



Over-Accountability

Citation

Jacob E. Gersen & Matthew C. Stephenson, Over-Accountability, 6 J. Legal Analysis 185 (2014).

Published Version

<http://jla.oxfordjournals.org/content/6/2/185.full.pdf+html;10.1093/jla/lau008>

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:16073949>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

OVER-ACCOUNTABILITY

Jacob E. Gersen* and Matthew C. Stephenson†

ABSTRACT

Although ensuring the “accountability” of agents to their principals is widely considered a core objective of institutional design, recent work in political economy has identified and elucidated an important class of situations in which effective accountability mechanisms can decrease, rather than increase, an agent’s likelihood of acting in her principal’s interests. The problem, which we call “over-accountability,” is essentially an information problem: sometimes even a fully rational but imperfectly informed principal (e.g., the citizens) will reward “bad” actions rather than “good” actions by an agent (e.g. the President). In these cases, not only do accountability mechanisms fail to remedy the agency problem inherent in representative government, they actually make the problem worse. This Article offers a conceptual and empirical overview of over-accountability problems, and also considers a range of potential solutions. By surveying both the distortions themselves and a range of possible responses, this article aspires to assist both public law scholars and institutional reformers in producing more effective solutions.

Politics and government are rife with principal–agent problems.¹ Government officials (the “agents”) are supposed to make decisions on behalf of, and for the benefit of, the citizenry (the “principals”), yet there is no guarantee that government officials will actually possess the superior wisdom that supposedly

* Professor of Law, Harvard Law School
Corresponding author: Jacob E. Gersen, Harvard Law School, 1563 Massachusetts Ave., Cambridge, MA 02138, Email: jgersen@law.harvard.edu. We are grateful for extremely useful comments from Rosalind Dixon, Richard Fallon, Josh Fischman, Justin Fox, Louis Kaplow, Randy Kozel, Ming Li, Martha Minow, Jennifer Nou, Jide Nzelbe, Dave Owen, Fu Qiang, Mark Ramseyer, Steve Shavell, Ken Shotts, Adrian Vermeule, and Jennifer Wriggins.

† Professor of Law, Harvard Law School

1 See, e.g., Besley (2007); Persson & Tabellini (2000); Bueno de Mesquita et al. (2004); Miller (2005); Moe (1984); Ferejohn (1986); Barro (1973). For examples of legal scholarship on public law that explicitly uses this framework, see, e.g., Nzelibe & Stephenson (2010a); Berry & Gersen (2008); Gersen (2007); Stephenson (2011); Samaha (2006); Levinson (2012); Levinson (2011); Posner & Vermeule (2007); Bibas (2009); Pierce, Jr. (1989); Law (2009); Amar (1991).

justifies their powers, and government agents may have their own interests that diverge from those of the people they represent. This agency problem afflicts not only the relationship between citizens and their elected officials, but virtually every relationship within government as well. Agency problems exist between legislators and staff, between the Supreme Court and lower courts, between the President and cabinet secretaries, between Congress and the bureaucracy, and between virtually any supervisor and subordinate.

The most familiar solution to principal-agent problems in government is to strengthen accountability mechanisms, such as elections, oversight, performance reviews, and the like. Indeed, “accountability” is one of the workhorse concepts of public law, lauded by courts and scholars as a core value of our constitutional democracy. Courts, scholars, and reformers ask questions like: How can constitutional law ensure that lawmakers will be accountable to the citizens they serve?² How can administrative law ensure that agencies enact policy consistent with legislative and presidential views?³ How can campaign finance and election law be designed so as to promote the right sorts of political accountability? More generally, because all accountability mechanisms are imperfect, a great deal of intellectual energy in the legal academy, in policy reform movements, and in the courts, is spent on finding ways to redress problems of under-accountability, either by strengthening existing accountability mechanisms or suggesting new ones. And those features of our constitutional system that lodge power in less-accountable actors, like courts or independent agencies, attract considerable criticism, precisely because they raise concerns about under-accountability.⁴

Yet accountability also has a dark side. In particular, recent work in political economy has identified and elucidated an important class of situations in which effective accountability mechanisms can *decrease*, rather than increase, an

2 For example, debates about federalism and localism are shot through with arguments as to which level of government is more accountable to the people. *See, e.g.*, Feeley & Rubin (1994); La Pierre (1985); La Pierre (1982); Merritt (1988); Hills (1998); Chemerinsky (1995); Friedman (1997); Stewart (1977).

3 Advocates of demanding administrative procedures argue that such procedures promote accountability, with one court going so far as to declare that “political accountability . . . is the very premise of administrative discretion in all its forms.” *Newman v. Apfel*, 223 F.3d 937, 943 (9th Cir. 2000).

4 For example, the expansion of the administrative state has been routinely criticized for allowing consequential policy decisions to be made by unelected bureaucrats. *See, e.g.*, Lowi (1969); Schoenbrod (1993); Epstein (2011); Aranson (1982). Relatedly, proponents of a “unitary executive” assert that a strong and accountable president is the solution to the concern about excessive delegation to the executive branch. *See, e.g.*, Calabresi (1995); Kagan (2001); Lessig & Sunstein (1994). As for courts, a stock criticism of aggressive judicial scrutiny of decisions by the legislature or executive is that such review is undemocratic. *See, e.g.*, Waldron (1999); Bickel (1962); Bork (1990); Tushnet (1999); Berger (1977); Kurland (1989) at 466; Farina (1989); Mashaw (1985); Horwitz (2008); Caminker (2003).

agent's likelihood of acting in her principal's interests. The problem, which we call "over-accountability," can arise even if—in fact, precisely because—both the principals and the agents act rationally. The over-accountability problem is essentially an information problem: sometimes even a fully rational but imperfectly informed principal (e.g., the citizens) will reward "bad" actions rather than "good" actions by an agent (e.g., the President). In these cases, not only do accountability mechanisms fail to remedy the agency problem inherent in representative government, they actually make the problem worse.⁵ Far from being anomalies, over-accountability problems may well be quite common. At a minimum, current debates should attend to the risks created by the mechanisms chosen to enhance accountability. Too many institutions are designed only with an eye towards the under-accountability problem when the over-accountability problem may be as serious or worse.⁶

5 In that sense, our analysis is closely related to other critical perspectives on political accountability that have emphasized that, even if one accepts a principal-agent framework and seeks to maximize faithful agency, this does not necessarily mean that any individual reform that enhances political accountability is a good idea. As several scholars have emphasized, the economists' "general theory of the second best," which establishes that in the presence of multiple market failures the correction of a subset of these failures may worsen rather than improve efficiency, see Lancaster & Lipsey (1956), has an institutional or political corollary: in the presence of multiple "democratic failures" (deviations from perfect political responsiveness), the amelioration of one such failure (say, by enhancing political accountability) can make the underlying principal-agent problem worse, not better. See Vermeule (2009); Vermeule (2003); Coram (1996) at 91; Goodin (1995) at 52–56; Solum (2008) at 311–112, 327–328; Ulen (1998) at 208–219; Stephenson (2007b).

6 We are hardly the first legal scholars to identify disadvantages of accountability. Indeed, the literature has identified a number of problems with accountability. Broadly speaking, existing critiques of accountability fall into five categories:

First, maximizing the accountability of government agents may not maximize overall welfare. For instance, democratic accountability creates risks of *majoritarian tyranny*. See, e.g., Komesar (1994) at 65–82; Guinier (1994). More generally, the outcome that prevails in a majority-rule system does not necessarily maximize the aggregate welfare of the citizens. See Stiglitz (2000) at 137–141. Alternatively, making agents more accountable may increase their responsiveness to parochial interest groups—a kind of "bad accountability" that reduces overall welfare. See, e.g., Vermeule (2007) at 11, 84; Lohmann (1998).

Second, some critics hold that political outcomes should be judged not (only) according to how well they promote the subjectively-perceived citizen welfare, but according to *extrinsic criteria* of justice and political morality. See, e.g., Dworkin (1977); Brown (1998).

Third, because the "principal" (the citizenry) in fact consists of a large number of individual principals (the citizens), notions of political accountability raise difficult *aggregation problems*, including concerns about the incoherence of collective preferences. See Arrow (1963); Riker (1982). *But see* Mackie (2003); Anderson & Pildes (1990). Yet another concern is whether a concept like "aggregate welfare" is meaningful in light of the difficulty of interpersonal utility comparisons. See Arrow, *supra*; Robbins (1932) at 125–126. *But see* Adler & Posner (1999) at 205–206; Markovits (1975) at 988.

Fourth, maximizing government responsiveness to the citizenry may be unwise because citizens exhibit forms of *irrationality* (temporary passions, persistently irrational beliefs, myopia, etc.). See, e.g., Caplan (2007); Kuran & Sunstein (1999); Somin (2004). *But see* Lupia & McCubbins

Section I of this article offers a conceptual overview of over-accountability problems. Some of the behaviors we describe will seem familiar from casual discussions of (and complaints about) American democracy; our goal is to describe with greater clarity and precision the underlying mechanisms that produce the perverse behavior, and how these mechanisms differ from other problems that might produce observationally similar phenomena. To this end, we draw on a set of models from political economy, previously ignored in most legal commentary.⁷ We identify five broad classes of over-accountability problems, which are rooted in the same underlying logic but have somewhat different characteristics: (i) *pandering* (the selection of a popular policy even when the agent believes the unpopular policy is in the principal's best interest), (ii) *posturing* (taking bold, risky action even when the agent believes the safe, conventional action is likely better), (iii) *persistence* (adhering to the policy the agent had originally adopted, even when subsequent evidence indicates the desirability of change), (iv) *populism* (adopting policies that injure some unpopular group or cause even more than the principal believes is appropriate), and (v) *political correctness* (adopting policies that benefit some sympathetic group or cause even more than the principal would view as appropriate).

Section II turns from problem to solution. Broadly speaking, there are two potential solutions to the over-accountability problem. First, and more obviously, one might simply limit accountability for certain kinds of decisions. This might be done by delegating to a less accountable decision maker (say, an independent agency or court). It might be done by extending or limiting the temporal scope of an agent's responsibility with term limits or a longer term of authority. Or, one might reduce the transparency of agents' decisions by limiting access to information or utilizing more complex decision-making processes. Second, using an additional agent or agents for oversight purposes can compensate for over-accountability. By employing a more complex

(1998); Schleicher (2008). On this view, an insulated agent may do a better job of advancing the citizens' true interests. See Miller (2005) at 219–222.

Fifth, the principals may suffer from a *time-consistency problem* (or a “credible commitment problem”), in which they would prefer to bind themselves to decisions in advance. See Holmes (1995); Elster (1979); Kydland & Prescott (1977); Rogoff (1985); Levinson (2011).

While these objections are all worth taking seriously, we (provisionally) accept the principal-agent framework as an appropriate lens through which to evaluate legal and political institutions, and we take a majoritarian conception of citizen welfare as our main normative criterion. We also put to one side concerns about aggregation, citizen irrationality, and time-consistency. We do this because we want to focus on a different sort of problem.

7 Although the literature from which we draw is technical, we strive to present the main ideas in non-technical language. As a result, some of the nuances of the technical literature may be obscured. Interested readers should consult the original sources, cited throughout.

decision-making structure with both a politically-accountable decision-making agent and one or more overseers, one can preserve the benefits of accountability while relying on the overseer to mitigate concerns about over-accountability. Yet oversight is no panacea, and indeed under some circumstances it can exacerbate rather than ameliorate over-accountability problems. We explore this issue, and we also explore oft-overlooked questions regarding the design of oversight institutions.

Our objectives are part descriptive and part constructive. Descriptively, we hope to clarify the nature of these potentially serious accountability-induced distortions, and to elucidate the mechanisms that may generate them. Constructively, we hope that by surveying both the distortions themselves and a range of possible responses, the analysis will assist both public law scholars and institutional reformers in producing more effective solutions.

1. OVER-ACCOUNTABILITY

Accountability mechanisms may ameliorate principal-agent problems in two related but distinct ways. First, principals can try to select good agents, where “good” is shorthand for an array of desirable characteristics, including competence, public-spiritedness, and motivation.⁸ In the electoral context, for example, voters use their limited information about the incumbent officeholder, if there is one, and about the other candidates seeking the office, to make judgments about which one would do the best job (or would be “best” for the voter in some other relevant respect). This *selection effect* means that on average, or over time, agents may be of higher quality than would be the case if they were chosen some other way. Second, accountability mechanisms can induce desirable behavior on the part of current agents.⁹ The idea behind this *incentive effect* is straightforward: because an incumbent agent knows that her principals will assess her fitness for continuation in office (or other rewards) based on her performance, she will make better choices than she would otherwise. Thus, not only can accountability mechanisms help principals select good agents, they can also induce bad agents to behave more like good agents. In the best case scenario, the selection and incentive effects both work to the principals’ advantage: the incentive effect induces *all* agents to try to act in ways that serve the principals’ interests, and when this sort of

8 See Fearon (1999) at 55. See also Ferejohn (1986); Maskin & Tirole (2004).

9 See Ashworth & Bueno de Mesquita (2008); Gordon et al. (2007); Ashworth & Bueno de Mesquita (2006) at 169; Gowrisankaran et al. (2008), Fearon (1999); Fiorina (1996).

discipline does not produce perfect convergence, the principals can use observed performance to replace less fit agents with more suitable ones.¹⁰

Sometimes, however, these two effects work at cross-purposes. In particular, sometimes an incumbent agent's concerns about the selection effect can induce her to make *worse* decisions than she would have made otherwise—a perverse incentive effect. The most general version of that basic point is quite familiar, not only in the scholarly literature but in everyday discussions of politics: elections, despite their many virtues, sometimes encourage politicians to make decisions contrary to voters' interests.¹¹ The reaction to these failures is often either a Churchillian shrug ("Democracy is the worst form of government, except for all those other forms that have been tried from time to time"¹²) or some form of Jane Addams' assertion that "the cure for the ills of Democracy is more Democracy"¹³—that is, the claim that democratic defects are a sign that we need accountability mechanisms that are stronger, deeper, and more robust. Contra these views, this Article shows why more accountability may sometimes make matters worse rather than better.

The basic problem is that principals are at an informational disadvantage relative to their agents in two respects. First, for many decisions, the agents have better information about the likely consequences of different courses of action.¹⁴ Indeed, this is one of the major reasons for a principal to delegate to an agent in the first place. Even after a decision is made, the principal may be unsure whether the agent did the right thing. This is particularly true when the "right thing" depends on probabilistic judgments about consequences, and when policy consequences are hard to discern.¹⁵ Second, the principals are often uncertain about their agents' preferences and abilities. In the conventional political economy jargon, they are uncertain as to each agent's "type," where "type" is a catchall term that might include a variety of relevant characteristics,

10 See Gordon et al. (2007); Ashworth (2005); Ashworth & Bueno de Mesquita (2008); Canes-Wrone et al. (2001); Banks & Sundaram (1998) at 294.

11 See generally Campbell et al. (1960). See also Zaller (1992) at 313; Riker (1982) at 136. See also Arnold (1993) at 401, 404–406; Downs (1957) at 219; Lupia & McCubbins (1998) at 79.

