legality, Justice, and Morality*, JURISPRUDENCE 1(2)(2010): 253-265.

[18] *See* Shapiro, supra note 12, at 72 (describing "the regulation of mass populations"). *See also* Jeremy Waldron, *The Rule of Law in Contemporary Liberal Theory*, 2(1) RATIO JURIS (1989): 79-96. The basic demographic contours of the U.S. population are available at United States Census Bureau, www.census.gov. It should be noted that U.S. law,

new duties, rights, or obligations? Even if each individual governed by a new regulation could be located, it is clearly beyond the capacity of the state to send officials door-to-door providing updates on new regulations.[19] As a substitute, the state crafts legal norms that provide citizens with *epistemic guidance*. Laws provide epistemic guidance when a citizen "is able to learn of his obligations or rights from [the relevant legal norm] without engaging in deliberation."[20] More colloquially, laws provide epistemic guidance to individuals when an individual is able to read the law and thereby learn how to act in compliance with the law. Such *impersonal* guidance is a substitute for the *personal* guidance rendered infeasible by the size of the modern state and its large-scale regulatory schemes.[21]

Epistemic guidance offers a solution to the problem of regulating large-scale populations. But not all laws provide epistemic guidance. A wide consensus holds among legal positivists that epistemic guidance is possible *only* if the legal norms that a state promulgates meet certain parameters. In order to guide epistemically, positivists hold, "there are certain conditions that the law has to meet."[22] Let us call these conditions – which positivists argue "state necessary conditions"[23] for the existence of epistemic guidance – the *attributes of guidance*.

Conventional wisdom among positivist legal philosophers holds that, in order to guide individual conduct, "there are certain conditions that the law has to meet."[24] A wide consensus holds that there are eight attributes that "state necessary conditions for the existence of [law]."[25] We need only briefly review the substance of these conditions. Laws must be (1) *general*, setting forth rules of conduct, applicable to some defined segment of the population, prohibiting or facilitating certain modes of behavior. These rules must be published, or (2) *promulgated*, so that they are available to those whose conduct they govern on a (3) *prospective*, rather retrospective, basis. Laws must avoid obscurity or unintelligibility; they must be (4) *clear*. A legal code must be (5) *non-contradictory*, in that its provisions do not conflict. Closely related is the demand that law be (6) *possible* to follow. Official enforcement of these laws should be (7) *congruent* with the laws as written. And, finally, laws should be (8) *stable*, which requires that norms not change too frequently. As a knife must be sharp if it is to cut well, so too must laws meet these criteria if they are to provide epistemic guidance.[26]

We, of course, are concerned here with the attribute of stability in particular. Ordinarily one might pause here and explicate the idea of stability in some detail. The difficulty is that positivists have remained rather coy about just what stability amounts to. Andrei Marmor captures contemporary sentiment about stability: "This requirement of the rule of law is basically a rough standard," argues Marmor, because "it would be absurd to assume that we can have a precise notion of the ideal pace of change."[27] Our inability to formalize the stability constraint seems to stem from the fact that, unlike other attributes of guidance, it is difficult, if not outright impossible, to define stability in the abstract. To characterize a law as stable seems, in some fundamental way, to involve an

---

in addition to governing those within the territorial United States, also applies to certain individuals outside its borders.

[19] Although, on certain occasions, the state does organize events, like community workshops, to help inform individuals of their new duties, rights, and obligations.

[20] *See* Shapiro, *supra* note 8, at 153.

[21] On occasion, law is used, of course, to regulate single individuals or small groups. Legislatures pass bills aimed at single individuals. *See*, *e.g.*, Jeffrey S. Hill and Kenneth C. Williams, *The Decline of Private Bills: Resource Allocation, Credit Claiming, and the Decision to Delegate*, 37(4) AM. J. POL. SCI. (1993): 1008-1031. Judicial decisions affect only the parties to the case, in the first instance. And administrative agencies adjudicate claims of single persons.

[22] Marmor, *supra* note 2, at 5.

[23] Scott Shapiro, *supra* note 12, at 395. Earlier, Shapiro observes that "regimes that flout these principles are simply not engaged in the basic activity of law." *Id*. at 394.

[24] Marmor, *supra* note 2, at 7.

[25] Shapiro, *supra* note 12, at 395.

[26] Raz helpfully compares law to a knife, in the sense that both can possess certain properties that render them, as tools deployed by individuals, more or less useful. Our question can usefully be framed: If sharp knives cut well, what sort of laws govern conduct well? *See* Raz, *supra* note 13, at 225.

[27] Marmor, *supra* note 2, at 34. It is instructive to observe that Marmor suggests that, in discussing stability, we are searching for an "ideal" pace of change. This interpretation of the attributes of guidance seems overly moralized. The items on the list explain how laws should be crafted if they are to guide individuals; they do not instruct officials how to make ideal policy.

assessment of the conditions in a society, and the expectations of the governed, in a way that the other items on the list do not. In the second half of this Chapter I will offer a conception of motivational guidance that cashes out stability in just these terms – but prior to reaching that point, we will need to investigate the conventional approach to stability.

The conventional wisdom holds that, if a legal norm meets the criteria stated by the attributes of guidance, then it provides epistemic guidance and, thus, is capable of *making a difference* to individuals' beliefs and conduct – "a difference, that is, in the structure or content of deliberation and action."[28] Epistemic legal guidance, then, and the attributes of guidance that make it possible, provide a crucial explanatory link for positivists. It explains how laws, which begin as strings of characters posited by officials, govern society.

Crucial to this explanatory chain is the idea that individuals, once they learn of their rights and obligations from law, take the content of law as reasons for action.[29] Law transmits its directives by providing individuals with things called reasons.[30] Individuals, in turn, are equipped to process and make sense of these reasons. Although philosophers dispute the exact set of processes that individuals use to process reasons – is reasoning machine-like or driven in part by emotions? – it is generally agreed that the processing of reasons involves some, most likely imperfect, process of *weighing*. Discrete considerations can be measured against each other, and added and subtracted, in order to help individuals arrive at this or that belief. Hence we often speak of a *balance* of reasons. Legal norms provide epistemic guidance when they possess the capacity to add a reason to one side or the other of the balance.

At this point one might well wonder, when we speak of "legal norms," if we are referring to all legal norms, or only a subclass of them. One could, following Joseph Raz, argue that only the legal norms governing the *processes* that produce legal rules must be stable, not the legal rules themselves.[31] Guidance, Raz argues, requires merely that "the making of particular laws should be guided by open and relatively stable general rules."[32] What distinguishes a general rule? Raz specifies that "two kinds of general rules create the framework for the enactment of particular laws," namely, "those which confer the necessary powers for making valid orders" and "those which impose duties instructing the power-holders how to exercise their powers."[33] Under Raz's framework, general rules govern the creation of the primary rules that regulate individual conduct. And it is general rules that must be stable in order that law guide conduct.

Notice, however, that it is also the case that any given process rule can be re-described as a discrete rule that applies to some legal official.[34] All process rules are discrete rules, in this sense. Thus, any argument we can run to show that stability is not a necessary condition for epistemic guidance works for both discrete norms and process rules. Insofar as stability is concerned, Raz's distinction doesn't make a difference.[35]

Moreover, instability of this sort – instability so pervasive that it affects the conventions sustaining the creation of laws – is the type of instability that undermines an entire legal system. Legal systems tend not to survive

---

[28] Jules Coleman, *Incorporationism, Conventionality, and the Practical Difference Thesis*, 4(4) LEGAL THEORY (1998): 383. We can test whether or not an individual was guided by a rule by asking about the counterfactual case: A legal rule, *P*, guides a person to do some act, *A*, only if that person "might not have done" *A* "had he not appealed to" *P* qua legal rule. *See* Shapiro, *supra* note 8, at 132.

[29] Enoch, *supra* note 1; Gerald J. Postema, *Implicit law*, LAW AND PHILOSOPHY, 13(3)(1994): 361–387; Stephen Perry, *Hart's Methodological Positivism*, in HART'S POSTSCRIPT: ESSAYS ON THE POSTSCRIPT TO THE CONCEPT OF LAW (Oxford: Oxford University Press, 2001).

[30] Pamela Hieronymi, *Reasons for Action*, PROCEEDINGS OF THE ARISTOTELIAN SOCIETY, 111(2011): 407–427.

[31] Raz, *supra* note 13, at 214-216.

[32] *Id*. at 215.

[33] *Id*.

[34] Ultimately, some official enforces the rules that govern process.

[35] Raz's reason for drawing this distinction is well taken. Raz is concerned that if we do not draw the distinction, most of administrative law will fail to be counted as law, since it is notoriously volatile. Raz singles out administrative regulations as the site at which stability of general legal processes are of particular import for the stability requirement. We will show later in the argument that one can sustain administrative law as law without resorting to the discrete rules versus process rules distinction.

such dramatic shake-ups. And legal systems, as such, are standardly defined as stable entities.[36] If Raz's distinction *does* make a difference, it proves too much. Thus, absent a compelling reason to distinguish types of legal norms, we shall use the term to refer indiscriminately to both general and primary rules.

It *is* important, however, to distinguish between legal and non-legal norms. Unsurprisingly, legal norms do not monopolize individuals' deliberations. *Non-legal* reasons compete with legal reasons for individuals' allegiance. In contemporary states, individuals receive guidance not only from legal norms, but also from customary norms, ethical norms, religious norms, and social norms. Each of these types of norms might *also* purport to regulate individual conduct.

The state assists individuals in locating the *legal* norm by adorning it with "an official marking that designates certain standards as those to which one must conform."[37] Common marks include "inscription in some authoritative text" or "declaration by some official."[38] Official marks allow citizens to identify that a norm is, in fact, a legal norm without having to engage in deep deliberation and speculation as to the pedigree of the norm. For positivists, a legal norm is valid just in virtue of certain social facts, usually certain political procedures, such as passage of bills, and the like. Adorning legal norms with official marks is designed to distinguish legal norms from other sorts of norms, and to ensure that individuals need not engage in the deep historical and conceptual tasks that would be involved in an investigation of whether a given norm is, in fact, a legal norm.[39] Official marks attempt to solve the informational problem that arises when competing norms purporting to regulate proliferate. An official legal mark conveys to the individual the idea, as far as the law is concerned, *this* norm is dispositive.[40]

Epistemic guidance, therefore, involves *presenting* individuals with the legal norm alongside whatever non-legal norms to which they adhere. In other words, epistemic guidance models a *thin* integration between the reasons created by legal norms and an individual's deliberative processes. Individuals pursue their antecedently chosen ends, and, in the pursuit of those ends, encounter legal norms. Background beliefs and desires determine whether, and to what extent, individuals comply with legal norms. Legal norms provide something of a map of possible routes for individuals, assuming that the individual knows her destination. Which route any given agent takes is a function of antecedently given desires. And the individual may choose to simply not consult the legal map at all, depending on these antecedently given desires. Epistemic guidance does not require that individuals choose their ends because of legal norms – only that individuals can make themselves aware of what the law requires.[41] This would seem to be

---

[36] Shapiro, *supra* note 12, at 65.

[37] Jules L. Coleman, The Practice of Principle: In Defense of a Pragmatist
APPROACH TO LEGAL THEORY (Oxford: Oxford University Press, 2003): 137.

[38] Scott Shapiro, *What is the Rule of Recognition (and Does it Exist?)*, YALE LAW SCHOOL PUBLIC LAW AND LEGAL THEORY RESEARCH PAPER SERIES (unpublished manuscript): 2, available at
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1304645.

[39] *See* Hart, *supra* note 13, at 95.

[40] Nonetheless, an individual might wonder why comply with the legal norm rather than some other norm. Official markers do not provide an all-things-considered reason to comply with the legal standards, but, in designating certain standards as those with which law requires conformity, official markers convey that legal reasons do not take themselves as subject to the balance of total reasons. Official markers demonstrate that legal reasons, simply in virtue of their status, not only are not subject to the balance of all reasons, but take themselves to have changed the balance of reasons in a Razian manner or do not take themselves to have changed the balance of reasons but do simply disregard the other reasons.

[41] Here is another way of putting the point. Assume an individual with a set of beliefs, B, about what the law requires. Artificially separate the individual's desires into some set of desires toward compliance with the law, D, and all of her other desires, O. Individuals' deliberations concerning the pursuit of O are constrained by the interaction of B and D. So, e.g., if the individual wants to get rich, but also wants to comply with the law, then things like robbing a bank will be ruled out. Now, in the ordinary course of things robbing a bank will also be ruled out by moral or social desires. In such a case, that the individual doesn't rob a bank is overdetermined. But let's assume that, for this individual, moral or social pressure doesn't rule out robbing a bank; the desire to comply with law is the only constraint. Say B changes, to B-prime. Now, O is constrained by B-prime. O does not change; only the content of the constraint is changed.

the case even on a Razian account premised on peremptory reasons.[42] For even if it is the case that law presents itself as a practical authority, and is "meant to replace the reasons on which it depends," it still may be the case that the citizen determines that the reputed practical authority is no authority at all.[43] By *the rules of the legal game*, legal norms purport to be practical authorities; but epistemic guidance does not require that citizens themselves conclude, as an all-in judgment, that legal rules are always authoritative.[44] If it did, it would imply that epistemic guidance is only satisfied when all citizens comply all the time. All that epistemic guidance requires is that legal norms be written in such a way as to convey their normative demands to citizens.

The conventional wisdom holds that epistemic guidance is necessary in order that individuals be made aware about what the law requires of them. In a world in which personal guidance is infeasible, individuals learn of their rights and obligations impersonally, through epistemic guidance. Because individuals' beliefs and actions are determined by weighing reasons, legal norms affect beliefs and action by creating new reasons for individuals. Thus, on this descriptive model, legal norms operate via "self-directed action."[45] In other words, individuals must apply general legal directives to their own particular situations, thus overcoming the problem of mass regulation. Law thus guides by seeking to affect individuals with reasons.

III.B. Does Epistemic Guidance Require that Law be Stable?

In order to guide by reason – in order to provide the epistemic guidance crucial for governing mass populations – must legal norms, in fact, meet the criteria set forth by the attributes of guidance? Given this picture of epistemic guidance, which mediates between laws on the books and individuals' deliberations and actions, is it, in fact, true that the attributes of guidance state necessary conditions for epistemic guidance? More formally, we may ask of any of the eight attributes:

What state of the world does epistemic guidance require such that some attribute, *X*, is necessary, assuming the operation of some other set of attributes, *Y*, where $X \notin Y$?

We answer this question by posing a counterfactual. An attribute is necessary for a conception of legal guidance when the absence of that attribute renders law incapable in principle of guiding individuals. The idea here is that there are certain states of the world that must obtain in order that legal guidance be possible, and the items constituting the attributes of guidance either uniquely cause those states or are necessary for those states to exist. Posing a counterfactual in which legal norms do not possess the attribute in question allows us to test whether guidance is possible in its absence. Applying this framework to the case of stability, we can generate two cases: one in which a legal norm is stable and a counterfactual case in which a legal norm is unstable. We then compare how these two cases affect epistemic guidance.

First, imagine a successful case of epistemic guidance, in which the applicable law, *P*, is stable. Say that Alice wants to invest some money in a stock that will grow in value over time. Alice is indifferent between stocks, so long as the one she picks will grow in value. One investment option is a company's stock that Alice read about in the newspaper. Another option is a company's stock that her friend, who possesses insider information, told her would grow in value because it is going to be acquired by another company shortly. Prior to consulting the applicable law, Alice is indifferent between the two options, since all she wants to do is invest in a money-making stock. However, because Alice is motivated to comply with law, she seeks epistemic guidance. Alice first seeks indirect guidance – from her associates, secondary sources, and so forth – and receives a uniform set of reports as to the content of the law. They tell her that trading stocks on insider information is illegal, and could land her in jail. Just to be sure, Alice reads the applicable law and applicable judicial opinions, and finds that her indirect guidance is supported by the direct guidance. One alternative, investing in the stock she read about in the newspaper, is within

---

[42] *See* Joseph Raz, *Authority, Law and Morality*, 68(3) THE MONIST (1985): 295-324. I am very grateful to an anonymous reviewer for raising this objection.

[43] *Id*. at 297.

[44] As Raz writes, "No blind obedience to authority is here implied. Acceptance of authority has to be justified …" *Id*. at 299. Epistemic guidance, for Raz, requires that law "must be capable of guiding the behavior of its subjects. It must be such that they can find out what it is and act on it." Raz, *supra* note 13, at 214.

[45] Postema, *supra* note 28, at 369.

the bounds of compliance, and the other alternative, trading on insider information, is prohibited. Alice chooses the legal alternative, and avoids jail.

Our successful case of epistemic guidance provides a baseline against which we can pose a counterfactual. Can we alter the case such that instability interferes with the transmission of legal content, thereby impairing epistemic guidance? It is difficult to imagine such a case. Even if the contents of *P* changed at random intervals, Alice could, in principle, locate the content of the law and thus receive epistemic guidance. This conclusion follows from the definition of epistemic guidance. No matter how volatile a discrete legal rule, if it is duly enacted, then it bears law's official mark. And the fact that a norm bears law's mark implies that the informational problem about which norm is authoritative has been solved, and epistemic guidance is, in principle, possible. Duly-enacted laws are published, of course, thanks to the promulgation requirement. Laws altered in any random interval are still subject to promulgation, and the fact of promulgation ensures that individuals can in principle locate the content of the law, which is all that epistemic guidance requires. Epistemic guidance simply does not require any state of the world for which stability is required. Contra the conventional wisdom, then, it seems that no level of instability interferes with the capacity of legal norms to provide epistemic guidance.

If our account thus far is correct, the conventional wisdom, which holds that stability is a necessary attribute of epistemic guidance, is incorrect. Where did those accounts go wrong? We need an error theory. Here, we will explore why it is commonly believed that epistemic guidance requires stability. The culprit is that instability raises the costs of epistemic guidance. Instability makes it more difficult to locate the law, thus making it more difficult to comply with the law. On this line of thought, instability sets off a chain reaction that ends with decreased compliance.

This line of thought dovetails with the most common way to justify stability as an attribute of guidance. As earlier noted, the most common way to do so is on grounds that, if law is volatile, individuals will not rely on it. This claim, or some version of it, might well be true. The problem with this argument is that epistemic guidance does not require increasing individuals' reliance on law or otherwise providing motivational guidance. It does not require legal officials to cultivate reliance or that individuals be disposed to rely. If the attributes of guidance *are* conceptualized as attributes that tend to increase guidance or compliance, then it would make sense to justify stability on reliance grounds. The difficulty here is that if we conceptualize the attributes of guidance in this way, then our list of eight attributes is, and always was, woefully incomplete. If the standard is attributes that tend to enhance compliance, then surely Pareto optimality, or something along these lines, is missing from the list.

Nonetheless, one might still be worried about the possibility that high volatility makes epistemic guidance more difficult, thus rendering compliance less likely. This worry is not altogether misguided. Even if stability is not required for epistemic guidance, it might possess a contingent, yet predictable, relationship to compliance and thus law's efficacy.[46] Sketching out such a contingent relationship does help to explain the error, and explain why we might care about stability even if it is not a property of legality. The key to sketching out this contingent relationship between instability and reduced compliance is to treat legal norms as costly information.

III.C. Legal Norms as Costly Information

Ideally, locating legal norms would be frictionless. The environment in which legal norms operate, however, ensures that locating law's content is rather complex. In contemporary states, individuals receive guidance not only from legal norms, but also from customary norms, ethical norms, religious norms, and social norms. Each of these types of norms purports to regulate individual conduct. Laws can only claim primacy over other types of norms if they make themselves known. Earlier, we observed that official marks *attempt* to solve the informational problem that arises when competing norms purporting to regulate proliferate. Here, we show that official marks are *signals* from the state to citizens. As signals, the information they provide is not always perfect. It is sometimes noisy, especially when laws are unstable. This situation raises the costs of obtaining legal information, and thus can lead to reduced compliance.

---

[46] Ultimately, the level of compliance sought by legal officials is a function of the regulatory regime. Thus, higher or lower search costs might be tolerated. We cannot answer that in the abstract, so we cannot specify how much non-guidance is compatible with a regulatory program. We can only observe that there will be, in the normal case, some desired level of guidance and compliance.

One way in which individuals might determine if a norm is, in fact, a legal norm is to engage in a historical inquiry into the pedigree of the norm. This historical inquiry would involve questions like, *Did Congress pass this norm?*, *Does the Constitution forbid norms of this sort?*, *Has any court modified the norm?*, and so forth. Such investigations, however, are likely to result in confusion. Instead, we need a proxy or a signal to stand in for all of the historical investigation that goes into determining if a norm is, in fact, a legal norm. The best proxy or signal of legality is that the norm "bear the mark" of legal authority. When a legal norm bears the mark of legal authority, it announces to individuals that it possesses the proper pedigree.[47] In theory, official marks guide individuals by signaling those norms, and only those norms, that are sanctioned by the state. As a signal, however, it is liable to convey incomplete or erroneous information.[48] In such cases, epistemic guidance suffers.

Signals can be erroneous because inscriptions in authoritative texts do not *stop* signaling if a law is superseded. A superseded statute does not report that it has been superseded. A superseded statute does not stop bearing the mark of law once it has been superseded. Ideally, we would devise a technique to signal the content of the law that could update itself. Official marks as they currently exist, however, are not self-updating.[49]

One might be tempted to argue that a superseded statute bears an invalid mark. But marks are not valid or invalid. Rather, they are accurate or inaccurate, but there is no way of discerning whether they are accurate or not from the mark itself. Thus, every change in the content of a law results in the creation of some erroneous signals, unless inaccurate signals are immediately destroyed. Frequently changing legal norms results in a proliferation of inaccurate signals. If, for instance, an official report of the law is printed once every year, and a given legal norm is changed every year for 10 years, then there will exist nine superseded reports, unless they are destroyed. These reports are inaccurate signals of the content of the law.[50] Instability thus recreates the informational problem that official marks were intended to solve. Superseded marks are, in essence, competitor norms. They state norms that conflict with the state's position on what counts as compliance with the law.