12 Winston Churchill, House of Commons speech on November 11, 1947.

13 Addams (1902) at 12. See also Dahl (1989) at 183–192.

14 This is not to say that government officials always (or even usually) know what they are doing, nor does it imply that all voters are completely ignorant. Rather, in most cases the average voter finds herself at a *relative* informational disadvantage vis-à-vis government decision makers with respect to the consequences of different possible policy choices.

15 See, e.g., Bueno de Mesquita & Landa (2014); Lohmann (1998); Nzalibe & Stephenson (2010b); Wolfers (2007).

including values, motivations, and competence at predicting consequences and implementing policies effectively. All else equal, a rational principal will prefer an agent who is a better type.¹⁶

The interaction between these two forms of uncertainty can create incentives for otherwise well-motivated agents to take actions contrary to their principals' interests. It is this dynamic that constitutes the core of the over-accountability problem. The action which a perfectly-informed principal would want the agent to take is not necessarily the action that would most effectively convince imperfectly-informed principals that the agent is a good type. In other words, accountability mechanisms create incentives for agents to make decisions that *principals expect good types to make*, which may not be the same as the decisions that are *actually in the principals' best interests*.

Lest the discussion become too abstract too quickly, consider a stylized example. Suppose an administrative agency must decide whether to impose regulatory burdens on some industry, such as automobile manufacturers. Recently, for example, the Department of Transportation considered whether to require the installation of rear-facing video cameras in all new cars to prevent potential accidents when cars are driven in reverse. For convenience, assume that the policy choice is binary: regulate (require the cameras) or do not. We will treat the Secretary of Transportation as the "agent" and the President (representing the citizens) as the "principal."¹⁷

The correct decision depends on what the political economy jargon refers to as the "state of the world" (or simply the "state"), which is a shorthand way of expressing all the complex factors that determine the impact of different policy choices on outcomes. Assume there are only two possible states, each of which is equally likely to occur: in one state the best policy for the citizens (and the President) is to require the cameras; in the other state the best policy

16 Indeed, that statement is essentially a tautology, in that the goodness or badness of an agent's type, in this context, is defined in terms of the principal's expected utility, going forward, from selecting that agent. There are two qualifications to the statement that the "goodness" of an agent's type is by definition a measure of the principal's interest in retaining that agent. First, in some of the political economy literature the agent's "type" refers *only* to those characteristics about which the principal is uncertain. That qualification is mostly semantic. Second, a principal often has an incentive to pre-commit to a retention strategy that requires her sometimes to retain an agent who is a worse "type" than a challenger, or to dismiss an agent who is a better "type" than the challenger, because committing to such strategies may create desirable ex ante incentives. There is an ongoing theoretical and empirical controversy over whether (or when and how) voters can credibly commit to such strategies.

17 Of course, there is also a potential agency problem between the citizens and the President, but we suppress that here for purposes of keeping the example simple.

is not to do so. The President, however, does not know the state. So, from her perspective, there is a 50 percent chance that the cameras should be required and a 50 percent chance that they should not.¹⁸

The Secretary of Transportation, let us suppose, can observe the state or at least learn about it.¹⁹ What the Secretary prefers to do with that information depends on her type. Suppose the Secretary may be one of two types: she may be “captured” by the industry, or she may be “public-spirited.”²⁰ A captured Secretary always chooses non-regulation (no imposition of costs on industry) regardless of the state or what this choice means for her ability to keep her job in the future. A public-spirited Secretary, by contrast, prefers the same policy that the (presumptively public-spirited) President would prefer if the latter knew the state. That is, the public-spirited Secretary wants to require the installation of cameras in all cars if, in some sense, it is “worth it.”

A public-spirited Secretary gets some benefit both from doing the right thing and from being perceived by the President as public spirited, as the latter perception enhances her prospects for keeping her job (or reaping other rewards). If this latter concern were not present—if the Secretary were not *accountable* to the President—then the public-spirited Secretary would always do the right thing (from the President’s perspective, as well as her own). But the accountability relationship changes her calculations: her desire to *appear* public spirited may cause her to make a decision that is in fact contrary to the President’s interests (and her own).²¹

18 To keep things simple, suppose the President’s utility is symmetric, in that its payoff from a correct decision to regulate is equal to its payoff from a correct decision not to regulate, and its (lower) payoff from an incorrect decision is the same whether that decision is erroneous regulation or erroneous non-regulation. Nothing of import turns on this symmetry assumption.

19 The assumption of omniscience is obviously unrealistic, but all that is important for the example is that the agent is *somewhat* better informed about policy consequences than the principal; the easiest way to represent that asymmetry, without introducing a lot of extra math, is to assume the agent’s information about the state is perfect.

20 We assume each type of agent is equally likely to hold office, and the President does not observe the Secretary’s type.

21 Two quick clarifications about our terminology: It may seem odd to say that a “public-spirited” agent would make a bad decision because of personal reputational concerns. But when we describe the agent as “public spirited” we mean only that her policy preferences align with those of her principal; she also may be self-interested, caring about her own reputation as well as (or possibly more than) getting the policy right. It may also seem odd that we assume the captured agent is not influenced by accountability concerns. This assumption is purely to simplify the example. Later we consider cases in which all agent types may care about both policy outcomes and about their reputations with their principal.

Figure 1. Predicted behavior if the public-spirited agent does the right thing, while the captured agent always chooses nonregulation.

	Regulation is the best policy (50% chance)	Non-regulation is the best policy (50% chance)
Agent is captured (50% chance)	<i>No Regulation</i> (25% chance)	<i>No Regulation</i> (25% chance)
Agent is public-spirited (50% chance)	<i>Regulation</i> (25% chance)	<i>No Regulation</i> (25% chance)

To see why, suppose that the public-spirited agent always does the “right thing” by picking the policy that is appropriate given the state of the world, while the captured agent always chooses no regulation.²² If so, then when the President observes the Secretary of Transportation choose to regulate (require the installation of cameras), the President can infer that the Secretary must be a public-spirited type, because in this example *only* a public-spirited agent ever chooses regulation. In contrast, if the Secretary chooses not to require the cameras, the President would conclude there is a 2/3 probability that she has been captured by the industry, and a 1/3 probability she is public-spirited. Why? Because the 75 percent chance that the agent chooses deregulation is composed of a 50% chance that the agent is captured and a 25 percent chance that *both* the agent is public-spirited *and* non-regulation is in fact the right policy.²³

22 If this is how the agents behaved, then the policy is “correct” 75 percent of the time: the public-spirited agent would always do the right thing, while the captured agent would (by happenstance) choose the right policy half the time.

23 More technically, this calculation calls for an application of Bayes’ Rule, according to which the conditional probability that X is true given that Y is true is equal to the conditional probability that Y is true given that X is true, multiplied by the unconditional probability that X is true, divided by the unconditional probability that Y is true. That is, $Pr(X|Y) = Pr(Y|X)Pr(X)/Pr(Y)$.

Figure 1 illustrates this. The principal cannot directly observe which row or column she is in, but she can observe the final policy outcome (regulation/non-regulation). So, if the principal observes regulation, she knows she must be in the bottom-left cell (meaning the agent is public-spirited), while if the principal observes non-regulation, she knows she must be in one of the other three cells, which implies a $2/3$ chance she is in the top row (captured agent) and a $1/3$ chance she is in the bottom-right cell (public-spirited agent).

The implication is that an incumbent agent is more likely to be rewarded if she chooses to regulate: if she regulates, the principal will conclude she is definitely a good type, while if she opts for non-regulation the principal will conclude she is probably a bad type. So, if the public-spirited type cares enough about her reputation with the principal—for example, if this is important to her chances of keeping her job—she will choose regulation even when she knows that everyone would be better off without regulation.²⁴ The important point of this example is that accountability can induce the “good type” to take the *wrong* action when the right action would be disproportionately preferred by “bad types” in the absence of accountability.²⁵

This example is but one illustration of the general phenomenon we are interested in: an agent’s incentive to *appear* to be a better agent can sometimes lead her to *act* like a worse agent—even though the agent’s decisions are rational responses to the principals’ equally rational decisions whether to retain or replace (or, more generally, to reward or punish) an incumbent agent. A variety of perverse behaviors arise from this same underlying “over-accountability” dynamic.

The remainder of Section 1 will explore two aspects of this sort of over-accountability. Section 1.1 considers how the desire of accountable agents to appear *competent* may lead them to take actions that they believe are not in the principal’s best interests. Section 1.2 considers how the desire of accountable

24 That is, there is a stable outcome supported by rational beliefs and behavior (an “equilibrium”) in which the captured agent always chooses non-regulation and the public-spirited agent always chooses regulation. Note that in this equilibrium, the agent makes the right decision (from the principal’s perspective) only 50 percent of the time.

25 The statement in the text is somewhat loose. More precisely, there is a reputational incentive for an agent to choose action *X* over action *Y* if, from the principal’s perspective, the conditional probability that an agent is a good type given that the agent chose *X* is greater than the conditional probability that an agent is a good type given that the agent chose *Y*.

agents to appear *unbiased* may likewise induce agents to take actions that are not in the principal's best interests.²⁶

1.1 Signaling Competence: Pandering, Posturing, and Persistence

1.1.1. *Pandering: Taking the Popular Action to Avoid Seeming Incompetent*

Despite the frequent veneration of the “will of the people,” our political discourse regularly bemoans the propensity of politicians to “pander” to the electorate.²⁷ “Pandering” has many meanings, but in the political context it is usually taken to mean catering excessively to the desires or predispositions of some group. The objection to “pandering” is that such behavior amounts to doing what is popular, rather than doing what is right.²⁸ Understanding the phenomenon of over-accountability may help us understand how and why this occurs.²⁹ In particular, we can understand pandering as occurring when an agent tries to signal that she is a “good type” by choosing the popular action, even when the agent believes that the unpopular action is more likely to be the one that the principals (e.g., voters), if fully informed, would have preferred.³⁰

What we label “pandering” behavior may arise even if the principal is confident that all agents have the “correct” policy preferences, because agents may nonetheless differ in their *competence*—that is, their ability to discern which policies will best serve those preferences.³¹ If the principal prefers a more competent agent, then an incumbent agent has an incentive to choose whichever policy will make her appear more competent. Under some circumstances, a less-competent agent's interest in appearing competent will lead her to

26 The division between considerations of competence and bias is somewhat artificial, but the expositional division between Sections 1.1 and 1.2 is useful in clarifying the different mechanisms at work. These mechanisms are by no means mutually exclusive, and often several may be in play at once.

27 See, e.g., David Brooks, *What Moderation Means*, N.Y. TIMES (October 26, 2012); *Mr. Romney Stumps in Israel*, N.Y. TIMES (July 31, 2012); *Republicans and the Gun Lobby*, N.Y. TIMES (April 14, 2012); *Mr. Sarkozy on the Low Road*, N.Y. TIMES (March 15, 2012). But see Lawrence R. Jacobs & Robert Y. Shapiro, *Debunking the Pandering Politician Myth*, THE PUBLIC PERSPECTIVE (APRIL/MAY 1997).

28 See, e.g., Jacobs & Shapiro (2000).

29 One difficulty with a generic political term like “pandering” is that such terms may describe a form of behavior that can have a variety of distinct causes. Some forms of political pandering involve emotional appeals rather than rational appeals, or may exploit voters' psychological bias for short-term results. In contrast, we suggest a rationalist explanation of pandering.

30 The discussion in this subsection is based primarily on Canes-Wrone et al. (2001), and Maskin & Tirole, (2004). See also Ashworth & Shotts (2010); Prat (2005); Scharfstein & Stein (1990).

31 Another possibility is that an agent feels compelled to try to signal to her principal that she shares the principal's preferences. We discuss this sort of behavior later, though under the different labels of “populism” and “political correctness.” See Section 1.2, *infra*.

“pander” by choosing the popular action, even though the agent thinks that some other, less popular action would better serve the principal’s interests.

A modification of our earlier example can illustrate the dynamic. Suppose that rather than a bureaucrat, an elected politician must choose between one of two policies, which again might be thought of as regulation and non-regulation. Suppose regulation is more likely to be the right policy: there’s a 60 percent chance that voters’ interests would be best-served by regulation, and a 40 percent chance that voters would benefit more from non-regulation.³² We can therefore call regulation the “popular” policy and non-regulation the “unpopular” policy, though it is important to emphasize that regulation is “popular” in this example not because the voters have an intrinsic taste for it, but rather because voters believe that regulation is more likely to be in their interests.³³

Suppose there are two equally likely types of agent: “normal” agents and “expert” agents. The experts (the good types) always know the right policy to adopt. Normal agents (the bad types) have more information than the voters, but are not as reliable as experts. Instead, normal agents get some indication of the right decision (a “signal”) that is accurate 75 percent of the time, but inaccurate 25 percent of the time. This means that if a normal agent gets the signal “regulation,” the probability that regulation is in fact the right decision is approximately 82 percent.³⁴ If, on the other hand, the normal agent gets the signal “non-regulation,” the probability that non-regulation is the right decision is approximately 67 percent.³⁵ Therefore, voters not only prefer that an expert agent does what she believes to be right, but also that a normal agent does what she believes to be right.

But do the agents have an incentive to take what they think is the correct action? If agents care enough about their reputations—that is, if they care enough about the voters thinking that they are experts—then the answer may be no. To see why, suppose that the politician in the preceding example always took the action that she believed to be correct. This means that experts would regulate 60 percent of the time (because regulation is in fact the right decision

32 As in the earlier example, we adopt the simplifying assumption that voters’ payoffs are symmetrical: Each voter gets the same payoff when the agent makes the “correct” decision, whether that decision is regulation or nonregulation; likewise, voters’ payoff from incorrect regulation is equal to their payoff from incorrect nonregulation.

33 It is unfortunately quite common—even in otherwise sophisticated scholarly analyses—to conflate first-order preferences over outcomes with beliefs about what actions would best serve those preferences.

34 Applying Bayes’ Rule, the conditional probability that the true state is “regulate” given that the signal is “regulate” is $(0.75)(0.6)/[(0.75)(0.6)+(0.25)(0.4)] = 9/11$.

35 Applying Bayes’ Rule, the conditional probability that the true state is “deregulate” given that the signal is “deregulate” is $(0.75)(0.4)/[(0.75)(0.4)+(0.25)(0.6)] = 2/3$.

60 percent of the time). Furthermore, if a normal (non-expert) politician always follows her signal—which, again, is what the voters would want her to do—then a normal politician would regulate 55 percent of the time.³⁶ So if all agents did what they thought was right, normal agents would choose non-regulation (the “unpopular” policy) more frequently than would experts, while experts would choose regulation (the “popular” policy) more frequently than would normal agents. Thus voters, upon observing that the incumbent chose to regulate, would revise their assessment of the probability the incumbent is an expert upwards from 50 percent (the voters’ initial probability assessment, or “prior”) to approximately 52 percent (the voters’ “posterior” assessment).³⁷ In contrast, if the politician chose non-regulation, voters would adjust their estimate of the probability she is an expert downward from 50 percent to approximately 47 percent.³⁸ If all agents did what they thought was right, then those who chose the popular action would be perceived as more likely to be experts.³⁹ Therefore, agents who care enough about their reputations will sometimes choose the popular action even though they believe the unpopular action is better policy. More specifically, a normal agent may sometimes “pander” by doing what is popular (*ex ante*), rather than doing what the agent believes is right.⁴⁰

From the voters’ perspective, this sort of pandering is bad for two reasons. First, the government is more likely to make the wrong decision. In the above example, a normal agent who thinks non-regulation is the right decision may nonetheless choose regulation. The second problem is that this sort of pandering makes it harder for the voters to sort good types from bad types. If all agents

36 The probability that the normal agent *correctly* chooses regulation is $(0.6)(0.75) = 0.45$, while the probability that the normal agent *incorrectly* chooses regulation is $(0.4)(0.25) = 0.1$; the sum of these probabilities is 0.55.