In addition to the possibility that an official mark provides erroneous information, official marks might also, or alternatively, provide *incomplete* information. Although a legal norm contains much of the information that is necessary to comply with law, it does not always contain all of the information required for compliance. *That* information is located in interpretations of the norm, in judicial opinions, and the like. In other words, an official mark is partial when it points to a string of text, and refers to that text as law, when, in fact, the text in question only conveys *part* of the information an individual needs to know in order to comply with the law. In the United States, many official marks are partial, because some legal content is contained in judicial opinions. For instance, much of the content of the Eighth Amendment's prohibition on "cruel and unusual punishment," is contained not in the text of the law but, rather, in the text of judicial opinions.

In theory, official marks, or signals, point to a string of text, and tell the citizen: This string of text contains all of the information you need in order to comply with the law. In practice, however, these official marks are partial. They convey only some of the information needed to comply with the law.[51]

---

[47] One might be tempted to avoid taking the term "mark" literally, since, at base, to be lawful simply means to have the correct pedigree, to be a duly enacted law. But the official mark is not equivalent to (or even a logical extension of) the fact that a norm is duly enacted. The mark is a proxy designed to solve informational problems and decrease deliberation. Only duly enacted norms receive official marks, but we can imagine a case in which a duly enacted law does not, for whatever logistical shortcoming, receive a mark but is nonetheless law. A law that is not published would not receive a mark, for instance. If the official mark is to solve an informational problem about which norms individuals are supposed to comply, then we must take the idea of a mark literally.

[48] The legal philosophy literature has not addressed these shortcomings with official marks. It has treated the idea of official marks quite generously.

[49] It is worth observing that the technology is readily available to make official marks self-updating, at least with regard to those official marks published electronically. A simple piece of code embedded in the metadata of the standing law might well be able to provide self-updating signals.

[50] Inaccurate, that is, assuming that the content of the change is not a reversion to some past norm.

[51] *Cf. Appalachian Power Co. v. E.P.A.*, 208 F.3d 1015, 1020 (D.C. Cir. 2000).

If instability increases the erroneousness or incompleteness of official signals, it increases what we can describe, following Jules Coleman, as *search costs*.[52] If these search costs are high enough, the efficacy of a particular legal norm is reduced. Efficacy is a necessary condition of legal systems, although not itself an attribute of legal norms. This line of thought does not show that stability is an attribute of legality, but it may help us to identify the point at which the costs of finding the law become such that compliance is compromised to the point that legal norms are rendered inefficacious.

Volatility can increase the costs of locating the law, thus ramping up the costs of compliance. Volatility can artificially reduce compliance by making it costlier than *standing* law would imply that it is. Assume some "natural" rate of compliance with a given law. The rate of compliance is natural in the sense that it is a function of individuals' background beliefs and desires. Volatility of the sort we have described here reduces compliance below this natural rate, even though the content of the standing norm is the same.

It's possible that instability creates costs sufficiently high that they deter individuals from locating valid law. Locating valid law in the face of instability can take time and effort, and can even result in individuals mistakenly acting on what they believe to be legal reasons. An example will illustrate how these costs could mount under non-ideal conditions.

Imagine a variation on Alice's successful case of epistemic guidance. Assume that Ben, acting in good faith, seeks epistemic guidance. The content of the applicable legal norm is highly volatile, but Ben is unaware of this fact. Ben seeks indirect guidance by consulting secondary sources, which provide mixed reports. Some treatises report that the law permits only option *P*. Reports from those in a related industry indicate that *P* and *Q* are permitted. Sensing that something is amiss (Ben is not a legal official, but is aware that the law is supposed to be clear and predictable), Ben seeks direct guidance from what he believes to be an authoritative legal text. But because the law is volatile, Ben encounters a superseded statute, which, of course, does not report that it has been superseded. The statute bears the mark of law, and indicates that both *P* and *Q* are permissible. Because *Q* is the preferable option all things equal, Ben adds a consideration to his stack of reasons and acts accordingly. In fact, Ben erroneously relied on a superseded statute. The valid statute prohibits *Q*. So Ben fails to comply.

The costs individuals are willing to incur in their search for epistemic guidance presumably have limits. To be sure, the absence of epistemic guidance does not imply a lack of compliance in every case. It might well be the case that most individuals can comply without guidance. But over the long haul compliance suffers when guidance decreases. And when volatility increases, compliance decreases because individuals prefer to save the costs to time and effort. If, at a certain point, the costs become high enough and the content of the legal norm is non-intuitive enough that individuals cannot figure out the compliant behavior absent the norm, compliance could decrease to such a degree that the system can no longer be described as efficacious.

A second way in which search costs can impair epistemic guidance is when individuals erroneously rely on a superseded official mark. Here, individuals desire to comply, and believe that they are complying, but in fact fail to comply because instability decreases the strength of the signal sent by officials to individuals. Unlike cases where individuals decide to stop the search due to the costs of searching and bite the bullet on compliance, here individuals believe that they have been guided and are complying.

In our hypothetical, for instance, Ben *took* himself to be relying on a reason created by law. In fact, he erroneously relied on a superseded official mark. In such cases, compliance suffers not because of a decision that is too costly, but, rather, because volatility leads to a situation in which individuals' *operative* reasons – the reasons on which they, in fact, act – are different from the reasons that do, in fact, *justify* the action from the perspective of law. The mistaken non-compliance induced by volatility raises the question of whether Ben was epistemically guided. We can see that there was no reason for Ben to act in the way that he did, yet it seems as though Ben acted for a reason.[53]

III.D. Three types of reasons

---

[52] Jules Coleman, RISKS AND WRONGS (Oxford: Oxford University Press) (2002): 125-131.
[53] *See, e.g.,* Jonathan Dancy, PRACTICAL REALITY (Oxford: Oxford University Press) 2002.

Was Ben epistemically guided when the reason on which he acted was, in fact, not a reason? By way of analogy, consider a game of chess between Beginner and Expert.[54] If, say, Beginner is considering whether to move a bishop vertically, he might consult the rules of chess and find that bishops are only allowed to move diagonally. The rules of chess tell Beginner that moving the bishop diagonally (or at least omitting to move it non-diagonally) is an appropriate action. The rules *justify* the action.

It turns out that although moving the bishop diagonally was a valid move, it was not a smart one, and Expert has placed Beginner in a very precarious situation. Beginner sees that his one way out of Expert's trap is to undertake what he believes to be a "castle." Beginner attempts to think through the rules of castling, wondering if he may do so despite the fact that he had already moved his king. Beginner mistakenly concludes that the rules of chess do, in fact, allow him to castle despite already having moved his king. Here, Beginner acts for what he took to be a reason, but, which, under the rules of chess, is not actually a reason at all. We can call such reasons *operative* reasons.

Beginner erroneously believed he could castle after moving the king, due to his lack of familiarity with the rules of chess and a general sense of anxiety. These reasons *explain* Beginner's choice. By introducing a distinction between considerations grounded in a normative system (justificatory reasons) and considerations that an agent took to be grounded in a normative system (operative reasons), we can see why Ben was not epistemically guided by the superseded statute.

Volatility may induce individuals to mistakenly believe that they are being epistemically guided by the legal system when in fact they are not. Ben, like Beginner, *sought* to comply with the rules. The non-compliance in both cases resulted from the complexity of the rules relative to the individuals' capacities to comply. And while we can now see how to avoid the awkward implication that Ben was in fact guided by an invalid legal rule, the practical worry that volatility decreases guidance remains.

We have our error theory, then. We have identified why one might be worried by unstable legal norms, even if stability is not an attribute of legality. The claim that stability is necessary appears reasonable if one improperly assumes that stability is necessary in order to maintain the efficacy of the legal system, rather than to facilitate legal guidance. The core concern is that volatility reduces the strength of the signal created by law's official mark. Without a strong signal, ordinary folk will have difficulty locating the content of the law, which leads to a decline in epistemic guidance, which reduces the efficacy of the system. This is a serious concern, yet it does not explain why stability is necessary for law to guide conduct. We are thus left with a dilemma.

This dilemma is this: We must reject *either* the claim that stability is an attribute of guidance *or* the claim that epistemic guidance is an accurate account of legal guidance. Stability is widely thought to be an attribute of guidance, and epistemic guidance is the most prominent conception of legal guidance. Perhaps the way out of this dilemma is to modify our conception of legal guidance. Thus, we will consider an alternative conception of legal guidance. On *this* conception, stability is, in fact, a necessary attribute of guidance.

IV. Integrated guidance

Integrated Guidance ("IG") reconceptualizes law's guidance function.[55] IG treats guidance as possessing both epistemic *and* motivational dimensions. In addition to providing epistemic guidance, law motivates individuals by disposing them to comply with it. IG models individuals as planning agents with the capacity to incorporate legally sanctioned means in their plans. Law guides when its norms, owing to their status as dispositive settlers of normative controversies, exert pressure on individuals to use the norms' content as the bases of plans. If the goal of individual deliberation is to efficiently coordinate a large number of beliefs and desires, legal norms should outperform any other type of norm, since no other type of norm promises to be dispositive of controversy. Broadly speaking, this process can occur via either of two routes: filtering and facilitating. Legal norms that prohibit certain behavior filter out unsanctioned means from individuals' plans. Legal norms that facilitate conduct offer individuals opportunities to employ sanctioned means to achieve their ends. The content of legal norms filters out some means,

---

[54] Here the account follows Pamela Hieronymi, *supra* note 29.

[55] We find strong hints of IG-like positions in the literature, most notably in the work of later Scott Shapiro, but also in the work of Postema and Raz.

and creates others. The result is a legally sanctioned set of means from which individuals can choose. If individuals create higher order intentions structured around this legally sanctioned set of means, then they will be more disposed to comply with the law. Because plans require a certain degree of stability, the means by which individuals accomplish those plans must also possess a degree of stability. But we cannot simply leave things there, since individuals will have different preferences over reconsideration. We will attempt to sketch how much specificity is required, given a population with certain preferences over reconsideration.