37 Applying Bayes’ Rule, the conditional probability that the agent is an expert given that she chose to regulate is $(0.6)(0.5)/[(0.6)(0.5)+(0.55)(0.5)] = 12/23$.

38 Applying Bayes’ Rule, the conditional probability that the agent is an expert given that she chose to deregulate is $(0.4)(0.5)/[(0.4)(0.5)+(0.45)(0.5)] = 8/17$.

39 Note, however, that the agent’s choice also changes the voters’ perception of which policy is likely correct. Perhaps somewhat counterintuitive, if the agent chooses to deregulate, the voter changes her mind and concludes that deregulation, though initially unpopular, was probably the correct policy after all, so that policy becomes popular after the fact. If the agent chooses regulation, then the voter would be even more confident that regulation was the right policy (that is, regulation becomes even more popular).

40 The normal agent will not pander all the time, at least not in this example. If she did, voters would conclude that an agent who took the *unpopular* action (deregulation) was more likely to be an expert. Instead, the normal agent will pander just often enough that the electoral benefits of choosing regulation offset the policy benefits of following her signal and choosing deregulation. That is, the equilibrium here involves the agent playing a “mixed strategy,” in which players randomize among several possible choices.

did what they thought was right, then experts (who in the above example choose regulation more frequently) would be more likely to be reelected than would normal agents. But if normal agents sometimes pander, the relative frequencies with which different types choose regulation or non-regulation become more similar; as a result the voters learn less about the politician's type from her choice.⁴¹ So, not only is regulatory policy distorted, but electoral selection becomes less efficient.

We have explored this sort of political pandering in using a stylized example, but such pandering behavior seems plausible in a range of real-world contexts of interest. The electoral context is the most obvious, but pandering of this sort may also arise within the government. Consider, for example, presidential control of the bureaucracy. Agencies have local policy expertise; the President would like the agents to take full advantage of their private information; agency officials derive utility from a mix of getting things right and keeping their jobs, and their probability of keeping their jobs (or getting promoted, or getting more authority) may depend on whether the President perceives them as highly competent. A rational President will update her beliefs about an agent's competence based on the policy the agent adopts. The result is the risk of intra-executive pandering behavior: subordinate agents may advocate the policy that the President believes is more likely to be correct, but that the subordinate actually believes to be incorrect. When this occurs, the policies selected by administrators will not always reflect true bureaucratic expertise, but rather the bureaucrat's desire to appear expert.

1.1.2. *Posturing: Taking Bold Action to Avoid Seeming Incompetent*

If pandering is a distortion in which agents do what is popular rather than what is right, then posturing might be thought of as a distortion in which agents do what is bold and attention-grabbing rather than what is conventional and sensible. At least one form of behavior that might be characterized as posturing involves agents taking or proposing bold, dramatic action for the sake of being seen as the kind of agent who takes or proposes bold, dramatic action, rather than because the agent actually believes such action is appropriate under the circumstances.⁴² Politicians may want to be seen as the sort of leader that

41 For instance, if in the above example normal agents were to pander frequently enough that both experts and normal agents select regulation 60 percent of the time, then voters would learn *nothing* from the agent's policy choice.

42 The discussion in this subsection is based primarily on Fox & Stephenson (2011) and Levy (2004). See also Canes-Wrone et al. (2001) (describing this phenomenon as "fake agentship"); Levy (2005).

John F. Kennedy celebrated in his hagiographic *Profiles in Courage* (Kennedy 1955)—willing to take unpopular actions that are in the country’s best interests—or as independent-minded “mavericks.” At first blush, pandering and posturing would seem like quite different, indeed contradictory, behaviors. Pandering involves catering to mainstream opinion and beliefs, while posturing (of the sort we describe above) entails strategic defiance of mainstream opinion and beliefs. But both sorts of behavior can spring from the same underlying accountability-induced incentive to appear more competent.

To illustrate, suppose that the incumbent agent is a Prime Minister (PM) in a parliamentary government. The PM shares the parliament’s policy preferences, and can choose one of two possible actions, with the correct action depending on the state of the world. Suppose that one of those actions—the “popular” or “routine” action—is the one that parliament believes is most likely to be appropriate; this might simply be “staying the course” by retaining the status quo. The other action—the “extraordinary” action—is appropriate only under certain conditions that the majority coalition views as unlikely but not unthinkable. Examples of such extraordinary actions might include imposing emergency price controls, quarantining citizens, initiating military intervention abroad, and so forth. Suppose there is an 80 percent chance that the routine action is best (“normal” circumstances) and a 20 percent chance that circumstances call for the extraordinary action (“extraordinary” circumstances), rank-and-file members of parliament (MPs) cannot tell the difference. Because normal circumstances are more likely, parliament would choose the routine action if it had to make the decision on its own, and would prefer the PM to take routine action unless she has compelling information indicating extraordinary circumstances.⁴³

“Expert” PMs can recognize extraordinary circumstances that call for extraordinary action, but—in contrast to our pandering example—“normal” PMs don’t know any more than rank-and-file MPs. This implies that if there were no reputational incentives, expert PMs would follow their (accurate) signal, while normal PMs would always take the routine action. If PMs behaved this way, the PM would choose the correct policy (from parliament’s perspective) 90 percent of the time: half the time the PM is an expert who makes the right decision, and half the time the PM is a normal type who takes the routine action, which is correct 80 percent of the time. In 10 percent of the cases—when the PM is

43 We are assuming here, as before, that parliament’s benefit from the correct action is symmetric, as is its cost from the incorrect action.

normal and the circumstances call for extraordinary action—the government makes the wrong decision, staying the course when bold action is required.⁴⁴

But suppose the PM, who is accountable to parliament, cares sufficiently about her reputation. Consider what inferences parliament would draw from each of the PM's possible decisions if all agents did the right thing (given their information). If the PM takes extraordinary action, then parliament would infer that the PM is an expert, because only an expert would ever have enough information to justify such action. If, however, the PM takes routine action, then parliament would believe there is approximately a 44 percent probability that the PM is an expert and approximately a 56 percent probability that she is normal.⁴⁵ Normal PMs who care sufficiently about their reputations might therefore sometimes “posture” by taking extraordinary action even though they lack sufficient information to justify such action—not all the time, but often enough that their behavior looks similar to that of experts. Posturing leads to more incorrect decisions; furthermore, because posturing entails normal agents mimicking the behavior of expert agents, principals can glean less information from the agent's decision, and prospective selection is less efficient.

1.1.3. *Persistence: Signaling Competence through Stubbornness*

Given the complexity of modern policy problems, there is much to be said for a flexible approach in which governments can learn from their mistakes and alter policies that are not working well.⁴⁶ Yet governments often persist in policies long after evidence comes to light that these policies are unsuccessful. While there are many possible explanations for stubborn adherence to a failed policy, one explanation has its roots in the same kind of over-accountability that gives rise to pandering and posturing. Even though principals may want agents to take an experimentalist approach and learn from their failures, principals may end up thinking more highly of agents who stick with a policy over time

44 It's worth noting here that, as in the pandering case, the popularity of an action changes after the agent makes her choice: although the extraordinary action is unpopular *ex ante* (there's an 80 percent chance that it's a bad idea), after the agent takes the action it becomes popular. If we assume that all agents do what they believe is correct, then the principal is 100 percent certain that if the agent chose the extraordinary action it must have been correct. We can think of this as a kind of “rally” effect, though the mechanism is distinct from the more familiar psychological accounts of such effects.

45 Applying Bayes' Rule, the conditional probability that the agent is an expert given that the agent chose the routine action is $(0.8)(0.5)/[(0.8)(0.5)+(1)(0.5)] = 4/9$, while the probability that the agent is normal given that the agent chose the routine action is $5/9$.

46 See Listokin (2008).

precisely because this can be indirect evidence that the agents in question are good at distinguishing good policies from bad policies.⁴⁷

Again, the idea is easiest to see in the context of a stylized example. Consider an agency—accountable to voters, political overseers, or both—that is considering experimenting with privatizing a service that had been traditionally provided by the government. For simplicity, we will again consider this as a simply binary decision (privatize/don't privatize). Suppose, as before, that there are two types of potential agents responsible for this decision: normal administrators and expert administrators, both of which share the principals' preferences (say, for more efficient, higher-quality services). Suppose there is a 60 percent probability that government provision (the status quo policy) is superior to privatized provision.

In contrast to the earlier examples, in which the agent made a single policy choice, here we assume that the agent makes an initial policy choice (at "time 1") and then has an opportunity later (at "time 2") to revise that policy. Before each of these decision points, the agent gets some additional information about which policy is better, which (again using the standard jargon) we will call a signal of the state. The expert's signal is correct 100 percent of the time, but the normal administrator's signal is correct 80 percent of the time. For simplicity, we assume that the state does not change and that the agent's signal does not depend on which policy is in place.⁴⁸ Nonetheless, the normal administrator might get a different signal in each time period, because her signal is not always accurate.

Because the state does not change and the expert always knows the state, the expert never has an incentive to reverse herself in time 2. This is not true for the normal administrator, however. If the normal administrator acted as a perfect agent, her behavior would be as follows: If her first-period signal indicates government provision (the status quo) is the correct policy, she will maintain that policy, no matter what her second-period signal indicates.⁴⁹ If, however, her first-period signal indicates privatization is better, she will privatize, which

47 The discussion in this subsection is based principally on Majumdar & Mukand (2004). See also Prendergast & Stole (1996).

48 We also assume, again purely for simplicity, that the principal does not get any additional signal of the state. The results we discuss here would be qualitatively similar if the principal also got some additional information about the success or failure of the policy, so long as the agent's information was sufficiently more accurate.

49 Even if the agent's time 2 signal indicates deregulation is the correct policy, her two signals will cancel out, so her best guess as to the right policy is the status quo regulation policy, which has a 60 percent chance of being correct.

she estimates has a 73 percent chance of being the correct choice.⁵⁰ If her second-period signal also favors privatization, she will adhere to that policy, but if her second-period signal indicates that public provision is the superior alternative—that is, if additional evidence suggests the privatization policy is not effective—then a normal administrator acting as a faithful agent would reverse course and re-establish public provision. This is because if the time 1 and time 2 signals point in opposite directions, the normal agent is back where she started, information-wise: there's a 60 percent chance that public provision is the right policy, and a 40 percent chance privatization is preferable. Thus, the principal would want a normal agent to experiment with privatization if her early information suggests this might be a good idea, but also would want the agent to reverse course if sufficient subsequent information cuts in the other direction.

However, the normal administrator's reputational incentives to appear expert may prevent her from acting this way. The reason is easy to see in this stylized example: Because expert administrators are more likely to get the policy decision right the first time, their decisions are more likely to be consistent over time than would be the case for less-capable administrators who are doing their best to respond to new information as it comes to light. This creates an incentive for the normal agent who chose to privatize government services in the first period to adhere to that decision in the second period, even when new information suggests that choice was a mistake. This behavior hurts the principals both by leading to suboptimal policies, and by weakening the efficacy of the hiring and firing as a selection mechanism.⁵¹

1.2. Signaling Lack of Bias

1.2.1. *Populism: Avoiding the Appearance of Excessive Sympathy to an Unpopular Group or Cause*

Section 1.1 focused on situations in which agents shared the principals' underlying preferences; the relevant difference between good and bad agents was

50 Applying Bayes' Rule, the conditional probability that deregulation is the correct policy given the normal agent's signal of "deregulate" is $(0.8)(0.4)/[(0.8)(0.4)+(0.2)(0.6)] = 8/11$.

51 This discussion implies an additional form of over-accountability, which might lead to excessively frequent policy changes. Suppose that instead of expert types and normal types, we assume that "bad" types have some intrinsic degree of stubbornness. In that case, adhering to a prior policy might be interpreted as a sign that the agent is a bad (stubborn) type. The desire to appear flexible, open to fresh ideas, and capable of self-criticism might then lead agents to change policy too often. This possibility clearly has a close kinship with the posturing behavior described previously, much as the persistence perversion described in the main text has a close kinship with pandering. (For a version of this argument, see Majumdar & Mukand (2004); Prendergast & Stole (1996).) We relegate this observation to this footnote, partly because casual observation suggests that this sort of perversion, though conceptually possible, seems less empirically relevant.

their competence in predicting policy consequences. However, principals are concerned not only about the possibility of a less *competent* agent, but also about the possibility of a *biased* agent. Such bias may arise either from the agent's own preferences, or because the agent has been captured, corrupted, or unduly influenced by a parochial interest group. Because principals have an incentive to weed out biased agents, an accountable agent has an incentive to signal that she is loyal to the principal and is not ill-motivated or biased or captured.

How might the agent do this? One possibility is that she might go out of her way to harass, burden, or persecute unpopular groups in order to credibly signal that she has not been captured by, and is not secretly sympathetic to, those groups.⁵² For convenience, we label this behavior "populism." It is important to emphasize that it is difficult to say whether this behavior is intrinsically problematic; this judgment may require an independent normative evaluation of the relative value of the unpopular group's interests. Even if one focuses only on the principal's welfare, however, an agent's incentive to engage in reputation-enhancing populism may lead agents to take actions that are harmful to the principal.

The stylized example in the introduction to Section 1 illustrates this phenomenon. In that example, some agents ("good types") were public-spirited, while others ("bad types") were captured by a parochial interest group. Because of this, selecting the policy that burdens the interest group may boost the agent's reputation because it suggests the agent is not beholden to that group. This means that even a public-spirited agent may sometimes take the "wrong" action (in the example, imposing burdens on the interest group even when doing so is not in the principal's interests). Imposing costs on the unpopular group is a way for an agent to signal that she is not secretly sympathetic to that group, and this creates a reputational incentive for agents to place burdens on unpopular groups even when those burdens go beyond what the principal, if fully informed, would actually prefer.⁵³

As was true of the distortions discussed earlier, this problem affects not only the government's policy on the particular issue, but also on principals' ability to sort good (unbiased) types from bad (biased) types. In the examples offered in

52 The discussion in this section is based principally on Acemoglu, Egorov & Sonin (2013) and Fox & Stephenson (forthcoming 2014b). See also Morelli & Van Weelden (unpublished manuscript, 2014); Negri (unpublished manuscript, 2011).

53 See, e.g., Lopez (2011) at 1014; Acemoglu, Egorov & Sonin (2013); Fox & Stephenson (forthcoming 2014b).

the pandering, posturing, and persistence subsections, the distortion in agent behavior not only led agents to make suboptimal decisions on the policy issue at hand, but also weakened sorting by making the expected choices of different agent types more similar. By contrast, in the example offered in the introductory overview—which was an example of the populism mechanism at work—the distortion in the public-spirited agent’s behavior facilitated sorting, even though it sometimes resulted in bad policy choices on the policy decision at hand. When this is so, there may be a trade-off between the principals’ desire to get the right answer on a particular policy issue and the principals’ desire to accurately distinguish good types from bad types (which might affect the welfare on any number of future policy issues). That said, it is also possible to construct scenarios in which the incentive to engage in populism worsens sorting.⁵⁴

The populism distortion may also manifest itself in agents’ choices as to which policy issues to address.⁵⁵ Although one might conjecture that more accountable agents would focus on those issues of greatest importance to their principals, the incentives that accountability mechanisms create for agents to distinguish themselves as “good types” may cause agents to focus on less important, more divisive issues. Consider a society with two groups, a majority and some politically unpopular minority. Suppose there are two issues the agent could focus on while in office: the economy, and the legal rights of the minority group. Suppose that voters care much more about the economy than

54 In the introductory example, this did not occur because of the assumption that the captured type *always* adopts the policy favored by industry. If even captured agents have strong reputational incentives to appear public-spirited, then even they might turn on the industry and adopt a stringent regulatory policy. If this happens, all agents choose regulation regardless of the state, and voters learn nothing. (This equilibrium is sustained by the principal’s belief—never challenged in equilibrium—that if an agent chooses deregulation she must be a captured type.)

Importantly, if the agent were not constrained in terms of how oppressive the policy can be toward the unpopular group, then this sort of convergence may not be possible. If a public-spirited agent is always more willing than the captured agent to impose burdens on the industry, then if a captured agent were to try to mimic a public-spirited agent by adopting an anti-industry policy, the public-spirited agent could distinguish herself by adopting a policy that was just a bit more anti-industry. If a captured agent were to try to mimic that behavior as well, the public-spirited agent could escalate a bit more. At a certain point, the captured agent will not be willing to go so far. The result is that the agent—who is “public-spirited” in the sense that she shares the principal’s policy preferences, but also “self-interested” in the sense that she also cares about her reputation and career—may end up adopting a policy that is more extreme in its anti-industry bias than what the principal would prefer. See Acemoglu, Egorov & Sonin (2013); Fox & Stephenson (forthcoming 2014b).

55 See Morelli & Van Weelden (unpublished manuscript, 2014); Negri (unpublished manuscript, 2011).

about the minority rights issue. Suppose further that on the minority rights issue, the preferences of the majority and minority groups diverge, and members of the majority group prefer agents who share their preferences (“good types”) rather than those who sympathize with the minority (“bad types” from the majority’s perspective). Suppose also that it’s harder for agents to translate effort on the big issue—the economy—into visible results than it is for them to secure results on the less important and more divisive minority rights issue.

Under these assumptions, agents may focus on the divisive minority rights issue rather than the economy, even though the latter is more important even to the majority group. The logic should by now be familiar: If both agent types did what the voters would prefer them to do, they would both focus on the economy. However, this would not necessarily guarantee good economic results, nor would pro-majority types be able to distinguish themselves effectively. If, however, a pro-majority agent were to focus her attention on imposing burdens on the unpopular minority, voters would be able to infer more confidently that the agent is indeed pro-majority. Paradoxically, even though the majority-group voters would have preferred that the agent focus on the economy, voters in this scenario may be more likely *ex post* to reward agents who devoted their energies to attacking a minority group.

We see a possible real-world illustration of this phenomenon in the behavior of elected state court judges in criminal cases. Empirical studies have found that judges render decisions differently as election time nears, handing out less lenient sentences immediately preceding an election than immediately after.⁵⁶ While one cannot confidently say whether sentences are ordinarily too lenient, or instead whether they are too harsh around elections, it is plausible that this judicial behavior derives from the form of over-accountability we have labeled “populism,” in which elected judges try to demonstrate that they are “law-and-order” types by imposing sentences that are too harsh relative to the underlying conduct. It is at least possible—though we cannot establish it here—that the harshness of the sentences goes beyond what mainstream voters would actually prefer, even though voters rationally reward such harshness at the ballot box.

Although our main focus has been on how agents treat unpopular groups, the same basic mechanism may also be at work when an accountable leader negotiates with outside parties; the above analysis of populism can be transposed, *mutatis mutandis*, to this context by replacing the “unpopular group” with the “adversary.” If the agent’s constituency is concerned that she may be a “bad type” who is overly sympathetic to the adversary’s interests, or somehow “captured” by that adversary, then the agent has an incentive to be overly

56 See Gordon & Huber (2004); Gordon & Huber (2007).

aggressive and uncompromising in order to signal that she is the good type (“tough” on crime, communism, terror, or what have you) rather than the bad type (“soft” on same). We probably wouldn’t refer to this sort of behavior as “populism” (it seems more like a kind of “posturing,” albeit one that differs from that discussed in subsection 1.1.2), but the mechanism is essentially the same.⁵⁷

1.2.2. *Political Correctness: Avoiding the Appearance of Antipathy to a Sympathetic Group or Cause*

Sometimes, citizens may be worried not so much about an agent’s bias *in favor* of an unpopular group or cause, but rather about the agent’s possible bias *against* a sympathetic group or cause. When agents are anxious to avoid the appearance of this latter sort of bias, another type of over-accountability distortion—a form of “political correctness”—may arise.⁵⁸ This distortion is structurally similar to the populism problem discussed previously, even though its consequences seem superficially the opposite.⁵⁹

Suppose an agent must choose between two policies, one of which would benefit a sympathetic group and the other of which would burden that group. For example, the decision might be whether to implement an affirmative action program for traditionally disadvantaged racial minorities. Suppose, for

57 That perspective may help explain the puzzling phenomenon whereby organizations, such as labor unions or other professional associations, sometimes go to seemingly self-defeating lengths to defend malfeasant rank-and-file members. For example, the New York City teachers’ union defended teachers credibly accused of extended absenteeism or other significant underperformance, despite the negative effect that these few bad teachers imposed on the vast majority of good teachers, see Campbell Brown, *Teachers Unions Go to Bat for Sexual Predators*, Wall St. J. (July 29, 2012), and the NFL players’ union vigorously contested the sanctions imposed on the New Orleans Saints defensive players who deliberately tried to injure opposing players, even though the union’s own members were the victims as well as perpetrators of the scheme, see *Union Asks for Saints Discipline Delay*, ASSOCIATED PRESS (March 13, 2012). One explanation may be an accountability perversion of the sort described in this section. Union members need to distinguish “good” union agents (those willing to fight hard for union interests) from “bad” agents (those overly willing to “sell out”). One way agents can signal that they are “good” is to refuse to give management what it wants without a hard fight in *every* case—even when the union agentship is aware that the underlying conduct that management wants to penalize was wrongful or warrants action. Even when high-quality union agents recognize that the best course of action for the union is to sanction (or at least not vigorously defend) certain members, that course of action increases the risk that the membership will conclude that the leadership is weak if the agentship is weak or in the pocket of management. Thus, a blanket rule to vigorously defend any and all members may be more desirable for union agents seeking retention, even when this imposes net costs on the rest of the membership.

58 The discussion in this section is based principally on Morris (2001). See also Loury (1994); Bernheim (1994); Avery & Meyer (unpublished manuscript, 2012).

59 In that sense, the relationship between populism and political correctness parallels the relationship between pandering and posturing discussed earlier.

purposes of this example, that most voters harbor no ill-will toward the minority group, and indeed favor policies that would promote nondiscrimination and some form of redress for past wrongs. Suppose further, however, that majority-group voters are uncertain whether the affirmative action policy in question is justified. After all, there is a vigorous debate about whether such policies are helpful or counterproductive.⁶⁰ For the sake of the example, suppose that majority-group voters think there's a 50 percent chance that the affirmative action policy under consideration is a good idea, and a 50 percent chance that it's a bad idea (from the majority group's perspective).

Now suppose the agent can be one of two types. There's an 80 percent chance that the agent is a "good type"—a fair-minded representative of the majority group—who would prefer the pro-minority policy under the same circumstances that the median voter, if fully informed, would prefer it. However, there's a 20 percent chance that the agent is a "bad type"—an anti-minority bigot who has nonetheless tried to conceal her anti-minority views. Suppose that the bad agent would *always* oppose the affirmative action policy (though she would always *claim* that she is fair-minded and that the affirmative action policy is simply a bad idea). Suppose also that the majority-group voters, while perhaps not as sympathetic to the minority group as would be ideal, have a strong preference for a fair-minded agent rather than a bigot, and are much more likely to reward the agent (say, by retaining her) if they think she is the former type. What is the effect of accountability on agent behavior in this case?

Absent accountability concerns, bigoted agents would always oppose affirmative action, while fair-minded agents would have a 50 percent chance of supporting it. (As a result, the government would adopt the correct policy 90 percent of the time.⁶¹) Thus, if the agent were to endorse the affirmative action policy, voters would know that the agent is fair-minded. If the agent were to reject the affirmative action policy, voters would conclude that there is a 1/3 chance that the agent is bigoted.⁶² If agents care enough about not appearing to be bigots, then some agents who privately disapprove of affirmative action might nonetheless publicly support it.

60 For a summary of positions in the context of *Fisher v. Texas*, a case involving affirmative action in higher education, see *Beyond Race in Affirmative Action*, <http://www.nytimes.com/roomfordebate/2012/02/22/beyond-race-in-affirmative-action>.

61 In the 80 percent of cases where the agent is fair-minded, the policy is always correct, while in the 20 percent of cases where the agent is a bigot, the agent's rejection of affirmative action is in fact the correct choice 20 percent of the time, so the probability of a correct outcome when all agents do what they think is right is $(0.8)(1) + (0.2)(0.5) = 0.9$.

62 Applying Bayes' Rule, the conditional probability that the agent is a bigot given that she chose to reject the affirmative action policy is $(1)(0.2) / [(1)(0.2) + (0.5)(0.8)] = 1/3$.

Whether this is a good or bad thing for the majority group depends on how these reputational incentives affect behavior. One possibility is that bigoted agents care more about opposing affirmative action than do fair-minded agents who have concluded that the affirmative action policy would do more harm than good. In that case, reputational pressures will lead fair-minded agents, but not bigoted agents, to endorse affirmative action even when they think it as bad idea. This corresponds most closely to “political correctness” in the pejorative sense: the fear of being perceived as biased leads fair-minded individuals to support policies that they think are unwise. Indeed, if fair-minded agents *always* chose to support affirmative action, then in the above example the likelihood that the government would make the right decision (from the majority group’s perspective) would fall from 90 percent to 50 percent.⁶³ However, this adverse effect on majority welfare might be partially or fully offset by improved electoral sorting. After all, if fair-minded agents always endorse affirmative action (regardless of its merits), while bigots always oppose it, then voters can use an agent’s position on this issue as an indicator of the agent’s type. If distinguishing good types from bad types is more important than getting the right answer on this particular issue, then “political correctness” may be desirable on net, as the issue in question becomes a kind of “litmus test,” in a positive rather than pejorative sense.⁶⁴

Of course, if reputational incentives are strong enough, then even bigoted agents may join fair-minded agents in supporting affirmative action.⁶⁵ Whether this more extreme political correctness effect leads to better or worse decisions depends on whether the gains from inducing bad types to support the policy (which they would otherwise oppose even when it’s a good idea) are larger or smaller than the losses from inducing good types to support the policy even when they think it’s a bad idea.⁶⁶ The effect of this extreme form of political correctness on selection is clearer: with this sort of convergence, voters can no

63 This is an artifact of the arbitrary assumption that the probability the affirmative action policy is in fact a good idea is 50 percent. If this probability were higher, then the adverse impact of this sort of political correctness would be less pronounced, though if the probability that affirmative action is desirable were lower, the adverse impact would be more pronounced.

64 See, e.g., Dorf (2006); Scherer (2005). Cf. Abrams, Fiorina, & Pope (2008) at 558; Hinich & Pollard (1981) at 328.

65 Such an equilibrium would be sustained by the principal’s belief that any agent who defied the consensus in favor of the pro-minority policy at issue must be a bad type. This may seem unrealistic, but consider the overwhelming congressional support for re-authorization of the Voting Rights Act, which numerous (southern conservative) legislators almost certainly opposed but felt politically obligated to support. See Carter, Jr. (2012) at 1142, n.105.

66 The numbers we have chosen imply that the latter costs are larger, but one could easily construct an example in which the opposite is true.

longer use an agent's position on affirmative action to make inferences about the agent's type, making selection less efficient.

As noted above, this analysis closely parallels the previous discussion of populism. Indeed, one can reframe a "populism" problem as a "political correctness" problem by re-identifying the relevant groups. In both cases, the same underlying dynamic is at work: reputation-minded agents have an incentive to avoid the appearance of undesirable bias, and so have an incentive to choose those policies that unbiased agents ("good types") are more likely to prefer *ex ante*, whether or not those policies are the ones that unbiased agents actually prefer *ex post*. This distorts agent behavior in ways that can be harmful to the principal: agents may sometimes take actions that fully-informed principals would disapprove, because such actions are more effective ways for the agent to signal her lack of bias. Sometimes the agent's incentive is to avoid the appearance of bias *in favor* of some unpopular group or cause (as in the populism case), while in other settings the agent is more concerned about avoiding the appearance of bias *against* some sympathetic group or cause (as in the political correctness case), but the underlying mechanism is essentially the same.

2. COMPENSATING FOR OVER-ACCOUNTABILITY PROBLEMS

Section 1 provided a conceptual overview of a family of over-accountability problems that all derive from a common root: the action that would optimize an agent's reputation with her principal may differ from the action that would actually be in the principal's best interests. This phenomenon can induce a range of undesirable behaviors by elected officials, bureaucrats, and other agents. When accountability mechanisms induce these sorts of behaviors, what should be done?⁶⁷

Sometimes, the answer is "nothing." The benefits of greater accountability—better selection of agents *ex ante* and better incentives for agents *ex post*—may outweigh the costs, just as the benefits of a drug therapy may outweigh the

67 Although we discuss solutions that a hypothetical institutional designer or policy reformer might adopt, we acknowledge that the very same problems identified in Section 1 of the article may also plague the task of implementing appropriate institutional reforms. Presumably citizens seeking to reform institutions to ameliorate principal-agent problems in government (including over-accountability) will usually need to rely on agents to develop and implement those reforms, but there can be principal-agent problems in that relationship as well. We acknowledge the problem, which is endemic to virtually all proposals for institutional reform. See Posner & Vermeule (2013). But in this part of the discussion we shift to an external perspective for purposes of analytic clarity, and because better understanding the nature of the problem may in fact help solve it.

side effects.⁶⁸ And sometimes the measures available to suppress over-accountability problems—to control the side effects, as it were—are costly, infeasible, or counterproductive. In such cases, over-accountability problems must be accepted as a fact of life, though awareness of their existence may be useful in evaluating and reforming institutions. In other cases, however, some form of institutional intervention may be appropriate.