IG is most naturally modeled around Michael Bratman's planning conception of individual agency.[56] Very roughly, the idea is that individuals possess certain desires, or outcomes that they hope to achieve. Most of the time, these desires are not immediately satisfied; certain intermediate steps need to be taken to render their realization more likely. The planning conception of agency explains how we structure means and ends together to help us to achieve our desires. Building on the planning conception of agency, IG explains how, given the way in which individuals ordinarily structure their means and ends, legal norms guide behavior. Individual action is primarily governed by a set of hierarchical plans, to which one is non-trivially committed, and by which one is guided in the making of more specific decisions. The existence of such plans is essential, because we all face severe resource limitations: limitations of time, energy, cognitive capacity, and so forth. And, yet, despite these limitations we nonetheless must coordinate a large set of diverse activities. Plans are special sorts of individual intentions that help to overcome these limitations.[57] Individuals weave plans of various levels of generality and complexity together in order to create an agenda for future action. Rather than lurch from decision to decision, unguided by any background agenda, individual action is designed to bring about the goals contained in the plans, and do so efficiently. Inefficient means, or impossible means, are "filtered out."[58] Individuals are left with plans that exert strong, although not absolute, control over any given decision that an individual makes. Plans can be reconsidered, as we shall shortly see, but IG treats reconsideration as the exception, not the rule. In case an individual does reconsider a plan, she might reach for a new plan from her "plan library," or a set of beliefs about which actions bring about which effects under a given set of circumstances.[59]

Because plans by their very nature strive to reduce the costs of accomplishing one's ends, impossible, or inefficient, means are filtered out. This is the feature of individual agency that allows law to dispose individuals toward compliance. Those subject to law tend to believe that (a) purposive agents, namely, legal officials, (b) create legal norms addressed to them (c) in order to settle questions of how to act.[60] If an individual has these three beliefs, legal norms will occupy a privileged place in an individual's deliberations.

Because such agents accept that a hierarchical relationship exists between themselves and the state, such agents take the content of the law to provide plans for their own actions and seek to ensure that other plans are consistent with legal norms.[61] The content of legal norms thus has the capacity to induce rational agents to integrate their conduct within the parameters of the law because such agents *take* the law to be an authority within its domain. Law guides when its norms, owing to their status as dispositive settlers of normative controversies, exert pressure on individuals to use their content as the bases of plans. Law is designed to settle normative debates. If law does what it claims to do, it will provide a stable foundation for derivative plans. If the goals of a deliberative system are to efficiently coordinate a number of desires, legal norms should outperform any other type of norm, since no other type of norm guides with the backing of the state. Individuals thus desire to incorporate its content into their plans. This process of incorporation can occur via either of two routes: filtering and facilitating.

V.A. Filtering and facilitating

---

[56] Michael E. Bratman, INTENTIONS, PLANS, AND PRACTICAL REASON (Cambridge: Harvard University Press) 1987; Shapiro, *supra* note 12.

[57] Michael E. Bratman, *Planning and the Stability of Intention*, MINDS AND MACHINES, 2(1992): 1–16.

[58] Michael E. Bratman, David J. Israel, and Martha E. Pollack, *Plans and Resource-bounded Practical Reasoning*, COMPUTER INTELLIGENCE, 4(1988): 350.

[59] *Id*. For instance, if Alice possesses a plan to travel to a party at 9 p.m., then she can filter out all options that are incompatible with the possibility of attending the party, such as planning a dinner for 9 p.m. Faced with a smaller set of alternatives, Alice can more easily focus on the relevant questions: how to get to the party, and so on. And when 9 p.m. arrives, Alice will, in fact, travel to the party, absent special circumstances.

[60] Shapiro, *supra* note 12, at 200.

[61] *Id*., at 141.

Legal norms that prohibit certain behavior filter out possible means for the accomplishment of desires. For instance, if an individual has a desire to "get rich quick," criminal prohibitions on insider trading might filter out the means of, inter alia, buying or selling securities on material, nonpublic information in breach of a fiduciary duty.[62] By filtering out certain possibilities, non-filtered possibilities face fewer competitor means. Legally sanctioned means are, in a sense, elevated as non-legal means are winnowed away.

Yet, as Hart famously observed, legal norms do more than prohibit conduct: they also create and facilitate new forms of conduct. By making available new forms of action, legal norms create new means by which individuals can accomplish their ends. Although it is odd to speak of "compliance" with legal norms that facilitate conduct, it is not odd to imagine that legal officials in some sense favor the use of such ends over the use of non-legal ends. It makes sense, then, to conceptualize guidance as disposing individuals to employ legally-created forms of action, even if the language of compliance is inapposite.

The content of legal norms filters out some means, and creates others. The result is a set of legally sanctioned means. If individuals create higher order intentions structured around this legally sanctioned set of means, individuals will be more disposed to comply with the law. Law influences dispositions by influencing the structural framework of – the higher level plans – individual deliberations. This dynamic has the effect of crowding out alternative sources of planning norms. Recall that legal norms compete against other norms to guide individual conduct. One of the primary competitive advantages of law is that it usually has the state's force behind it. Nonetheless, as the fact of non-compliance shows, legal norms do not always claim victory. IG shows how legal norms stay competitive by reducing the foothold that alternatively-sourced norms might have in individuals' planning mechanisms.

Raz hints at such an idea in his discussion of stability. ``[O]nly if the law is stable are people guided by *their knowledge of the content of the law*," Raz argues.[63] Implicit in Raz's argument is the idea that, in the absence of legal guidance, individuals will plan according to alternatively-sourced normative standards. If law is unstable, individuals are guided by their knowledge of *some other* norm. Legal norms, though, can crowd out competing norms, and thus make them less probable to influence individual behavior. By creating a set of legally sanctioned means, law disposes individuals toward creating higher order intentions that are compliant with law. The more individuals craft their intentions in a legally compliant manner, the more dependent they become upon legally compliant behaviors. As with many systems engineered to produce certain outputs, the more engaged one is with a system, the more dependent one becomes on that system, the harder it becomes to exit that system. And so it is with law.

Plans allow individuals to coordinate vast swaths of desires in an efficient manner. They are efficient solutions in part because they allow individuals to incur the costs of weighing options, deciding on means, and coordinating those choices with other of their plans and other individuals. It thus seems as though plans only efficiently structure decision-making if they are relatively stable. After all, "if we were constantly to be reconsidering the merits of our prior plans they would be of little use in coordination and in helping us cope with our resource limitations."[64] "Nonreconsideration," in other words, "will typically be the default," even in cases where, were the individual to reconsider the decision, she would plan differently.[65]

Under IG, legal norms must be capable of forming the basis of plans through legally-sanctioned means. And plans, we just saw, are in the normal course of things stable. Legal norms, then, too, it would seem, must be stable. Predicating plans on unstable means seems like a recipe for incurring the costs of reconsideration. The role of stability under IG is, at bottom, about the way in which reconsideration affects individuals' ability to construct plans around legally-sanctioned means. Plans are premised on the existence of certain conditions. When the conditions upon which a plan is premised change, an agent may reconsider her previously formed plans, but she need not. Non-

---

[62] Such prohibitions may also filter out certain *ends*. One might be enticed by the thought of a life of crime, for instance. Our focus on means here should not be read as denying that prohibitions also filter ends.

[63] Raz, *supra* note 13, at 229 (emphasis in original).

[64] Bratman, *supra* note 56, at 3.

[65] *Id*.

reconsideration is the default position, but we can sketch with slightly more specificity the conditions under which individuals override this default.

Changes in the world exact one of two types of costs in relation to individual plans. On the one hand, changes can exact costs related to the agent's deliberative structure. Reconsideration of plans premised on changes takes time, energy, and inhibits coordination with others. If an agent prefers to avoid those costs, he might incur a different type of cost: the costs of attempting to bend a changed world back to a state in which it is compatible with his original set of plans. For all individuals interested to keep their plans and the state of the world in some sense connected, change exacts one or the other of these types of costs. Individuals choose different strategies to manage these costs.[66] At one end of the spectrum, individuals' plans are only as strong as the set of beliefs and desires that formed the original basis for the plan. If the agent's plans are non-robust in the face change, then the agent will frequently be attempting to bring his plans in line with the changed state of the world. The costs typically borne of such an agent include lost efficiency gains, since he expends resources recalculating means and ends, and lost coordination gains, since others that rely on his previous plans will no longer reliably do so.[67] Such plans will struggle to serve the usual purposes of a plan, since they filter few actions out and ask the agent to continually deliberate over new options. Nonetheless, some agents may hold such a view, at least about some matters.

At the other end of the spectrum, plans possess an "intrinsic stability," which commits the agent to attempt to bring about their satisfaction conditions, come what may.[68] If an agent never reconsiders his plans, then volatility will impact the agent in a characteristic way; namely, the agent will be forced to attempt to bring the world in line with her pre-established plans. The costs typically borne of this type of agent include those involved with the open-ended challenges of attempting to make whatever aspect of the world is no longer compatible with his original set of plans once again compatible. Outside of these two poles, individuals tolerate varying levels of reconsideration. This variance need not be considered the product of irrationality. Some individuals may be quite adept at shouldering the costs of reconsideration, while others may be skilled at bending the world to their wills.

IV.B. Deliberative Types

We see, then, that individuals prefer to allocate change costs in different ways. This allows us to imagine a spectrum of approaches to the allocation of change costs.[69] We will identify two intermediate types, producing four types in total.[70] We can label these types as follows:

1. Non-Robust: Prior plans are always subject to reconsideration, given a relevant change.
2. Weakly Robust: Prior plans are subject to reconsideration, given a relevant change, only if the anticipated costs of reconsideration are less than the anticipated benefits of reconsideration, with regard to a specific action.
3. Strongly Robust: Prior plans are subject to reconsideration, given a relevant change, only if the anticipated costs of reconsideration are less than the anticipated benefits of reconsideration, with regard to an agent's overall dispositions about reconsideration.
4. Intrinsically Stable: Prior plans are never subject to reconsideration, no matter the change.

The existence of deliberative types complicates IG's account of the value of planning. Recall that future-direct intentions, or plans, allow us to deliberate more efficiently and coordinate intra- and inter-personally.[71] It would seem to follow that an agent will not adopt or will tend not to adopt a plan unless the plan facilitates efficient deliberation and intra- and inter-personal coordination. Any plan, the content of which is likely to induce inefficient

---

[66] Bratman, *supra* note 54.
[67] For instance, consider the deliberative difficulties facing an agent with a higher order plan grounded in act consequentialism, or some other set of values that requires that an agent constantly calculate the value of her action.
[68] *Id*. at 11.
[69] This section draws heavily on Bratman's discussion of stability of intentions. *See* Bratman, *supra* note 56. Any given individual might be a certain type with regard to one issue, but a different type with regard to another issue, since the costs of being one type or the other will vary based on the issue and individual capacities.
[70] We will assume that individuals know their deliberative types, and find it rational to be that type. If an individual knows his deliberative type, and the costs associated therewith, he will tend to structure his plans in such a way as to bring about the satisfaction of rational desires.
[71] Bratman, *supra* note 55.

deliberation or inhibit coordination is likely to remain in an agent's "plan library" of possible, but unadopted, plans.[72] Given an agent's knowledge of her type, she will rule out some plans in the plan library on grounds that they are subject to too much volatility. Given a potential plan from the plan library, there is some threshold of probability at which the agent, given her tolerance for reconsideration, will reject the plan. Given the choice of plan *A*, there is some probability, $\Pr(A)$, that, out of the set of possible future events, at least one event that forces the agent to reconsider will materialize. When that $\Pr(A)$ reaches the threshold that the agent sets based on her deliberative type, that potential plan is rejected.