One way to address the problem might be to adopt institutional reforms that ameliorate the underlying principal–agent problem, for example by reducing the information asymmetry between principals and agents. While measures to reduce the information asymmetry are likely desirable, in most cases it is impossible to completely eliminate the principal–agent problem. For that reason, we focus on institutional responses to over-accountability that treat the underlying principal–agent problem as given. Proceeding under that assumption, potential mechanisms for “fixing” or compensating for over-accountability problems fall into two broad categories. First, if accountability creates perverse incentives, then an obvious solution would be to curtail or eliminate accountability. Second, and more interestingly, additional oversight agents that monitor or constrain the primary decision maker may mitigate the over-accountability problem without sacrificing the first-order benefits of accountability—though poorly designed oversight may actually make the problems worse.

2.1 Reducing Accountability

If too much accountability is a problem, then less accountability might be the (partial) solution. At the risk of belaboring this obvious point, a number of institutional devices can be used, and indeed often are used, to mitigate over-accountability problems.

First, even when an accountable agent retains authority over some general policy area, certain decisions can be ruled out in advance (assuming the polity can credibly pre-commit to such limitations).⁶⁹ Alternatively, authority can be assigned (or reassigned) to a less-accountable agent, like an independent

68 For example, in addition to the selection and incentive effects of democratic accountability, such accountability may also have information effects that reduce the information asymmetry between principals and agents. Such information effects may arise if rival politicians or others have a stronger incentive to provide information, and citizens have a stronger incentive to become better informed about policy issues, when decision makers are elected.

69 See Fu & Li (2014).

agency or a court.⁷⁰ This practice, though frequently criticized as undemocratic,⁷¹ is often defended—in terms broadly consistent with our analysis—as a way to insulate important policy decisions from the distorting influence of politics.⁷² Our discussion complements and bolsters this view by demonstrating how insulation may generate behavior and policy that is better for all involved.⁷³

Another possibility is to leave the decision in the hands of an accountable agent but to weaken the relevant accountability mechanisms. Perhaps sometimes a modicum of accountability is desirable, but too much is harmful. Devices like term limits are one possibility. If elections are more likely to produce faithful agents than other means of selection, but the incentives created by the threat of electoral sanctions are perverse rather than productive, then the optimal system may involve electing officials to term-limited offices so that career concerns cannot distort agent behavior.⁷⁴ Other ways to retain accountability mechanisms, but to weaken them so as to limit over-accountability distortions, might include longer terms of office, greater “bundling” of responsibilities (so that each individual issue has less of an impact on overall reelection probabilities),⁷⁵ indirect elections, randomly timed elections, and so forth. All these mechanisms involve complexities and trade-offs; the only point we wish to make here is that the attractiveness of institutions that limit political accountability increases with the severity of over-accountability problems.

Another fix would be for principals to design a more refined “contract” with their agents, with variable rewards and punishments for different observable

70 To be clear, our argument is an argument about *relative* accountability: Even if judges and independent agency heads are not *perfectly* insulated, they are nonetheless *more* insulated than are other bodies.

71 See, e.g., Calabresi & Prakash (1994) at 638–639; Lessig & Sunstein (1994) at 101.

72 See Lewis (2008); Patashnik (2000); Huber & Shipan (2002); Falaschetti & Miller (2001); Waller (2000). For a classic statement, see Landis (1938).

73 Nothing in our analysis necessarily indicates that delegating to independent agencies or courts is welfare improving on net. Nonetheless, the possibility discussed in the text is significant given the tendency of some commentators to invoke “political accountability” (or, equivalently, “democracy”) as a kind of talisman. If the objective is faithful agency, and political accountability mechanisms may increase or decrease faithful agency, then pointing out that independent agencies or courts are less accountable, without more, is a vacuous truism.

74 See generally Besley & Case (2003); Garrett (1996); Cohen & Spitzer (1992). In addition to attenuating incentive effects, term limits may also make selection less efficient, because voters have less information about a challenger or an incumbent with a short history than they do about an incumbent with a long tenure in office.

75 See Ashworth & Bueno de Mesquita (2014); Berry & Gersen (2007); Gersen (2010); Besley & Coate (2003); Besley & Coate (2008).

actions or outcomes.⁷⁶ To unpack this a bit, the over-accountability problem arises because the agent's benefits from different actions are determined by two things: the agent's own evaluation of the merits of those actions, and how those actions affect the principals' assessment of the agent's type. The principals might do better if they could raise or lower the costs to the agent of certain actions, independently of what those actions would otherwise tend to imply about the agent's type, as this could discourage less-informed or biased agents from trying to mimic the expected behavior of the good types.

For example, consider our earlier posturing example, in which legislators would prefer the expert bureaucrat always to do what she thinks is right, but would prefer that normal bureaucrats always to take the "routine" action rather than the "extraordinary" action. If agents care enough about appearing to be expert, the normal type may sometimes "posture" by choosing extraordinary action.⁷⁷ But the normal type's interest in extraordinary action is likely to be weaker than that of an expert inclined to choose such action: after all, while both types would receive an equal reputational boost from that choice, the expert *also* prefers this action on the merits. Thus if legislators could somehow increase the costs of the extraordinary action—say, by adding procedural hurdles or other decision costs,⁷⁸ or by committing to reduce the political rewards associated with this action⁷⁹—they could get each agent to behave appropriately. This approach, however, might not be feasible, both because it requires pre-commitment, and because these sorts of refinements are much more difficult to calibrate and implement in the context of political agency than in the context of a more traditional contractual relationship, such as that between an employer and employee.

Finally, one can also weaken accountability mechanisms, and thus attenuate over-accountability problems, by reducing the transparency of the agent's decisions (what we call "decision transparency"⁸⁰). There are many ways to reduce decision transparency, ranging from limiting access to information

76 The term "contract" is a metaphor in this context, but the contract theory literature nonetheless provides useful insights. See Bolton & Dewatripont (2005).

77 See *supra* Section 1.1.2.

78 Cf. Stephenson (2007a); Stephenson (2008a).

79 See Nzelibe & Stephenson (2010b); Downs & Rocke (1994).

80 We distinguish this "decision transparency" from transparency as to the *outcomes* associated with those decisions, which we refer to as "outcome transparency." See Section 2.2.1, *infra*. See also Prat (2005) (distinguishing between transparency of actions and transparency of consequences); Fox (2007) (also noting this distinction).

about government action,⁸¹ to making the policy process sufficiently complex that it is difficult for principals to assign responsibility for decisions to particular actors (thus creating a “clarity of responsibility” problem),⁸² to altering the laws and institutional structures that govern the media and other watchdog organizations that might investigate and report on government policy.⁸³ The case for limiting decision transparency is the flip-side of the more conventional case for enhancing such transparency: greater decision transparency enhances political accountability.⁸⁴ When the costs of accountability outweigh the benefits, limiting the principals’ ability to observe the agent’s action can be beneficial on net.

2.2 Third-Party Oversight

Another institutional mechanism that might be used to preserve the best aspects of accountability while constraining undesirable over-accountability involves *third-party oversight*. The idea is that the principal can make use of a second agent (an overseer) whose responsibility is to police the behavior of the first agent (the primary decision maker). Of course, the introduction of a second agent is no panacea, not least because it introduces a second agency problem between the principal and the overseer. But under some circumstances the use of multiple agents may mitigate the over-accountability problem, while preserving much of the benefit associated with accountability.

Conceptually, oversight bodies may differ along two primary dimensions: (1) whether the overseer is itself accountable; and (2) whether the overseer can block (or limit or modify) the primary agent’s decision. Some overseers wield this sort of veto power, while others can merely provide their evaluation (a “second opinion”⁸⁵), which can then be used by the principals to decide whether to retain or fire the primary agent. While overseers may vary continuously along each of these dimensions, as a simplification we treat each

81 See Samaha (2006); Stone (2012). Sometimes, of course, limiting information about what the government did may not be possible. If the underlying question, for instance, is whether to invade a country with military force, a decision “to invade” is hard to make opaque. Still, in many settings something less than full transparency is surely possible, and increasing transparency is likely to exacerbate rather than ameliorate most accountability perversions.

82 See Bueno de Mesquita & Landa (2014); Berry & Gersen (2008); Hayward (2010); Gersen (2010). For example, when one agency has the authority to mandate that another agency enforce a rule, or one agency may not take an action without another’s consent, ambiguity about accountability is introduced into the system.

83 See Ashworth & Shotts (2010) at 844–845.

84 See generally Fenster (2006) (discussing relationship between transparency and accountability).

85 See Vermeule (2011).

dimension as a dichotomous variable. This allows us to sort overseers into four “ideal types,” summarized in Figure 2.

One category of third-party oversight (upper-left cell in Figure 2) involves an overseer that is politically insulated and lacks the power to veto the primary agent’s decision. The independent media, or other civil society “watchdog” groups, may fit into this category.⁸⁶ Likewise, in a “departmentalist” system of constitutional interpretation, in which judicial rulings are not binding on the political branches, judicial review might also fall into this category.⁸⁷ Other government institutions might fit this category as well. For example, the Government Accountability Office (GAO) and Congressional Budget Office (CBO) are relatively insulated entities which frequently provide analyses of legislative or executive initiatives, but cannot block those initiatives.⁸⁸

A second type of overseer (upper-right cell in Figure 2) has no formal power to block or alter the primary agent’s policy proposals, but is itself politically accountable. For example, a minority party in the legislature (the “loyal opposition”) may be unable to block the majority coalition’s initiatives, but may nonetheless have access to information and a platform from which to criticize (or to endorse) those initiatives.⁸⁹ Turning from inter-party monitoring to inter-branch monitoring, even when the legislature cannot block an executive branch action, members of the legislature can make speeches or hold hearings praising or criticizing executive decisions, and the legislature may also pass nonbinding resolutions of support or disapproval.⁹⁰ Likewise, even when the President is practically unable to veto particular provisions of a legislative proposal (as when these provisions are part of a larger bill that

86 It is no doubt true that media outlets and other civil society groups care about their reputations for competence and lack of bias (or at least “correct” bias from the relevant consumers’ perspective), in part due to market forces. See Ashworth & Shotts (2010) at 844–845; Gentzkow & Shapiro (2006). Yet because these overseers are not elected, and the opinions they provide are “bundled” together with other forms of valuable information (such as “hard news”), it is likely that they are *relatively* more insulated from reputational concerns than are elected officials. A media outlet or other third-party overseer that cared sufficiently about its reputation (but still lacked the formal power to block action) would fall into the upper-right cell in Figure 2.

87 This sort of departmentalism has advocates on both the political right, see, e.g., Paulsen (1994) at 225–226, and the political left, see generally, e.g., Tushnet (1999), though their arguments have not gained much traction in the U.S. federal system. Compare Lawson & Moore (1996), with Alexander & Schauer (1997).

88 As with the media and civil society watchdog groups, one must be careful not to overstate the political insulation of the GAO, CBO, etc.

89 See Schmitz (1988).

90 See Gersen & Posner (2008).

Figure 2. A typology of ideal-type overseers.

		Is the overseer politically accountable?	
		NO	YES
Can the overseer block the decision?	NO	<ul style="list-style-type: none"> • <i>Independent media or other civil society watchdogs</i> • <i>Constitutional judicial review in a “departmentalist” regime</i> • <i>Independent intra-governmental monitors</i> 	<ul style="list-style-type: none"> • <i>Public statements by the minority party (“the loyal opposition”)</i> • <i>Nonbinding resolutions of legislative approval or disapproval of executive action (“soft law”)</i>
	YES	<ul style="list-style-type: none"> • <i>Judicial review by insulated judges empowered to issue binding rulings</i> • <i>Regulatory review by a politically independent agency</i> 	<ul style="list-style-type: none"> • <i>Presidential veto of legislation</i> • <i>Legislative veto of executive proposals</i> • <i>Judicial review by elected judges</i>

the President feels compelled to sign), she may use signing statements or other forms of communication to express her views.⁹¹ In all these cases, the “overseers” are themselves accountable officials who care about how the second opinions they offer may affect their reputations.⁹²

Third, some overseers are both politically insulated and have the ability to veto the primary agent’s policy proposals (bottom-left cell in Figure 2). The most familiar example of this sort of oversight is independent judicial review in a system with judicial supremacy.⁹³ In addition to domestic judicial review, this category might also include international tribunals (existing or proposed) with the power to review legislation enacted by national governments.⁹⁴ One could also imagine similar bodies in the administrative system. For example, some areas of U.S. regulatory policy, such as worker safety, are characterized by a “vertical split enforcement” scheme in which the decisions of a primary enforcement agency (which, as part of a cabinet department, is relatively more accountable) are adjudicated by an independent commission (which is more insulated).⁹⁵ Furthermore, although centralized regulatory review as currently practiced in the USA does not really fit in this category (because the main reviewing entity, the Office of Information and Regulatory Affairs (OIRA),

91 See Bradley & Posner (2006); Garber & Wimmer (1987); Cross (1988).

92 As noted earlier, even unelected overseers may fall into this category if they care sufficiently about their reputations. For example, media outlets, watchdog groups, and other nominally “independent” entities might care so much about their reputations that they would fit better into this category than into the first category.

93 We must add the appropriate caveats here. In our typology, overseers in this category are not politically accountable. That strong assumption is not entirely accurate when applied to actual federal judges, who may care about their reputations both because they aspire to higher judicial office, and because they derive benefits from the esteem of certain constituencies. Furthermore, though judicial supremacy is fairly well-established, the power of the courts to obstruct the political branches cannot be taken as given. Yet as we have emphasized, our ideal-type categorization is meant as a heuristic device that uses simplified *absolute* categories to capture important *relative* differences across oversight institutions. Notwithstanding the above qualifications, it is almost certainly true that federal judges are significantly less concerned with the short-term reputational consequences of their decisions than are elected politicians, and that judicial decisions have greater power to block government initiatives than do op-eds or nonbinding resolutions of disapproval.

94 See generally Shelton (2003) (surveying the increasingly formal mechanisms that secure the independence of international tribunals).

95 See Gersen (2007) (discussing split enforcement regimes). For example, the Occupational Safety and Health Act gives rulemaking authority for workplace safety standards to the Occupational Safety and Health Administration in the Department of Labor, while granting an independent commission, the Occupational Safety and Health Review Commission, adjudicatory authority for violations.

is closely affiliated with the President⁹⁶), it is possible to imagine an alternative scheme where many of OIRA's review functions are performed by a more politically insulated body.

The final oversight category in our typology (bottom-right cell in Figure 2) covers accountable overseers with veto power. Most of the standard checks-and-balances between and within the political branches would fit into this category. The power of the President to veto legislation, for example, involves an accountable overseer (the President) blocking the proposals of another accountable decision maker (the legislature).⁹⁷ The same would be true of institutions that require legislative approval of executive branch proposals, such as a legislative veto,⁹⁸ or more generally settings in which the President needs to get congressional authorization.⁹⁹ Though the roles are reversed, these settings also involve an accountable overseer (Congress) empowered to block the proposals of an accountable primary decision maker (the President or her subordinates). Intra-branch checks—such as legislative bicameralism,¹⁰⁰ OIRA review of agency decisions,¹⁰¹ and mandatory

96 Miller III (2011) at 101; Bressman & Vandenberg (2006) at 49 (describing OIRA as the President's "regulatory review arm"); Croley (2003) at 830 (equating OIRA and White House review); Kagan (2001) at 2318 (discussing President George W. Bush's pick for OIRA as a form of presidential "control of administrative action"). Of course, there remains substantial daylight between OIRA and the President. See Bagley & Revesz (2006) (noting that "OIRA is not the President," *id.* at 1307, and that "[t]he White House has limited resources to expend on regulatory oversight — including OIRA oversight — and only the most salient or politically consequential regulations will invite explicit White House attention," *id.* at 1308); Barron (2008) at 1113 ("OIRA review is not simply a means of ensuring White House control.").