Strongly Robust and Intrinsically Stable types may be wary of incorporating plans that they anticipate will be subject to volatility that destabilizes those plans. This wariness may result from the fact that they tend not to reconsider plans. Bending the world to one's plans can be difficult, even where one prefers that course of action to reconsidering her plans. Of course, some changes in the world may induce such high costs that reconsideration is likely even for Strongly Robust and Intrinsically Stable types. As a percentage of all changes, however, these cases are likely to be low, given these types' policies about change.

Strongly Robust and Intrinsically Stable types, precisely because they make it a policy not to reconsider, may plan for contingencies. If this is so, then these types may, in fact, *not* be considerably more wary of potentially unstable plans. Yet, contingency planning is costly, just as buying insurance is costly. Thus, it will be avoided where possible. In general, then, we can hypothesize that Strongly Robust and Intrinsically Stable types will possess a comparatively low volatility threshold. In other words, for these deliberative types, a comparatively low level of volatility will cause them to leave plans in their planning libraries, rather than adopt them. For instance, say an Intrinsically Stable type is considering whether to buy a piece of land on which sits a large lake, for the purposes of fishing. An assessment of the regulatory climate, however, suggests that fishing from lakes such as this one might soon be prohibited. The question is, given our individual's policy of not reconsidering plans, our individual decides not to buy the property. It seem as though, because of his policy to not reconsider plans, our individual will decide not to buy the property. But, as an Intrinsically Stable type, he might be comfortable incurring the risk and attempting to bend the regulatory process to his preferences. For symmetric reasons, Non-Robust and Weakly Robust types will possess a comparatively high volatility threshold.

Law's capacity to guide is typically thought of as a capacity that extends uniformly over all individuals. IG requires a differentiated approach to guidance, since it is premised on the idea that individuals possess different tolerances for the risks associated with volatility. By introducing Deliberative Types, we have introduced differences in the population. We must find some way of discussing guidance as a capacity possessed by law while also taking account of these individual differences. The solution requires analysis of the dynamic and cooperative character of legal guidance.

The state faces a limit on how often it can change the content of these legal norms, if it desires individuals to use them as part of their plans. Given a potential plan from the plan library, there is some threshold of probability at which the agent, given her tolerance for reconsideration, will reject the plan. This probability threshold, however, is not permanently fixed or wholly independent of individuals' wills, as we have thus far imagined.[73] Even if we hold one's deliberative type constant – even if one maintains that an individual's deliberative type is beyond that individual's capacity to change – the individual can, nonetheless, structure his affairs such that the possibility of legal change is less likely to reach the threshold beyond which the legal norm is perceived to be too volatile to form the basis of a plan. In other words, individuals can plan for legal change itself. But why do so? Purchasing insurance, after all, is costly. Strongly Robust and Intrinsically Stable types, precisely because they make it a policy not to reconsider previously established plans, may choose to plan for certain sorts of contingencies. If this is so, then these types may *not* be considerably more wary of potentially unstable plans. However, all things equal, such planning is costly. Why incur these costs?

---

[72] *Id.*

[73] Whether an individual is of a certain deliberative type is, in some sense, a function of that individual's desires. If an agent really is committed to act consequentialism, then it makes sense to treat prior plans as always or almost always subject to reconsideration. But whether an agent is in fact committed to act consequentialism is itself subject to reconsideration. Likewise, individuals can seek to reject fewer legal norms on grounds that they are insufficiently stable. An individual can lower the probability threshold at which he rejects possible legal norms due to its likely volatility if he moves closer to the Non-Robust type.

IV.C. Legal Guidance as Assurance Game

We can explain why individuals incur these costs by observing that the state's provision of legal norms takes the form of an assurance game. Both the state and individuals benefit when the state responds to widely held problems *and* individuals are disposed to adopt the regulations. But both of these endeavors are costly, if unreciprocated. One might hypothesize, then, that legal change will tend to occur when both sets of parties are sufficiently assured that the other will cooperate.

Let us assume that the state faces a choice between enacting a legal norm that is responsive to some widespread problem. The state can either change the content of the law or not change it. Citizens, meanwhile, can choose to plan for volatility or not plan for it. Of course, some individuals, namely, those Non-Robust types, need not plan for volatility in order to cooperate. They are cooperative by their very type. Outside of the special case of Non-Robust types, however, the actions of the state and individuals are interdependent. The state is better off changing legal norms in response to widespread problem if and only if individuals tend to plan for volatility. Likewise, individuals are better off incurring the costs of planning for volatility if and only if the state enacts responsive legal regulations. Otherwise, they have spent resources – deliberative and perhaps material – preparing for legal guidance that never arrives. We can represent legal guidance as an assurance game:

		Individuals	
		Change	Not-Change
State	Change	5, 5	-5, 0
	Not-Change	0, -5	0, 0

In order that the state be willing to change a legal norm, it likely anticipates that individuals will be guided by it. As we observed earlier, there is some volatility threshold above which individuals of a certain deliberative type will not incorporate a legal norm into their planning apparatus. Individuals encourage the state to enact new regulations when they *raise* that threshold.

An unreciprocated attempt at cooperation is worse than maintaining the status quo. This is true of both the state and its subjects. If individuals believe that the state cannot or will not enact responsive regulation, it is preferable not to incur the costs of planning, all other things equal. Likewise, the state will seek to avoid attempting to provide legal guidance if it is likely to fail.

The dynamic conception of stability incorporates the possibility that individuals will *raise* their volatility thresholds. Although, as we observed, raising one's volatility threshold is costly, we have now seen why individuals may be inclined to do so. Individuals cooperate with the state by planning for beneficial legal change. Where this is true, volatility thresholds decrease. If so, then we have a solution to our problem of differentiated guidance: volatility thresholds are no longer so differentiated. If individuals plan for legal change, their volatility thresholds clump together, thus providing officials with a target range of stability that must be met. And, thus, we have derived an answer to the question of why stability is a necessary condition in order that law guide individual behavior. The degree of stability necessary in order that law guide individuals is just that degree required to meet the volatility threshold of some significant portion of the population.

V. Conclusion

The conventional wisdom among positivists holds that instability generates uncertainty and this uncertainty *itself* impairs guidance. This Chapter rejected that claim, and suggested, instead, that individuals' *expectations* about the volatility of legal norms is far more important to legal guidance. If individuals anticipate volatility, or have non-legal reasons to prefer volatility, then unstable legal norms will pose few problems.

Raz argues that stability requires "that the making of particular laws should be guided by open and relatively stable general rules."[74] The concern is that instability produces uncertainty, which results in individuals

---

[74] Raz, *supra* note 13, at 213.

failing to guide their conduct by law. When this is not the case, Raz argues, it becomes "difficult for people to plan ahead on the basis of their knowledge of the law."[75] Our analysis suggests a more complicated picture. Whether individuals are able to plan ahead on the basis of probabilities about what the law will be depends on their general deliberative type and the level of resources that they have devoted to the issue. If the issue matters to individuals, they will devote resources to understanding the probabilities that the law will change; they will take out insurance to compensate against change; and they will seek to push the law in one direction or the other. What's more, as we have shown, individuals can recognize that the state's responsiveness to issues can be a public good, and respond to this realization by planning for legal change.

As one of the primary tools of democracy, law is deployed to respond to problems. The sorts of problems that contemporary democratic regimes solve do not arrive in neat intervals; they are often significant and difficult, and the solutions partial and temporary. Instability is high because long-term solutions are technically or politically impossible. To demand stability in the name of legal guidance is to effectively curtail the process of problem solving. Because it is usually preferable that the state be flexible in its approach to problem solving, it is important to know just how much instability is compatible with legal guidance. This Chapter has tried to move past the simple refrain that law must be stable, so common in the legal philosophy literature, and explain just how much stability is, in fact, compatible with legal guidance.

---

[75] *Id*. at 216.

CHAPTER FIVE: STEPS TOWARD REESTABLISHING ACCOUNTABILITY

I. Introduction

So far, this dissertation has explained how contemporary practices *undermine* electoral accountability. In this Chapter, we turn to ways of reestablishing accountability. This Chapter does not provide a complete theory of electoral accountability. Rather, the goal of this Chapter is to point to some general directions that a theory of contemporary electoral accountability might take. The Chapter makes three primary recommendations, corresponding to each of the three critiques of the traditional model of electoral accountability.

First, this Chapter uses the findings that individuals under-blame elected officials to reject what is the conventional view about administrative accountability – the derivative liability theory of accountability. On this approach, agencies are held accountable by holding elected officials accountable. In other words, the derivative liability approach depends on the traditional model of electoral accountability. We have shown, however, that individuals under-blame elected officials. Thus, if derivative liability is to remain a valid approach to administrative accountability, we will need to find ways to de-bias individuals' perceptions of blame.

Second, this Chapter uses the findings that individuals under-blame agencies for nudging (as compared to coercing via law) to propose new review processes to mitigate the under-blaming. Reviewing the primary forms of agency action, the Chapter makes recommendations for ways to de-bias individual perceptions of blame and coercion. Both of these steps should be undertaken by the Office of Information and Regulatory Affairs (OIRA), I argue. OIRA, which ensures the efficiency and transparency of regulations, is well positioned to explain to elites and the public which agencies are responsible for the development of policy and how coercive policies are likely to be.

Third, and finally, this Chapter applies the approach to legal change and stability developed in the previous Chapter. Using the primary forms of agency action, it shows how agencies might chart a middle path between upsetting reasonable expectations and ignoring democratic mandates to govern. The hope is that, by the end of the Chapter, we will have an integrated set of recommendations concerning the *who*, *what*, and *when* of regulatory accountability.

II. De-biasing derivative liability

Chapter Two argues that evaluations of policy are not independent of the structures that produce them. Rather, delegations of policy-making authority systematically induce voters to under-assign blame to legislatures. Legislatures face less blame even in cases where they delegate to non-autonomous, "instrument" agencies. In effect, legislatures can delegate the final, public-facing component of a policy-making process to an agency in order to soften the blame

they face. This finding coheres with a body of economic research that shows that attributions of blame diminish when a principal employs an agent.

Post-*Chevron* approaches to administrative accountability have sought, in varying degrees, to give "maximum authority to the most politically responsive decision maker," on grounds that doing so "maximizes the responsiveness of policy to majoritarian preferences."[1] This is the core insight of the "derivative liability" approach to accountability. Agency heads, of course, are unelected. If, however, some elected office that possesses leverage over agency behavior can be reliably held to account for agency performance, voters might be able to hold agencies derivatively accountable. Agency policy choices that draw the ire of the electorate, on this view, can be expressed through punishment of the president, who is directly accountable to voters.