97 U.S. Const., art. I § 7 cl. 2.

98 The legislative veto was ruled unconstitutional in *Immigration & Naturalization Service v. Chadha*, 462 U.S. 919 (1953). However, Congress has continued to enact—and the executive branch has continued to respect—legislative veto provisions, despite their formal unenforceability. See Fisher (2000) at 180–182; Fisher (1993). From time to time various proposals have been floated that would create the equivalent to a legislative veto, without the formal constitutional problem identified in *Chadha*. See, e.g., Breyer (1984) at 793–796; Sohoni (2012) at 1593.

99 The canonical account of the relationship between presidential power and congressional authorization is Justice Jackson's tripartite framework in *Youngstown Sheet & Tube Co. v. Sawyer*, 343 U.S. 579, 637 (1952) (Jackson, J., concurring); see also *Medellín v. Texas*, 554 U.S. 749 (2008). Examples of presidential reliance on or constraint by Congress abound, especially in the national security context—e.g., congressional authorizations for use of military force, see generally Bradley & Goldsmith (2005), and the War Powers Resolution, 50 U.S.C. §§1541–1548 (2006), see generally Grimmett (2010).

100 Levmore (1992).

101 See *supra* note 99.

inter-agency consultation¹⁰²—are trickier cases, but may also fit into this category, at least sometimes.¹⁰³ This category would also cover judicial review in jurisdictions where judges must stand for re-election, as is true in many U.S. states.¹⁰⁴

Equipped with this typology, we can consider how each of these four forms of oversight might ameliorate—or exacerbate—the over-accountability problems described in Section 1.

2.2.1. *Unaccountable overseers without veto power*

The simplest place to begin is with the upper-left cell in Figure 2: politically insulated overseers without veto power. These overseers' only role is to offer a second opinion, and their political insulation implies that this second opinion will be an honest one (that is, the overseer will say what she thinks, without fear or favor).¹⁰⁵

Most straightforwardly, this overseer's opinion can provide principals with more accurate (though still imperfect) information about the true state of the world, and this can give the primary agent a stronger incentive to do what she thinks is in the principals' best interest.¹⁰⁶ So long as the overseer's evaluation is better than random, and her bias is not too extreme, then if the overseer agrees with the primary agent, this endorsement has a "legitimation effect," improving the primary agent's reputation.¹⁰⁷ Likewise, overseer opposition to the primary agent's policy has a delegitimizing effect, hurting the latter's reputation. Therefore, the primary agent has a stronger incentive to do what she believes is right, rather than what would (in the absence of oversight) send the most

102 Sometimes a statute requires an action by one agency prior to an action by another agency. For example, section 7 of the Endangered Species Act requires the Fish & Wildlife Service to make a "no jeopardy" finding before the "action" agency can move forward. The Secretary of Energy can propose rules, regulations, and statements of policy in areas that fall under the jurisdiction of the Federal Energy Regulatory Commission. DeShazo & Freeman (2005); Gersen (2007); Marisam (2011).

103 These cases are harder to classify in part because there is more negotiation involved than in a simple proposal-veto structure.

104 See Huber & Gordon (2004); Gordon & Huber (2007).

105 Technically, *any* announcement by the overseer would be sequentially rational. We assume here that the overseer has a mild preference for honestly stating her views.

106 This discussion is based principally on Ashworth & Shotts (2010).

107 See Fox & Stephenson (2011).

favorable signal about her type.¹⁰⁸ This is true even if the overseer's information is relatively inaccurate, though the beneficial effect of the overseer's signal declines as her accuracy worsens.¹⁰⁹

The same logic applies to any other measure that improves principals' ability to assess the quality of the primary agent's policy decision, including measures that increase the transparency of policy *outcomes* (as distinct from measures that increase the transparency of policy *decisions*). One can conceive of the overseer's "subjective" endorsement or criticism as a noisy signal of the quality of the policy chosen, but a similar effect might be achieved with "objective" reports that provide better information about the actual consequences of a given policy—such as fact-based reporting on how a war is going or about the state of the economy. Such information is inherently imperfect, yet increasing the accuracy of this sort of information—outcome transparency as opposed to decision transparency—may reduce the primary decision making agents' incentive to behave incorrectly, because citizens will have a better sense whether the agents actually acted in the citizens' interests.¹¹⁰ In sum, this kind of independent oversight may substantially mitigate over-accountability problems,

108 Although we focus on settings in which the overseer offers its opinion on the desirability of the chosen *policy*, a similar analysis might apply if the overseer provided a more general (and at least somewhat informative) assessment of the primary agent's type. The reason is that if the agent knows that her policy choice has less of an effect on her reputation (because the principals can use the overseer's announcement as a separate source of information), then her policy decisions will be driven relatively more by policy considerations and less by reputational considerations.

109 The fact that the overseer is not accountable may reduce the accuracy of her signal, at least insofar as principals would try to select oversight agents who were highly competent and dismiss those who were consistently wrong. The fact that the overseer's announcement does not have a direct policy effect may also reduce the quality of her signal, relative to overseers with veto power, because the overseer's lack of power may reduce her incentives to invest costly effort in acquiring and analyzing relevant information. See Stephenson (2011).

110 See Fearon (1999); Prat (2005). An important qualification to this observation concerns short-term and long-term welfare effects, or more generally policy outcomes comprised of several separate outcomes, some of which may be more transparent than others. Increasing the transparency of *some* aspects of the policy outcome (such as short-term effects), while leaving other aspects (such as long-term effects) opaque, may exacerbate rather than ameliorate over-accountability, in much the same way that increasing decision transparency may exacerbate these problems. The reason is that voters may rationally use the observable outcome as a proxy for the outcome more generally, which creates incentives for agents to take the action that generates the best results on the observable dimension, even if the agent knows that some other action—which generates less-good observable results—would be better for the principals overall. The seminal paper on agency problems in multi-task settings is Holmstrom & Milgrom (1992). For applications to government institutions, see, e.g., Bueno de Mesquita (2007); Bueno de Mesquita & Stephenson (2007).

simply by providing voters with more information about whether the primary agent did the right thing.¹¹¹

However, there are some potential complications that might undermine, or even reverse, the beneficial informational effects of this sort of third-party oversight. Perhaps most importantly, although outcome transparency and decision transparency are conceptually distinct, they may tend to move together, and institutions that strengthen the former may strengthen the latter. In other words, it may be difficult in practice to disentangle revelation of information about the government's policy decisions from a substantive evaluation of those decisions; indeed, often the same entities perform both functions simultaneously. The media, for example, provides both "hard news" about government decisions (thus improving decision transparency) and editorial commentary or other kinds of hard news that facilitate more accurate assessment of government performance (thus improving outcome transparency). These functions are not easily separable, and indeed the credibility of the evaluative statements may depend on the accuracy and clarity of the more objective descriptions of what the government has in fact done.

This is a problem because, as noted in the preceding section, when over-accountability distortions exist, greater decision transparency is not necessarily a good thing: the more accurately principals can observe policy choices (as opposed to outcomes), the stronger the incentive for agents to make those choices with regard to their short-term reputational effects. So, when an institutional designer is considering reforms that would create or eliminate a particular government oversight body (like the GAO or CBO), or that would strengthen or weaken outside overseers (say, through laws or regulations that make it easier or harder for the press or civil society organizations to monitor government), the threat of over-accountability problems may give rise to a trade-off between limiting the transparency of the particular decisions made and enhancing informative evaluations of the quality of those decisions.¹¹² The former effect tends to make over-accountability

111 This is one reason that the traditional bar on Article III courts issuing advisory opinions can be detrimental. Even if a judicial opinion were not enforceable, the advisory opinion would nevertheless provide information about the desirability of a decision by Congress or the President. See Fox & Stephenson (forthcoming 2014a).

112 This is essentially a trade-off between reducing decision transparency and increasing outcome transparency, both of which may be desirable if accountability perversions are severe. (Honest and independent second opinions from third-party overseers are, in effect, a form of greater outcome transparency insofar as the overseer's judgment is correlated with the truth.) The institutional design difficulty highlighted in the text is that decision transparency and outcome transparency may be positively correlated—the measures that increase one tend to increase the other as well—and this means the net effects of increasing "transparency" (or, equivalently, strengthening third-party oversight) are ambiguous.

problems worse, while the latter effect tends, in most cases, to redress these distortions.¹¹³

2.2.2. *Unaccountable overseers with veto power*

Over-accountability problems may sometimes be sufficiently pronounced that it makes sense to empower an insulated oversight agent to block the primary agent's decision.¹¹⁴ Independent judicial review may be the most obvious example, though there are others as well. The most important advantage of giving the overseer veto power is straightforward: If over-accountability problems are sufficiently severe, and if the overseer is a good enough agent (reasonably competent, not too biased), then giving the overseer veto power can prevent bad policies from taking effect. (And, to the extent that the overseer's decision remains a credible source of information about the quality of the policy choice—an important qualification we will explore in a moment—this sort of oversight also may improve outcome transparency.) A second advantage is that giving the overseer veto power may actually improve its performance (that is, improve the accuracy of its signal) because the overseer knows that its decision will have immediate consequential effects on policy, and so has a stronger incentive to invest effort in getting the right answer.¹¹⁵

Such a scheme has at least two disadvantages, however. First, the overseer may block some desirable policies. Second, because the overseer is not itself accountable, the overseer is likely, on average, to be a worse type than the primary agent. That said, one need not assume that the overseer is a *better*

113 A second, more subtle problem has attracted a great deal of attention in the theoretical political economy literature, though its practical significance is unclear. In certain circumstances, there may be “threshold effects,” such that although most changes in an incumbent's reputation have relatively modest effects on her re-election chances, there are thresholds at which relatively small changes in the incumbent's reputation can translate into very large increases or decreases in her re-election probability. In these cases, the introduction of independent and honest third-party oversight (or, equivalently, an improvement in outcome transparency) can create a new over-accountability problem. Consider an incumbent who, in the absence of third-party oversight, is certain to win (or lose) her upcoming re-election, no matter what she does. Such an incumbent, though nominally accountable, is *de facto* independent. The introduction of the third party's evaluation may be enough to push the incumbent below (or above) the threshold at which the election becomes seriously contested. If so, then the introduction of oversight may induce accountability—and perhaps over-accountability—that would not have otherwise existed. See Canes-Wrone et al. (2001); Ashworth & Shotts (2010). The practical significance of this theoretical possibility, however, is unclear, given that the result arises only if there are strong threshold effects in the relationship between an agent's reputation and her probability of retention, and only if the relevant thresholds are in exactly the right place.

114 For simplicity, we describe this as “veto power,” though that characterization is inaccurate in some applications. For example, U.S. federal courts cannot “veto” a statute as unconstitutional before it goes into effect. Rather, the court must await a post-enactment challenge to the statute in a concrete case.

115 See Stephenson (2011).

agent than the primary decision maker (in terms of bias or competence or both) in order for the principals to benefit from giving the overseer veto power. Under the appropriate circumstances, giving even a flawed overseer veto power may lead to better decisions overall. But this is not guaranteed, and as a first cut the decision whether to introduce an independent overseer with veto power may depend on a comparison of the errors that the overseer would correct to the errors that it would introduce.

There are some additional subtleties not fully captured in this first-cut comparison of relative error costs, however. First, the overseer's decision to uphold or block the primary decision maker's proposal will have a legitimizing or delegitimizing effect similar to the sort discussed in the previous section. This legitimation effect may mitigate over-accountability problems for all the reasons given in the preceding section.¹¹⁶ However, in contrast to the case where the overseer can only issue a nonbinding opinion, the legitimation effect is likely to be more attenuated when the overseer has veto power. After all, an independent overseer who lacks veto power has no incentive to dissemble with respect to her assessment of the primary agent's decision: such an overseer can "call them as she sees them," partly because she is not worried about her own reputation, but also because there are no significant policy costs to being wrong. This sort of overseer—like many an academic—is liberated by her own (ex post) irrelevance: she has no incentive to "defer" to the judgment of the primary agent, even if the overseer thinks that the primary agent is actually a better judge of policy. Giving the overseer veto power changes her calculations, because she must now be worried about making the wrong call. This can make her reluctant to veto the primary agent's policy choice, even when the overseer's own analysis indicates that the policy is a bad idea.

Another stylized example can illustrate this. Suppose the primary agent is an elected official, such as a President, who must decide whether to nationalize an industry that provides vital wartime materials.¹¹⁷ If the President nationalizes the industry, there is an 80 percent chance that this is in fact the right decision from the perspective of the voters (the principals). Now suppose we add an overseer who can either approve or disapprove of the nationalization, and assume that the overseer's independent assessment is correct 60 percent of the time. Compare two types of overseer: one (say, a newspaper) can publish an editorial approving or condemning the President's decision to nationalize the industry; the other (say, a court) can actually invalidate the decision. Both

116 See Fox & Stephenson (2011). However, if strong threshold effects are present the perverse result discussed previously may also obtain See *supra* note 117.

117 The example is drawn from *Youngstown Sheet & Tube Co. v. Sawyer*, 343 U.S. 579 (1952).

would approve of the nationalization if the overseer's independent assessment indicated that nationalization was the right thing to do. But what if the overseer's assessment indicates the decision was wrong? The newspaper would condemn it, because the newspaper can't change the policy and so it might as well simply announce its true views.¹¹⁸ The court is in a quite different position. If the President believes nationalization is the right decision but the court's independent evaluation indicates that it's the wrong decision, then a rational court would conclude that the President's decision was probably correct.¹¹⁹ Thus a rational court would "defer" to the President by upholding the nationalization even when the court's independent analysis suggests the decision was wrong. The voters benefit from the court's deference, in that they would not want the court to strike down a decision that is probably correct. However, the voters lose the benefit of the legitimation effect, because the court's approval no longer conveys useful information, and this also may have adverse effects on the President's incentives to choose nationalization even when this may not be the best decision for the voters.¹²⁰

So, one important difference between settings in which an unaccountable overseer has veto power and one in which such an overseer lacks veto power is that veto power may attenuate the legitimation effect, because an overseer with veto power may be more cautious and deferential. In some circumstances this attenuation may be a benefit,¹²¹ but probably more often it must be counted as a cost. Although veto power enables an overseer to combat the adverse consequences of over-accountability distortions more directly, it may also make such distortions more likely or more severe, relative to a setting with a comparable overseer who lacks veto power. For this reason, giving the overseer veto power may also make the selection effect a less effective means for improving the quality of the primary agent.

118 The fact that the primary agent (in this example, the President) can anticipate the overseer's action may have a significant influence on her initial decision, because of the legitimation effect elucidated in Section 2.2.1. This may mitigate (or in certain special cases exacerbate) over-accountability. For purposes of this example, though, we fix the probability that the primary agent makes the correct decision at 80 percent.

119 By Bayes' Rule, the probability that the President's proposal is correct, given that the court's analysis indicates otherwise, is $(0.4)(0.8)/[(0.4)(0.8)+(0.6)(0.2)] = 8/11$.

120 This claim assumes voters cannot distinguish situations in which the court's own analysis accords with the President's from situations in which the court upholds the President's decision out of deference rather than agreement. This assumption is often reasonable, yet courts do sometimes declare they are upholding a government action out of deference rather than agreement. The analysis here suggests that the attempt to distinguish deference from agreement is important, as such a distinction could preserve more of the informational value of the court's assessment without the court taking action that is likely wrong.

121 See *supra* note 117 (discussing threshold effects).

Giving the overseer veto power can also exacerbate over-accountability problems in a second way: Such oversight creates a kind of “safety net” for the primary decision maker, reducing the probability she will have to live with the consequences of making a decision she views as unwise but politically advantageous. In the absence of an overseer with veto power, the incumbent decision maker may have to trade off policy benefits against political benefits, and may sometimes be unwilling to choose the more politically advantageous action because the expected policy costs are too high. But if the incumbent knows that an overseer may bail her out by vetoing the unwise but politically advantageous policy, she may adopt such a policy more often. This “bailout effect,” which has been discussed in the context of judicial review,¹²² is another potential cost of endowing an overseer with the power to block, instead of merely to announce disapproval of, the primary agent’s decision. Sometimes this cost may be relatively mild, but in other circumstances it may exacerbate over-accountability distortions to the point that the introduction of an overseer with veto power does more harm than good.¹²³

Additionally, the bailout effect and the incentive to defer—both consequences of endowing the overseer with veto power—can create self-reinforcing dynamics in which aggressive oversight exacerbates over-accountability problems (via the bailout effect), and the extent of these over-accountability distortions in turn makes aggressive oversight (rather than deference) rational for the overseer. By the same token, more deferential oversight may reduce bad behavior of this sort (because there is little or no bailout effect), and the absence of significant over-accountability distortions may in turn make deference a rational strategy for the overseer. This possibility—which gives rise to what game theorists call “multiple equilibria”—can make it difficult to predict the impact of introducing oversight, because exactly the same institutions (at least with respect to their “fundamentals”) can lead to quite different outcomes, depending on which equilibrium is selected.¹²⁴

122 See Vermeule (2006); Tushnet (1999); Thayer (1893); Fox & Stephenson (2011); Rogers (2009).

123 See Fox & Stephenson (2011). The two effects just discussed are in some tension. If the overseer is extremely deferential to the primary agent, then not only is the legitimation effect attenuated, but there is not likely to be much of a bailout effect. However, there is no necessary contradiction: One or the other effect might be present, depending on the circumstances. Moreover, both effects might arise in the same setting if the overseer’s confidence in her own judgment varies from case to case. See *id.*

124 See Fox & Stephenson (2011). Self-limiting dynamics are also possible. Suppose that extreme deference, by eliminating the legitimation effect, would trigger more severe over-accountability distortions. The severity of these distortions might make extreme deference irrational for the overseer. But, if the overseer relies on its own independent assessment, the associated legitimation effect may sufficiently mitigate over-accountability distortions to the point where deference becomes rational

In sum, the impact of a politically-insulated overseer on over-accountability problems may depend on whether the overseer has veto power or not. Veto power has a direct effect, in that it empowers the overseer to strike down policies that she is sufficiently confident are bad ideas. This can correct some of the adverse consequences of over-accountability problems. This direct effect, however, entails the risk of significant error costs (vetoing policies that are good ideas), particularly since an unaccountable overseer is likely to be a flawed agent. Beyond this direct effect, endowing an insulated overseer with veto power has more subtle indirect effects. First, an overseer with veto power may have stronger incentives to invest in improving the accuracy of its signal, which strengthens the case for conferring such power. Second, however, an overseer with veto power has stronger incentives to defer to the primary agent, and this may attenuate the informational value of the overseer's decision (and this incentive to defer may counteract the incentive to invest in information). Third, the overseer's veto may create a kind of safety net, reducing the policy costs to the primary decision maker of unwise but politically advantageous decisions. These latter two indirect effects may exacerbate over-accountability distortions, and their interaction may also generate additional complications that make the effect of oversight on behavior hard to predict.

2.2.3. *Accountable overseers without veto power*

The preceding two subsections considered how one might mitigate the over-accountability problems that afflict the primary decision maker by employing overseers that are *not* themselves accountable to the principals. What about overseers that *are* accountable? If the main concern is that the primary decision maker may fall victim to one or more of the over-accountability problems explored in Section 1, should the overseer be insulated or accountable? Consider first an overseer that lacks veto power. How do the costs and benefits of oversight change if the overseer cares about its reputation in much the same way that the primary decision maker does?

Many of the advantages and disadvantages of making the overseer accountable parallel the advantages and disadvantages of making the primary decision-maker accountable. The most obvious advantage is that an accountable overseer

for the overseer. Such settings may therefore lack a “pure strategy” equilibrium in which each agent behaves consistently and predictably. Instead, the equilibrium may involve some degree of inherent unpredictability: the overseer sometimes defers and sometimes follows her own independent judgment (essentially at random, or at least not following any pattern that the primary agent can anticipate), and this unpredictable “standard of review” induces an intermediate degree of over-accountability distortion that makes either deference or non-deference rational approaches for the overseer. *See id.*

may well be a better agent—less biased, more competent—than an insulated overseer, because of the joint operation of the selection effect and the incentive effect (when the latter is operating productively rather than perversely). If the only difference between a second opinion from an accountable overseer and one from an unaccountable overseer were that the former is more likely to be accurate, then the choice would be clear. However, as Section 1 established, sometimes the incentive effect operates perversely, leading an agent to take the action that principals expect good types to choose more often, regardless of what the agent actually believes. All the undesirable behaviors described in Section 1 may afflict an overseer, just as they may afflict the primary agent.

This problem is important, but we will not belabor it. Rather, we focus on a separate (though related) set of concerns that arise particularly in a context where one accountable agent gives an opinion on the choice of another accountable agent. In particular, the overseer's reputation is affected not only by her decision, but also by whether she *agreed or disagreed* with the primary agent. This is the flip-side of the legitimation/delegitimation effect discussed earlier. If the overseer agrees with the primary agent, this not only improves the reputation of the primary agent (the legitimation effect), but it also may improve the reputation of the overseer; likewise, disagreement may hurt *both* of their reputations. Several things follow from this.

First, the reputational consequences of agreement and disagreement may induce accountable overseers to be more deferential. The end result is similar to the deference phenomenon discussed in the previous subsection with respect to insulated overseers with veto power,¹²⁵ but it arises for a different reason. In the earlier case, the overseer deferred not out of concern for her own reputation, but out of respect for the primary agent's superior judgment. Here, the overseer's concern for her own reputation is the reason for deference.

To illustrate with a stark example, consider a setting in which both the primary agent and the overseer each have a 70 percent chance of being good types,¹²⁶ the good type always knows the right decision, and the bad type's guess about the right thing to do is correct 60 percent of the time. Suppose the primary agent must choose regulation or non-regulation, with each choice having an equal *ex ante* likelihood of being correct. Suppose the primary decision maker chose regulation, but the overseer's independent analysis indicates non-regulation is the correct choice. If the overseer condemns the primary

125 See *supra* TAN 123-125.

126 For simplicity we also assume these probabilities are independent.

agent's choice, then the citizens know that at least one of the two agents must be a bad type, so each agent's perceived probability of being a good type drops from its prior level (70 percent) to approximately 40 percent.¹²⁷ By contrast, agreement increases each agent's perceived probability of being a good type to approximately 78 percent.¹²⁸ This means there is a reputational incentive to agree, which may induce deference on the part of the overseer (who, in this setting, moves second). Thus even if an accountable overseer is more competent than an unaccountable overseer, she may be less likely to honestly announce any disagreement for fear of the adverse reputational consequences. This problem—a kind of “herd behavior”¹²⁹—may be particularly pronounced if the primary agent has more expertise than does the overseer, as will often be the case. If principals perceive the primary decision maker as relatively more likely to be a good type, the reputational consequences of disagreement, when it occurs, will fall more heavily on the overseer.

This same basic problem may apply even to overseers that are not directly accountable, like media outlets or watchdog groups, if these overseers care about developing a long-run reputation, and if enough information will eventually come to light to enable citizens to alter their assessment of the overseer's reliability. Although one might think that an overseer's desire to acquire a reputation for making the right calls (that is, condemning policies that turned out to be failures, while endorsing those that turned out to be successes) would always strengthen the overseer's incentive to give its honest opinion, this is not always the case. Again, the reason is that the overseer usually announces her judgment only after the primary agent has made her choice,¹³⁰ and the fact that the primary decision maker concluded a given policy was a good idea will influence the overseer's evaluation. If the primary agent's judgment is more informative than the overseer's independent assessment, then an overseer interested in maximizing accuracy (perceived after the fact) may have a reputational incentive to act as a “yes man” by always endorsing the primary agent's decision.¹³¹ When this occurs, the overseer will not provide reliable

127 By Bayes' Rule, each agent's probability of being a good type given that the two of them disagree is $(0.3)(0.4)(0.7)/2[(0.3)(0.4)(0.7)+(0.3)(0.3)(0.6)(0.4)] = 35/88$.

128 By Bayes' Rule, each agent's probability of being a good type given that they agree is $[0.7 + (0.3)(0.6)](0.7)/[(0.7 + (0.3)(0.6)](0.7) + [(0.7)(0.6) + (0.3)(0.6)(0.6) + (0.3)(0.4)(0.4)](0.3)] = 385/493$.

129 See, e.g., Banerjee (1992); Bikhchandani, Hirshleifer & Welch (1992); Effinger & Polborn (2001). We could describe this sort of deference as a kind of “pandering” by the overseer.

130 This is not always the case, and we will consider below how issues of timing may affect the efficacy of this form of oversight.

131 See Prendergast (1993); Effinger & Polborn (2001).

additional information to the citizens, and therefore will not do much to counter over-accountability problems.¹³²

The discussion thus far has assumed that each agent cares about her *absolute* reputation—that is, the principal’s assessment of the probability that the agent herself is a good type. That makes sense in some contexts, but in other contexts agents may care more about the principal’s *relative* assessment of each agent. This might be the case, for example, when the principals are voters, the primary decision-making agent (say, the President) is controlled by one political party, while the overseer (say, the Congress) is controlled by a rival party. Even if disagreement hurts the reputation of both agents, it may hurt the reputation of one more than the other; likewise, even if agreement helps both agents in an absolute sense, it may help one more than the other. If disagreement hurts the overseer more than the primary decision maker, the former’s incentive to defer strengthens. However, if disagreement hurts the primary agent’s reputation more than the overseer’s reputation, then an overseer who cares about her relative standing has less incentive to defer. In that latter case, politically-accountable oversight may be more effective when the overseer is “partisan” in the sense that she cares more about her standing *relative to the opposition* than about her standing in an absolute sense.¹³³ That can be good for the principals in that it improves the informational value of oversight. However, it can go too far: If the overseer cares overwhelmingly about relative standing, and if disagreement hurts the primary agent more than the overseer, then the overseer may have a powerful incentive to publicly condemn the pri-

132 The equilibrium dynamics here can be quite complex. For example, suppose that, absent the overseer, (1) serious over-accountability distortions arise; (2) honest oversight would eliminate these distortions; (3) the overseer is more interested in accuracy than honesty; and (4) the primary agent’s information is so much better than the overseer’s that the overseer would always defer to the primary agent’s decisions if—but only if—the overseer is confident the primary agent is acting consistently with her true preferences and beliefs. In this case, if the primary agent always did what she thought was right, the overseer would always defer—but by assumption, if the overseer always defers, the primary agent suffers from over-accountability. If the primary agent’s behavior is distorted by over-accountability, the overseer would always honestly announce its own assessment—but, again by assumption, if the overseer always gives an honest assessment, the primary agent always does what she thinks is right (that is, the over-accountability distortion is eliminated). In this sort of setting, there may not be a stable equilibrium in which all agents’ behavior is consistent and predictable—rather, there may be a “mixed strategy” in which the primary agent sometimes takes inappropriate actions (frequently enough to make the overseer indifferent between honest criticism and deference), while the overseer sometimes gives her honest opinion rather than deferring (frequently enough to make the primary agent indifferent between doing what she thinks is right and distorting her behavior). This complication, though significant, does not undermine the main point in the text, which is that the overseer’s interest in accuracy can lead to excessive deference to the primary agent, which in turn can weaken the efficacy of this sort of oversight in combating over-accountability.

133 See Fox & Van Weelden (2010).

mary agent's decision even when she privately agrees that it was the right call.¹³⁴ That is, the overseer may have incentives to act as a “no man” rather than a “yes man.”¹³⁵

Of course, both these tendencies may have self-corrections built into them: The more the overseer acts as a yes man (to avoid the reputational consequences of disagreement), the harder it is for voters to draw inferences about either agent's type from the fact of agreement, and this weakens the incentive for the overseer to act as a yes man. Likewise, if the overseer derives a significant *relative* reputational benefit from disagreeing—that is, if disagreement makes the primary agent look bad—then the overseer may have a reputational incentive to act as a no man, but the very fact of this reputational incentive means disagreement does less damage to the primary agent's reputation, and this lessens the overseer's incentive for contrarian behavior. Whether these self-correcting tendencies will kick in is difficult to predict. It is possible for the overseer to always support or always oppose, giving the citizens no information and making oversight a costly but useless exercise. But it is also possible, in less extreme cases, for these self-corrections to constrain, though perhaps not eliminate, undesirable yes man or no man behavior.

In sum, taking an unaccountable overseer without veto power as a baseline, making the overseer more accountable (or, more generally, strengthening the overseer's concerns about her reputation) has a range of possible benefits and costs. On the benefit side, an accountable overseer may be more likely to be a good type, both because of a selection effect and because the incentive effect can sometimes operate in a productive manner. However, the same over-accountability distortions that afflict the primary decision maker may afflict the overseer as well. Moreover, because citizens may draw inferences about the overseer from the fact of agreement or disagreement, an accountable overseer may have excessive incentives to agree, or to disagree, with the primary agent's decision. The significance of this latter problem may in turn depend on whether the agents care more about their absolute or relative standing. A basic trade-off

134 See Baliga & Sjoström (2001); Glazer (2007).

135 No-man behavior can also occur if subsequent information may come to light after the overseer announces her position, and the overseer's only hope of improving her competitive standing vis-à-vis the primary agent is to correctly oppose one of the primary agent's initiatives. In this case, even if the overseer expects that her opposition to the primary agent's decision will prove unfounded (thus worsening the overseer's relative reputational standing), this reputational hit will not translate into a significant political cost; however, criticizing the primary agent's decision and having that opposition validated might significantly boost the overseer's standing. (This presumes “threshold effects” of the sort discussed earlier, see *supra* note 117.) In this case, the overseer has an incentive to oppose even policies she thinks are probably good ideas. This tendency reduces the overseer's informational value.

involved in making the overseer accountable, then, is that accountability may produce overseers capable of making more accurate assessments of the primary agent's policy choices, but more reluctant to announce those assessments honestly.