Proponents of derivative accountability argue that the best way in which to legitimate the administrative state is to give expansive interpretive powers to the executive branch. *Chevron* and its progeny, the theory runs, accomplish this by maximizing transparency. A "fundamental precondition of accountability in administration" then-Professor Elena Kagan argued, is "the degree to which the public can understand the sources and levers of bureaucratic action."[2] Although agencies are, by their very nature, "the ultimate black box of government," if "clear lines of command" running from the executive to the agencies are established, then administrative power can be exercised "in ways the public can identify and evaluate."[3] In a similar vein, Stephen Holmes argues that separation of powers can "sort out unclear chains of command, and help overcome a paralyzing confusion of functions."[4] If courts endow the executive branch with responsibility over making policy choices and interpretive choices, and the president accepts responsibility for this role, then the electorate will link the two sets of actors. Most importantly, the electorate will evaluate the actions of the agency, then assign those actions to the executive, as a part of its evaluation of the executive. The electorate makes a judgment as to the credit or blame of the agency, then assigns that score to the executive. Derivative accountability, then, holds that transparency is improved, and information is increased, through the process of assignation: the electorate treats the actions of agencies as the actions of the president. In this way, *Chevron* is believed to maximize the responsiveness of policy to majoritarian preferences.

With the demise of other modes of administrative accountability—interest group pluralism, for instance—it is derivative accountability that animates approaches to administrative accountability. The findings reported in Chapter Two, however, suggest that derivative electoral accountability is subject to bias caused by the dynamics of blame associated with delegation. If derivative accountability is thought to operate via the executive branch—that is, if citizens are to blame the president for agency action—the public's evaluation of the policies of agencies may tend, in the case of less powerful agencies, to be inflated, and, in the case of more powerful agencies, to be deflated. In other words, delegation tends to clump attributions of agency blame,

---

[1] Matthew Stephenson, *Optimal Control of the Political Bureaucracy*, 107 MICH. L. REV. 53, 54 (2008).
[2] Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245, 2332 (2001).
[3] Id.
[4] Stephen Holmes, PASSION AND CONSTRAINT 165 (1995).

which will result in relatively blameless agencies facing undue heat from the public, and relatively blameworthy agencies getting a pass.

If derivative accountability is thought to run through the legislature, the problems are even more straightforward: delegation reduces the blame the principal faces. Delegation tends to artificially inflate the public's opinion of legislative action, relative to a baseline of direct legislative action. Such skewed valuations of agency action sit uneasily alongside the *Chevron* theory of electoral accountability.

II.A. Reasons for optimism

Individuals' ability to accurately blame officials for bad policies is at the heart of the Supreme Court's post-*Chevron* jurisprudence. In the most recent major post-*Chevron* case, *City of Arlington v. F.C.C.*, the Court took up the question of whether a reviewing tribunal should apply *Chevron* to an agency's determination of its own jurisdiction. Most agree that Chevron generally gives agencies, rather than the courts, the authority to interpret ambiguous provisions in statutes that they clearly possess the authority to administer, but what of agencies' authority to decide if they possess the authority to administer the statute? *City of Arlington*, like many of the foundational administrative law decisions that preceded it, answered this question by relying, in part, on *Chevron*'s claim that agencies are better positioned to resolve "competing policy interests" than are the courts. This idea—that the legislative and executive branches together should work out competing policy interests where possible—is rooted in *Chevron*'s logic of political accountability via the administrative state.

In order to realize this promise, we will need to de-bias individual perceptions of blame. If we identify the way in which attributions of blame are sensitive to the legal structure governing the relationship between legislature and agency and the type of policy-making authority exercised by the agency, we can begin to consider institutional designs that better accomplish the twin goals of robust agency policy making and democratic accountability. Most optimistically, we can assess the tendencies that push observers to blame more in one direction or another and institutionally harness these tendencies to enhance accountability.

The dynamics presented here could induce the executive branch to publicly clarify its role vis-à-vis agencies. Recall that attributions of blame deviated less from baseline cases as agencies acted with more policy-making authority and less legislative oversight. Thus, broad delegations of power to the executive may enhance the performance of agencies by inducing the executive branch to take extra care of policy outcomes in these situations. By contrast, attributions of blame deviated the most from the baseline scenarios in those cases where agencies exercised less policy-making authority and were more subject to legislative oversight. The error in these cases was to blame agencies more, not less. Thus, derivative accountability may result in undeserved over-blaming of the executive branch where Congress has delegated away only minimal authority. In these situations, the executive branch might emphasize the relative power of Congress over the agency.

II.B. Responsibility "worksheets"

A more interesting solution would be to visibly link political principals to their agents, in order to establish a more visceral connection in the public's mind between the two. Given what we know about information processing and information overload, it seems unlikely that merely providing *more* information about what institution control what aspect of the decisionmaking process would be helpful. One can imagine, for example, OIRA creating a "responsibility worksheet" for any significant policy action. This would be akin to the way that OMB provides accounting checks on Congressional proposals.

As a matter of institutional design, OIRA is well-positioned to take on the burden of establishing responsibility worksheets. Currently, OIRA operates as a sort of "meta" regulator. It seeks to make regulation more efficient and ensure that the process is as transparent as possible. OIRA works closely with the policy development offices of each major agency. Therefore, a structure is already in place where agencies work alongside OIRA in the development of policy. In the course of preparing its efficiency guidance, it is not unreasonable to imagine OIRA also developing a responsibility worksheet that shows which departments are responsible for which parts of the policy. In fact, OIRA is best positioned for this task, since ideally the worksheet would be created *alongside* the development of the policy, by an independent third party.

II.C. Visualizing responsibility

While the worksheets might assist elites in understanding which institution is responsible for what part of the policy, it is unlikely to make much of an impact with the public. Visual representations, on the other hand, of responsibility might offer a much more compelling way to debias perceptions of blame. We are increasingly a visual culture, with more and more visual media available. An institution like OIRA might take a look at finding ways to represent the true dispersal of responsibility through mobile applications and video. That is, after the responsibility worksheet is created, that information might be translated to an easily-digestible visual representation.

III. Debiasing soft paternalism

Chapter Three showed that individuals under-blame agencies for nudging as compared with enacting hard law. This section considers some recommendations for de-biasing individuals in their evaluations of soft versus hard paternalism. In the interests of providing concrete solutions, we will consider these recommendations in the context of the administrative policymaking process.

First let us consider the three prominent means, laid out in Section 553 of the Administrative Procedures Act, by which agencies might enact soft paternalistic policies. (These will be relevant in the next section, on stability, as well.)

> Notice and Comment Rulemaking: Provides the public an opportunity to be involved in significant agency actions. The regulator initiates an annual cycle of notice and comment rulemaking. Interested parties participate and the agency provides reasons for its actions.

Interpretive Rulemaking: Rather than initiate notice and comment rulemaking, the regulator decides to act by interpreting the meaning already inherent in recognized sources of authority, such as statutes. The agency announced its updated interpretations periodically on its website.

Policy Statement Rulemaking: Taking advantage of its ability to announce new nonbinding norms, the agency periodically announces its stance on matters of importance to its stakeholders.

To the extent that these tools are used to enact soft paternalism, steps might be taken to debias individual perceptions that soft paternalism is less coercive than law.

Recall that our discussion of soft paternalism is focused more on elites than on the broader voting public. This is due to the fact that most nudges are designed not to be perceived by the end-users. So the focus is shifted to elites who follow the policymaking process and might be able to expose the use of soft paternalism to the public. This focus on elites makes the debiasing process easier, perhaps, than in the previous case. With elites, we can assume that if the type of "responsibility worksheet" discussed earlier was distributed, there would be uptake of the information included therein. In the case of soft paternalism, however, we don't need a responsibility worksheet, but, rather, a "coercion worksheet."

The idea behind the coercion worksheet would be for OIRA to establish coercive equivalencies between different regulatory tactics. For example, if an agency sought to deploy a nudge, the OIRA coercion worksheet could estimate the comparable coercive effects of other regulatory tools. The coercion worksheet would serve, in effect, as a way to compare the coercive effects of a range of policy alternatives. As such, it would also indirectly improve the policy development process.

Each of the major three modes of agency action would need to package the worksheets differently, to maximize impact. For notice and comment rulemaking, the worksheets could be packaged alongside the other materials agencies circulate to stakeholders. Note the important function that such a document would play. On the one hand, it would alert interested parties to the actual degree of coercion involved in the soft paternalistic policy. That, in turn, would presumably help to equalize levels of blamed incurred by regulators using nudges and those using laws. In addition to that primary function, however, the worksheet would also help interested parties compare different types of regulations. Having an explicit measure of the coerciveness of a particular set of possible regulations might well be a valuable analytic tool in comparing different regulations. For the latter two types of rulemakings, the worksheets would need to be circulated alongside the agency's statement of its understanding of the new norm.

IV. Establishing an acceptable rate of legal change

Finally, let us consider how to apply our thinking about the proper rate of legal change. In Chapter Four, we established a basic framework for thinking about these issues. Our framework is a step toward establishing a means of assessing, controlling, and allocating the costs associated with each species of action. We build this framework by situating each of the three modes of

agency action within the legal environment that governs how agencies issue new norms or change existing norms. Once we have situated the modes of action within this legal environment, we can create scenarios describing how, all things equal, a changed norm will affect law's capacity to guide behavior. In creating these scenarios, we move past the conventional wisdom that draws a sharp dichotomy between stability, which is perceived as necessary for the rule law but bad for a citizenry in need of responsive regulation, and "change," which is deemed harmful for the rule of law but good for citizen interests.[5]

The policy questions surrounding stability are primarily hashed out in federal court. Theoretical concerns of this sort we considered in Chapter Four have recently become relevant to legal doctrine governing the ability of administrative agencies to rapidly respond to changing environments. As Kozel and Pojanowski argue, "stability is usually best promoted by limiting the frequency and degree of legal change."[6] In order to promote stability, they call for heightened judicial oversight "as a braking mechanism to prevent an endless string of" changes.[7] There are a number of components to the emerging body of law governing administrative change that led to this conclusion. The most important for our purposes is a provision in the Administrative Procedures Act instructing courts to reject any agency action it finds "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law."[8] It is this prong of the law of agency change that will most likely be the site of judicial policy change.

The conventional wisdom undergirding the Court's arbitrary and capricious doctrine is that, when an agency changes one of its rules, that new rule is "more likely to be arbitrary in several senses: ... [it] is more likely to unsettle reliance interests, is more likely to reflect a partisan political judgment not carefully moored to carrying out Congress's statutory purposes, and is more likely to raise new interpretive difficulties for regulated interests."[9] By creating a framework that reveals how different types of agency action have differential effects on regulated parties, our analysis speaks directly to the first and third of these concerns.

The leading case on the APA's arbitrary and capricious standard is *F.C.C. v. Fox Television Stations*, in which the Court held that agencies' changing a "prior policy [that] has engendered serious reliance interests" must set forth a "more detailed justification" than it otherwise would, or risk a finding that the agency's change was arbitrary and capricious.[10] When an agency changes a policy, the Court reasoned, it can disrupt expectations that it created. And failing to consider such disruptions are grounds for rejecting the change. In short, agencies must analyze the consequences of their proposed changes on those who have relied. Closely related to the question of reliance interests is the matter of interpretive difficulties. One advantage of longstanding policies is that they have been interpreted, by the agency that enacted them, courts,

---

[5] Building such a framework is thus useful for two reasons. First, agencies might select modes of action with an eye to lessening the costs to legality incurred by flexible regulation. Second, our analysis can inform courts' application of the Administrative Procedures Act arbitrary and capricious standard, a legal standard that might well define the degree of freedom agencies possess to regulate responsively.

[6] Kozel, *supra* note 2, at 157.

[7] Id.

[8] 5 U.S.C. section 706(2)(A) (2006).

[9] Eskridge and Baer, The Continuum of Deference: Supreme Court Treatment of Agency Statutory Interpretations from Chevron to Hamdan, 96 GEORGETOWN L. REV. 1083, 1151 (2008).

[10] 129 S. Ct. 1800, 1811

and by the parties. Changed policies have yet to be interpreted. Their meaning and effects are thus much less certain than longstanding policies.

While in some cases changing rules or interpretations undermines reliance interests and certitude of interpretation, the matter is more complicated. First, different modes of agency action – considered above – can upset reliance and certainty interests in different ways. Second, agencies have the tools at their disposal to make the disruptions that are inevitable less severe. Third, parties' reliance and certainty interests are contingent on their prior plans, and these prior plans are, in some instance, subject to reconsideration. If we can show that one or more of these claims is plausible, it suggests that courts should consider updating the arbitrary and capricious doctrine to account for this more nuanced take on reliance and certainty interests.

Here, we will see that the way in which different types of agency action – NC rules, interpretive rules, and policy statements – differentially affect the quality of signals, and thus epistemic guidance. If we can show some plausible hypotheses about the way in which different types of regulatory action differentially affect epistemic guidance, we will have grounds to proceed with empirical studies attempting to bear out these hypotheses. More importantly, if we are able to identify predictable dynamics, we can focus on how to minimize the costs to epistemic guidance that might otherwise attend to the type of volatile regulatory orders necessary to address rapid societal, economic, and political changes.

We will pursue these dynamics by teasing out scenarios that assume some set of baseline conditions. Each scenario is structured around two assumptions, and the outcomes of the scenarios are dependent upon these. Each scenario is described in terms of an action type (NC rule change, e.g.) and a legal environment into which that action type is launched. The structure of the legal environment into which action types are launched captures the background web of rules created by the agency prior to it taking some action. These two assumptions factor together to affect the two components of epistemic guidance: erroneousness and incompleteness. We can ask of each set of assumptions whether they create competitor norms, thus increasing the chances that an official mark is erroneous, and whether they increase or decrease the completeness of the information referred to by the official mark. We characterize the outcomes as "Good," "Neutral," or "Bad." These measures are far from precise, but they do establish a basic framework for thinking about epistemic guidance.

This approach allows us to create hypotheses about how the different types of agency action work in concert with each other. Our framework is a step toward establishing a means of assessing the costs associated with each species of action, such that modes of action might be selected in order to lessen the costs to legality incurred by flexible regulators.

IV.A. Scenario 1: Stable Notice and Comment Rules

Consider an ideal case of epistemic guidance. An agency makes a rule, using NC rulemaking procedures, which are the most authoritative procedures used in everyday practice by agencies, and that rule remains in place, unchanged, for some extended period of time. Further, the agency does not dilute the informational value of the official mark by employing interpretive rules or policy statements, which have the effect reducing the percentage of the total information

an individual needs to know in order to comply with the rule. The rule produced using NC procedures disposes of the legal question.

Such a pattern of action best facilitates epistemic guidance because (1) it does not create any competitor norms by proliferating erroneous signals and (2) it consolidates the information that an individual needs to know in order to comply under a single official mark. Accordingly, it is the gold standard for modes of agency action designed to create epistemic guidance.

IV.B. Scenario 2: Changed Notice and Comment Rules

NC rulemaking can affect the quality of legal signals in either of two ways. In Scenario 2-A, we consider how changing NC rules increases the number of competitor signals, thus possibly reducing guidance. In Scenario 2-B, by contrast, we see that NC procedures are likely to enhance the quantity of information contained in signals in cases where NC rules supersede other types of less authoritative or murky information put out by agencies with authoritative rules.

When an agency must create new norms to deal with rapidly changing circumstances, the ideal pattern of epistemic guidance we witnessed in Scenario 1 is upset. In Scenario 2-A, imagine that an agency has created a NC rule in the past to regulate a situation, and now seeks to create a new NC rule to supersede the old rule. Each change to a NC rule creates a superseded NC rule. These superseded rules continue to bear the official mark of law, creating the potential for ``confusion,'' as the Seventh Circuit cogently observed. Each change, then, increases the number of signals that compete with standing law. In other words, in addition to the social, religious, and moral norms that purport to guide individual conduct, so too do superseded legal norms claim the power to guide individual conduct. All things equal, then, each NC change means the agency's norms must guide conduct amid a larger din of competing norms.

If we change our assumption about the nature of the legal environment in which the agency acts, the consequences for epistemic guidance of altering a NC rule change. Assume not that the agency has regulated a topic with NC rules, but, rather, with a disjointed collection of nonlegislative rules. In this situation, NC change can facilitate epistemic guidance. Contrary to conventional wisdom, instability can enhance law's ability to guide conduct.

Here's how. If, in the past, agencies have tended to guide through nonlegislative rules, then NC rulemaking can supersede those nonlegislative rules. As a result, a higher percentage of the information required to know how to comply with the law is referenced by an official mark. Likewise, the portion of the norm contained outside that text bearing with the official mark is reduced. Because the legal status of interpretive rules and policy statements are always subject to dispute, superseding them through NC rulemaking eliminates uncertainty as to just what the law is.

Consider, for instance, the S.E.C. investigation of Netflix C.E.O. Reed Hastings for publishing material information on his Facebook account. The underlying regulation, Regulation FD, was enacted in 2000; the S.E.C. issued a guidance document in 2008, which is styled as an interpretation of Regulation FD. Opinions differ on the extent to which the 2008 document covered a situation like Hastings' own, but it clearly did not address the situation directly. As a

result, onlookers called for the S.E.C. to engage in NC rulemaking in response to the Hastings matter. A rulemaking would allow the agency to take a unified approach to a fast evolving question. Rather than govern via piecemeal nonlegislative rules and enforcement actions, the S.E.C. might unify the rules governing Reg FD under one, major pronouncement. A NC rulemaking would unify the laws governing selective disclosure under one mark, and thus clarify the meaning of the law to affected parties.

IV.C. Scenario 3: Changed Interpretive Rules

Shifting our focus from NC rules to nonlegislative agency activity, we turn to the second major mode of agency action, interpretive rules. Interpretive rules, recall, allow an agency to provide its gloss on extant laws. They are especially useful in cases where a law is vague. Interpretive rules allow an agency to inform the public how it understands the extant law. When changed, interpretive rules have almost the opposite effect on epistemic guidance as NC rules. Changing interpretive rules does not harm epistemic guidance by increasing the number of competitor marks, since legislative rules do not bear official marks. However, changing legislative rules does reduce the informational quality of the official marks that do exist.

Assume an agency has acted to regulate a topic in the past with through NC rulemaking. Where changes to NC rules bring a larger portion of information required to comply with the law under the umbrella of the official mark, changes in interpretive rules do just the opposite. Say an agency makes a rule via NC rulemaking governing some new area of concern; say, the permissible uses of 3-D printing. The agency then makes an interpretive rule spelling out its interpretation of the rule it previously made.

IV.D. Scenario 4: Changed Policy Statements

Unlike NC rules, which are law, and interpretive statements, which are interpretations of law, policy statements are, at their core, unconnected to standing law. To the extent that these types of norms guide individuals, then, it is not *as law*. Policy statements, then, must act by setting expectations about what the agency will do in the future.

From the perspective of epistemic guidance, policy statements, precisely because they lack any connection to standing law, are neutral. Policy statements do not bear the official mark of law, so they their proliferation does not proliferate official marks. Nor are they interpretations of law, so they do not disperse the information needed to comply with law. Policy statements operate via a mode of engagement with regulated parties that is outside of law: expectations. Epistemic guidance, recall, is a necessary function of law. This mode of agency activity, which guides via expectations, does not impact law's ability to guide, so it does not impact epistemic guidance. Its effects are neutral.

IV.E. Packaging Modes of Agency Action

Together, our scenarios create a framework for agencies to mitigate the costs to epistemic guidance that result from changing regulations. Agencies can mitigate the risks that instability poses to epistemic guidance by taking advantage of the fact that they can act in these three ways,

each of which has a different effect on epistemic guidance. The fact that these different modes of agency action affect epistemic guidance differentially implies that agencies possess some leeway in assessing how to impose the costs that result from changing rules.

The conventional wisdom, which we encountered at the start of the Chapter, holds that instability impairs epistemic guidance, and, thus, that agencies should hold their hands when they might otherwise regulate in order to maintain the quality of their epistemic guidance. We have shown that agency action is diverse; and not only is agency diverse, it affects epistemic guidance in diverse ways. Here, we will suggest some ways that agencies can use different types of action together in order to improve the quality of both of their epistemic guidance. Doing so, it seems, will give us a basis for questioning, if not outright rejecting, the conventional wisdom that agencies should hold their hands in order to enhance stability.

Recall that, in Chapter Four's discussion of the costs of instability, we observed that instability might bring about either of two situations, neither of which bodes well for epistemic guidance. Instability, we saw, might individuals might stop the search for the content of law. If, at a certain point, the costs become high enough and the content of the legal norm is non-intuitive enough that individuals cannot figure out the compliant behavior absent the norm, individuals may simply stop searching for law. Further, instability might induce mistaken non-compliance. Here, individuals erroneously rely on a superseded official mark. Despite a desire to comply with the law, and a belief that, in fact, they are complying, we witness a failure to comply. This failure is due to the fact that instability decreases the strength of the signal sent by officials to subjects.

If agencies see evidence of either of these difficulties, or anticipate them, then they might deploy their three modes of action in a concerted effort to improve the quality of their guidance. Doing so might well mitigate the epistemic costs of instability. The idea is straightforward: Agencies can take advantage of the fact that they possess the power, due to the APA, to issue norms that are not laws. These nonlegislative rules can reduce the costs typically associated with instability while providing the guidance individuals need to keep up with the changes.

At the most general level, agencies might take note of the ways in which different sorts of action impact the two major sources of epistemic guidance. Because, as noted above, individuals will respond to signals in different ways, and agencies are apt to know those that they regulate, it falls to the regulators to be sensitive the needs of the regulated as they concern epistemic guidance. We may be able to say something more specific, however. Agencies can issue interpretive documents, which interpret, explain, and give the rationale behind agencies' legal norms. They can also issue policy statements, which set expectations about the direction of future agency action. These nonlegislative rules can play three important roles with regard to stability.

IV.F. Substituting Legislative for Nonlegislative Rules

There are some situations in which changes to NC rules can straightforwardly improve epistemic guidance. Where nonlegislative rules have proliferated, and, in effect, govern some area, changes to NC rules enhance epistemic guidance. Substituting NC rules for either

interpretive rules or policy statements enhances epistemic guidance, because NC are law, not merely interpretations of law or policy statements about law. The regulated thus know that they are guided by the same norms that will be applied to them by a court or administrative tribunal. This helps law epistemically guide in a number of ways. First, there are simply fewer documents to locate in order to be able to find out what behavior constitutes compliance. This feature is of more importance for less sophisticated parties. Second, multiple interpretations need not be reconciled, as is the case when multiple interpretive rules effectively govern an area. This enhances the certainty of the guidance. This is beneficial for sophisticated and unsophisticated parties alike, as well as for the courts that must interpret the documents.

IV.G. Substituting Nonlegislative for Legislative Rules: Interpretive Rules

If, on the other hand, an agency is disinclined to use NC rulemaking, because it is wary of proliferating superseded NC rules, or for some other reason, interpretive rules can be employed. The costs to this approach to epistemic guidance are noted above. There are, however, some benefits to epistemic guidance that might flow from using interpretive rules. In some situations, interpretive rules can enhance epistemic guidance by demonstrating to the deep legal reasoning behind agency decisions. As their name suggests, interpretive rules are statements about the agency's interpretation of an extant legislative rule. These documents, however, do not summarily state the agency's position; they provide the legal reasoning behind the agency's position. This reasoning is a dialogue between the changed circumstances in the world that necessitate the interpretive rule and the text of the extant legislative rule. The marriage of fact and law can update regulated parties' views of what is required of them to comply. For agencies disinclined to issue fresh NC rules, this updated guidance is preferable to guessing at how the agency perceives the old NC rule, at least from an epistemic guidance perspective. Accordingly, interpretive rules can allow agencies to fill regulatory gaps in rapidly changing regulatory environments while also providing epistemic guidance.

IV.H. Substituting Nonlegislative for Legislative Rules: Policy Statements

Policy statements, as noted above, advise the public about how the agency intends to exercise the powers it possesses. Because policy statements are not law, nor are they interpretations of law, in one sense policy statements have no bearing whatsoever on epistemic guidance, as we have defined it. Epistemic guidance, in its purest form, and in the form we've addressed it here, is about knowing how to comply with the law. Because policy statements are not law, nor interpretations of law, they bear on indirectly on the matter of epistemic guidance. For instance, policy statements might do the work of reconciling the reasoning contained in multiple interpretive statements, thus providing some measure of clarity as to the agency's perspective on a body of law.

From another perspective, however, policy statements are very important components of epistemic guidance. For although they do not tell individuals how to comply with the law, they do tell individuals how those officials holding public power plan to exercise law. In practice, if not in theory, policy statements might be the most useful documents issued for some matters.

To a first approximation, we have located our stability-based constraint on legal norms that seek to provide motivational guidance: Legal norms cannot tend to take content that breaches the volatility threshold for any given individual. Individuals use legally-sanctioned means as the basis of plans, which, in turn, requires legal norms that are stable enough not to breach any given individual's threshold of volatility.

Law's capacity to guide is typically thought of as a capacity that extends uniformly over all subjects possessing the capacity to reason. The conventional wisdom among those that argue that agencies should seek to motivate individuals by keeping law stable assume a homogenous individual. Here, by contrast, we offer a differentiated approach to motivational guidance. Ours is one premised on the idea that individuals possess different tolerances for the risks associated with volatility. By introducing deliberative types, we have introduced differences in the population.

The primary worry of advocates of stability concerning motivation is that instability disrupts expectations, which then disrupts law's capacity to motivationally guide. We should clarify exactly what the claim is. Advocates of stability sometimes fail to address the fact that it is not law, in the first instance, that is disrupting expectations; the unsettling of expectations results from change in some non-legal aspect of the world. Imagine, for instance, that there is some business whose model it is to scour corporate press releases and disseminate them to the larger public, and this business opposes the S.E.C. updating Regulation FD to allow for dissemination over social networks, thus largely obviating the need to scour press releases. Everyone can agree that this company's expectations have been unsettled. It made investments assuming one set of facts, and then sought to deploy those investments in a world governed by a different set of facts. The mere fact that this company's expectations have been unsettled is not problematic from a rule of law perspective. What proponents of stability object to is a situation where the unsettling of expectations results directly from an unjustified agency change.

Now, all the weight of the claim falls on what counts as a justified legal change. As the Supreme Court held in *Fox*, an agency reversing course must "show that there are good reasons for the new policy."[11] The Court held that, where an agency's own position induced serious reliance interests, the agency would need to adduce better reasons than it ordinarily might. The Court implied, however, that changed facts would suffice to explain a change of course. On this matter, we don't seek to provide an abstract answer. We can simply observe that advocates of stability seek to draw a more restrictive line when it comes to defining what reasons are good enough.

Our analysis shows that where one draws this line is --- or ought to be --- a function of the volatility thresholds of the regulated parties. Agencies regulating industries where parties are quite adept at reconsidering their plans might have more leeway to change regulations. Agencies regulating more staid industries might have comparatively less room. The responsibility falls to the agency, in the first instance, to know the interests that it regulates. Likewise, when courts evaluate the decisions of agencies, they might show sensitivity to the differing ways that regulated parties react to the different modes of action. In the following section, we offer a

---

[11] F.C.C. v. Fox Television Stations, Inc. 556 U.S. 502, 515 (2009)

framework explaining how the different modes of agency action affect parties with varying volatility thresholds.

Framework for Analyzing Effects of Agency Action on Motivational Guidance

	Action type	Structure of the Legal Environment *prior* to Agency Action	Creates (New) Binding Law?	Prefer Certainty?		Enhanced Motivational Guidance?	
				Non-Robust Types	Robust Types	Non-Robust Types	Robust Types
Scenario 1	Stable N&C	No pre-existing agency rule	Yes	No	Yes	No	Yes
Scenario 2-A	Changed N&C	One or more N&C rule	Yes	No	Yes	No	Yes
Scenario 2-B	Changed N&C	One or more non-legislative rule	Yes	No	Yes	No	Yes
Scenario 3	Changed Interpretive Rules	One or more N&C rule	No	No	Yes	Maybe	Maybe
Scenario 4	Changed Policy Statements	One or more N&C rule	No	No	Yes	Maybe	Maybe

IV.I Scenario 1 Revisited: Stable Notice and Comment Rules

In Scenario 1, an agency makes a rule, using NC rulemaking procedures, which are the most authoritative procedures used in everyday practice by agencies, and that rule remains in place, unchanged, for some extended period of time. The rule produced using NC procedures disposes of the legal question. Scenario 1 embodies the ideal of stability. Interestingly, however, this

stability might not always enhance motivational guidance for all parties. Non-Robust types, as addressed above, *do not* prefer legal certainty. Thus, the stability evinced in Scenario 1 may not enhance motivational guidance for such parties.

IV.J. Scenario 2 Revisited: Changed Notice and Comment Rules

All things equal, changes to NC rules create certainty for regulated parties in cases where NC rules supersede nonlegislative rules and thus unify an agency's regulations. The certainty results from the fact that legislative rules are legally binding, and nonlegislative rules are not. Thus, to the extent that a party is motivated by a desire for certainty, volatility will, in some situations, bolster motivational guidance, not impair it, as is commonly imagined.

This is especially true of what we have termed Intrinsically Stable and Strongly Robust types. These types most prefer certainty, because they hold fast to their plans in the face of change. So, for instances, companies with business models that have proven successful in the past and thus are being counted on to produce into the future might well prefer the certainty that accompanies NC rules. These are often incumbents in an industry, and they prefer to proceed according to carefully constructed long term plans, which depend on regulatory stability. Such parties have comparatively lower volatility thresholds; that is, they will tolerate less uncertainty in the regulatory environment. New NC rulemaking, then, even though it introduces a change in the set of governing norms, is more likely to induce motivational guidance for these types of parties.

On the other hand, Non-Robust and Weakly Robust types might be hurt by the change that results from added certainty. This claim might seem nonsensical; after all, don't we all prefer certainty? In short, no. In contrast to the incumbents who prefer stability, new entrants to an industry may thrive in an unstable environment. The changing regulatory environment may not impact the ability of law to guide motivationally, since these parties are willing to tolerate highly volatile environments. Non-Robust and Weakly Robust types, however, may well benefit from the uncertainty.

We can quickly draw two conclusions that defy conventional wisdom. First, volatility can generate certainty. This is due entirely to the fact that agencies have three modes of action available to them. The conclusion generalizes only insofar as other sorts of institutions have analogous modes of action available to them. Second, and also contrary to the conventional wisdom, legal certainty does not impact all parties equally.

Most interestingly, certainty does not ensure increased motivational guidance. Some types of parties may well believe they can flourish in uncertain environments, and, if we see law as motivationally guiding by providing a platform of means, which parties can use to accomplish their ends, those types that flourish in uncertain environments may flock to the legal platform. Some regulated parties may prefer that their own, and others', reliance interests are disrupted. This is of course not reason for agencies or courts to, in fact, disrupt those interests. But it does suggest, at minimum, legal doctrine relying on reliance interests needs to be grounded on something other than the preferences of the parties.

IV.K. Scenarios 3 and 4 Revisited

Scenarios 3 and 4 both involve the creation of new, nonbinding rules, layered on top of extant NC rules. It is thus difficult to describe in the abstract whether or not they enhance regulatory certainty. In theory, interpretive rules should always be certainty-enhancing. Why? Because interpretive rules take an extant law and provide an explanation of how the agency understands that law. Interpretive rules, in theory, single out the one meaning that the agency intends to give to a law that could support a variety of meanings. When the agency picks out its preferred meaning, other possible interpretations of the law fall to the side, thus creating certainty for regulated parties. In theory, then, we should be able to classify interpretive rules alongside NC rules as modes of action that enhance certainty. However, even if interpretive rules are actually, as in theory, certainty-enhancing with regard to the content of law, there is the question of whether that rule will withstanding judicial scrutiny. In sum, then, the effects of interpretive rules and policy statements on motivational guidance are difficult to state in the abstract.

We have established a basic framework for thinking about the role of agency action with regard to certainty and motivational guidance. But this framework assumes that individuals' preferences over reconsideration are fixed. In this final section, we consider how policy statements can induce parties to prepare for the future by shifting their preferences over reconsideration. In short, individuals can structure their affairs such that the possibility of legal change is less likely to reach the threshold beyond which the legal norm is perceived to be too volatile to form the basis of a plan. Individuals can plan for legal change itself.

V. Electoral accountability and the power gap at the center of the state

Will the measures recommended here solve the problems raised in Chapter One? It is doubtful; in fact, the answer is almost certainly no. The problems raised in Chapter One were strongly influenced by the failure of the traditional model of electoral accountability, but are also strongly influenced by changes in political norms, our deeply connected social world, and the decreasing power of elites. None of those issues will be resolved by "responsibility worksheets."

At a more abstract level, however, the dissertation recommends retrofitting our existing institutions to reflect what we continue to learn about political behavior. As mentioned, there are profound social and economic dynamics that are helping to reduce the efficacy of our political system. So it as not as though studies on political behavior are somehow silver bullets. However, as evidence accumulates, we might do well to experiment with using these studies to influence institutions. For example, it would be infeasible to immediately initiate OIRA review of soft paternalism, as recommended above. However, convening a working group to consider that solution at the local or state level might be feasible.