2.2.4. *Accountable overseers with veto power*

Finally, consider an accountable overseer with veto power. Although this may be the most complex of the four ideal types to analyze, our discussion can be brief because most of the relevant effects have already been discussed. The fact that the overseer is accountable suggests—though does not guarantee—that it is likely to be a better agent (more competent, less biased) than a comparable unaccountable agent. However, the fact that the overseer is accountable implies the possibility not only that the overseer may be vulnerable to the over-accountability distortions discussed in Section 1, but more particularly that the overseer will be excessively deferential to (or, in certain cases, excessively critical of) the primary agent's decisions. As noted in subsection II-B-3, even if accountability makes the overseer's assessments more accurate, it may also make the overseer less willing to act on those assessments.

The fact that this overseer also has veto power introduces additional benefits and costs along the lines discussed in subsection 2.2.2. The main benefit of granting veto power is that, if the overseer acts as a good-enough agent, the veto will enable the overseer to prevent inappropriate policies from going into effect, rather than relying only on indirect influence via the legitimation effect. The main costs are, first, that veto power may undermine the informational value of the overseer by making her less willing to disagree with the primary agent, and, second, that veto power may create a “bailout effect” that can exacerbate the over-accountability distortions afflicting the primary agent's decisions. The foregoing is simply a restatement of the main insights developed in subsections 2.2.2 and 2.2.3. However, these effects may interact in important ways. For example, giving an accountable overseer veto power could mitigate some of the problems with overseer accountability discussed in subsection 2.2.3. Suppose, for instance, that an accountable overseer without veto power would act as a “no man,” always opposing the primary agent's decisions. This might occur because opposition has neither policy costs (because opposition does not actually change the policy) nor reputational costs (because even though disagreement hurts the overseer, it hurts the primary agent more¹³⁶). Giving the overseer veto power may mitigate this problem, because now the

136 Alternatively, this may occur if, due to threshold effects, the overseer can only improve its electoral chances if it opposes a policy that subsequent events reveal was misguided.

overseer's decision to oppose has tangible policy consequences: the policy, which the overseer may well view as a good idea, does not go into effect if the overseer opposes it.

Consider another example. Suppose that when the overseer has veto power but lacks accountability, a strong bailout effect produces substantial over-accountability for the primary decision maker. The overseer might be better off if it could pre-commit to being more deferential, but such pre-commitment may be hard to achieve: after the primary agent has acted (and with the understanding that the primary agent often panders, postures, or what have you), the overseer's rational decision is to exercise her independent judgment and sometimes to veto the primary agent's decision. This may hurt both of their reputations, but the overseer does not suffer any cost from this (because she is not accountable) and the primary agent is willing to bear the cost (because the bailout effect is large relative to the legitimation effect). Now suppose the overseer is accountable, and suppose one consequence of this is that the overseer is more deferential because of the reputational damage associated with disagreement. This will increase deference relative to the case of an insulated overseer, and consequently will mitigate the bailout effect, which in turn may ameliorate over-accountability problems.

Thus, the effects of overseer accountability and veto power are not merely cumulative, but may be interactive. In the examples above, giving the overseer *both* veto power *and* political accountability was better than giving the overseer only one or the other. One can also construct examples in which the opposite is true—examples in which an insulated overseer with veto power, or an accountable overseer without veto power, are both superior to an accountable overseer with veto power.¹³⁷ There are other possibilities as well, which we do not explore here due to time and space limitations. The larger point is that although oversight may be one way to mitigate over-accountability problems while preserving the benefits of accountability for the primary decision makers, the design of such oversight may implicate complex trade-offs that defy easy generalization. The discussions above are intended as a rough guide to some of the key issues that can arise.

137 For instance, consider an overseer with veto power, and imagine that making the overseer accountable would improve her expected ability to distinguish good policies from bad policies. If this effect is stronger than the impact of accountability on the overseer's incentive to defer, then an accountable overseer may become more aggressive about using its veto, and this could strengthen the bailout effect and lead to more over-accountability. See Fox & Stephenson (2011).

3. CONCLUSION

This Article has described a range of over-accountability problems—systematic undesirable behaviors triggered or exacerbated by the accountability of agents to their principals—and considered a range of legal-institutional responses. Our analysis makes three main contributions to current debates in legal scholarship about public law and institutional design.

First, our analysis adds to the literature that challenges the invocation of “accountability” as a rhetorical trump card that is sometimes used to justify particular institutional arrangements or legal/constitutional positions.¹³⁸ Although accountability is often desirable, it is not an end in itself. Rather, accountability is valuable only insofar as it helps to redress the principal–agent problem inherent in representative government or other similar relationships. Sometimes accountability, and the various mechanisms designed to promote accountability, serve this end, but sometimes they do not. Indeed, under some circumstances, promoting accountability can actually make the principal–agent problem worse.

Second, we provide a more sustained analysis of a particular kind of accountability problem, which we call over-accountability. In particular, we draw on relevant work in political economy to better understand how an agent’s interest in appearing to be competent and unbiased may lead her instead to act in ways that the agent believes do not in fact serve the principals’ true interests. A range of familiar undesirable behaviors—which we have labeled pandering, posturing, persistence, populism, and political correctness—may all derive from this same underlying dynamic. Institutional designers (including legal reformers) should attend to the prevention and remediation of these distortions where possible. While we are hardly the first to point out these problems, our hope is that labeling them, and elucidating their underlying causal mechanism, will facilitate the diagnosis of such problems when they exist. Of course, these behaviors may have other causes as well, including non-rational emotional or psychological causes, and we have not attempted here to distinguish empirically these different possibilities. Still, understanding the potentially rational-instrumental roots for seemingly irrational political behavior may be helpful in designing appropriate legal-institutional responses.

Our third contribution is to explore some of the ways that institutional design choices might help mitigate over-accountability problems. While the most straightforward response is to curtail or eliminate accountability, the more interesting questions concern whether supplemental institutional

138 See, e.g., Bressman (2003); Brown (1998); Stephenson (2008b).

checks, such as third-party oversight, might help redress certain over-accountability distortions, while at the same time preserving some of the benefits associated with accountability. It turns out that this may sometimes be possible, but the relationship is far from straightforward. Oversight—whether by courts, the political branches, the media, other governmental or non-governmental watchdogs, or some combination—may sometimes mitigate over-accountability problems, but may sometimes make those problems worse, or introduce new ones. Much depends on the nature and type of overseer, as well as the nature of the underlying over-accountability problem. We have attempted to provide a preliminary discussion of some of the considerations at play, but future research will need to delve more deeply into this issue in order to determine the appropriate institutional response in any given context.

In sum, over-accountability problems are a pervasive feature of our legal-political system, and institutional designers, legal scholars, and others should pay careful attention to both their accurate diagnosis and appropriate treatment. Otherwise, mechanisms that are intended to make our governing institutions more effective and more representative may end up having the opposite effect.

REFERENCES

- Abrams, Samuel A., Morris P. Fiorina, & Jeremy C. Pope. 2008. Polarization in the American Public: Miconceptions and Misreadings. *J. Politics* **70**, 556.
- Acemoglu, Daron, Georgy Egorov, & Konstantin Sonin. 2013. A Political Theory of Populism. *Quart. J. Econ.* **128**, 771.
- Addams, Jane. 1902. *Democracy and Social Ethics*. Norwood, MA: Norwood Press.
- Adler, Matthew D., & Eric A. Posner. 1999. Rethinking Cost-Benefit Analysis. *Yale L. J.* **109**, 165.
- Alexander, Larry, & Fred Schauer. 1997. On Extrajudicial Constitutional Interpretation. *Harvard L. Rev.* **110**, 1359.
- Amar, Akhil Reed. 1991. The Bill of Rights as a Constitution. *Yale L. J.* **100**, 1121.
- Anderson, Elizabeth S., & Richard H. Pildes. 1990. Slingshot Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics. *Columbia L. Rev.* **90**, 2121.
- Aranson, Peter H. 1982. A Theory of Legislative Delegation. *Cornell L. Rev.* **68**, 1.

- Arnold, R. Douglas. 1993. Can Inattentive Citizens Control Their Representatives? In Oppenheimer Dodd, *Congress Reconsidered*, 5th edn. Washington, DC: CQ Press.
- Arrow, Kenneth. 1963. *Social Choice and Individual Values*, 2nd edn. New York: Wiley.
- Ashworth, Scott. 2005. Reputational Dynamics and Political Careers. *J. L., Econ., Organ.* **21**, 441.
- Ashworth, Scott, & Ethan Bueno de Mesquita. 2006. Delivering the Goods: Legislative Particularism in Different Electoral and Institutional Settings. *J Politics* **68**, 168.
- . 2008. Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage. *J. Politics* **70**, 1006.
- . 2014., Multitask, Accountability, and Institutional Design. Unpublished manuscript.
- Ashworth, Scott, & Kenneth W. Shotts. 2010. Does Informative Media Commentary Reduce Politicians' Incentive to Pander? *J. Public Econ.* **94**, 838.
- Avery, Christopher, & Margaret Meyer. 2012. Reputational Incentives for Biased Evaluators. Unpublished manuscript.
- Bagley, Nicholas, & Richard L. Revesz. 2006. Centralized Oversight of the Regulatory State. *Columbia L. R.* **106**, 1260.
- Baliga, Sandeep, & Tomas Sjostrom. 2001. Optimal Design of Peer Review and Self-Assessment Schemes. *Rand J. Econ.* **32**, 27.
- Banerjee, Abhijit V. 1992. A Simple Model of Herd Behavior. *Quart. J. Econ.* **107**, 797.
- Banks, Jeffrey S., & Rangarjan Sundaram. 1998. Optimal Retention in Agency Problems. *J. Econ. Theory* **82**, 293.
- Barro, Robert J. 1973. The Control of Politicians: An Economic Model. *Public Choice* **14**, 19.
- Barron, David J. 2008. From Takeover to Merger: Reforming Administrative Law in an Age of Agency Politicization. *The George Washington L. Rev.* **76**, 1095.
- Bauerly, Cynthia L., & Eric C. Hallstrom. 2012. Square Pegs: The Challenges for Existing Federal Campaign Finance Disclosure Laws in the Age of the Super PAC. *New York University J. Legislation and Public Policy* **15**, 329.
- Berger, Raoul. 1977. *Government by Judiciary*. Cambridge: Harvard University Press.
- Bernheim, B. Douglas. 1994. A Theory of Conformity. *J. Political Econ* **102**, 841.
- Berry, Christopher R., & Jacob E. Gersen. 2008. The Unbundled Executive. *The University of Chicago L. Rev.* **75**, 1385.

- Besley, Timothy. 2007. *Principled Agents: The Political Economy of Good Government*. New York: Oxford University Press.
- Besley, Timothy, & Ann Case. 2003. Political Institutions and Policy Choices: Evidence from the United States. *J. Econ. Lit.* **41**, 7.
- Besley, Timothy, & Stephen Coate. 2003. Elected versus Appointed Regulators: Theory and Evidence. *J. Eur. Econ. Association* **1**, 1176.
- Besley, Timothy. 2008. Issue Unbundling via Citizens' Initiatives. *Quart. J. Pol. Sci.* **3**, 379.
- Bibas, Stephanos. 2009. Prosecutorial Regulation versus Prosecutorial Accountability. *Uni. Pennsylvania L. Rev.* **157**, 959.
- Bickel, Alexander M. 1962. *The Least Dangerous Branch: The Supreme Court at the Bar of Politics*. Binghamton: Vail-Ballou Press, Inc.
- Bikhchandani, Sushil, David Hirshleifer, & Ivo Welch. 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades. *J. Pol. Econ.* **100**, 992.
- Bolton, Patrick, & Mathias Dewatripont. 2005. *Contract Theory*. Cambridge, MA: The MIT Press.
- Bork, Robert H. 1990. *The Tempting of America: The Political Seduction of the Law*. New York: Touchstone.
- Bradley, Curtis A., & Jack L. Goldsmith. 2005. Congressional Authorization and the War on Terrorism. *Harvard L. Rev.* **118**, 2047.
- Bradley, Curtis A., & Eric A. Posner. 2006. Presidential Signing Statements and Executive Power. *Constitutional Commentary* **23**, 307.
- Bressman, Lisa Schultz. 2003. Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State. *New York Uni. L. Rev.* **78**, 461.
- Bressman, Lisa Schultz, & Michael P. Vandenbergh. 2006. Inside the Administrative State: A Critical Look at the Practice of Presidential Control. *Michigan L. Rev.* **105**, 47.
- Breyer, Stephen. 1984. The Legislative Veto after *Chadha*. *Georgetown L. J.* **72**, 785.
- Briffault, Richard. 2001. *Nixon v Shrink Missouri Government PAC*: The Beginning of the End of the *Buckley* Era? *Minnesota L. Rev.* **85**, 1729.
- Brown, Rebecca L. 1998. Accountability, Liberty, and the Constitution. *Columbia L. Rev.* **98**, 531.
- Bueno de Mesquita, Bruce *et al.* 2004. *The Logic of Political Survival*. Cambridge, MA: The MIT Press.
- Bueno de Mesquita, Ethan. 2007. Politics and the Suboptimal Provision of Counterterrorism. *Intl. Organ.* **61**, 9.
- Bueno de Mesquita, Ethan, & Dimitri Landa. 2014. *Political Accountability and Sequential Policy Making*. Unpublished manuscript.

- Bueno de Mesquita, Ethan, & Matthew C. Stephenson. 2007. Regulatory Quality under Imperfect Oversight. *Am. Pol. Sci. Rev.* **101**, 605.
- Calabresi, Steven G. 1995. Some Normative Arguments for the Unitary Executive. *Arkansas L. Rev.* **48**, 23.
- Calabresi, Steven G., & Saikrishna B. Prakash. 1994. The President's Power to Execute the Laws. *Yale L. J.* **104**, 541.
- Campbell, Angus *et al.* 1960. *The American Voter*. Chicago: The University of Chicago Press.
- Caminker, Evan H. 2003. Thayerian Deference to Congress and Supreme Court Super Majority Rule. *Indiana L. J.* **78**, 73.
- Canes-Wrone *et al.* 2001. Agentship and Pandering: A Theory of Executive Policymaking. *Am. J. Pol. Sci.* **45**, 532.
- Caplan, Bryan. 2007. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton, NJ: Princeton University Press.
- Carter, William M. Jr. 2012. The Paradox of Political Power: Post-Racialism, Equal Protection, and Democracy. *Emory L. J.* **61**, 1123.
- Chemerinsky, Erwin. 1995. The Values of Federalism. *Florida L. Rev.* **47**, 499.
- Cohen, Linda, & Matthew Spitzer. 1992. Term Limits. *Georgetown L. J.* **80**, 477.
- Coram, Bruce Talbot. 1996. Second Best Theories and the Implications for Institutional Design. In Robert E. Goodin, *The Theory of Institutional Design*. New York: Cambridge University Press.
- Croley, Steven. 2003. White House Review of Agency Rulemaking: An Empirical Investigation. *Uni. Chicago L. Rev.* **70**, 821.
- Cross, Frank B. 1988. The Constitutional Legitimacy and Significance of Presidential "Signing Statements". *Admin. L. Rev.* **40**, 209.
- Dahl, Robert A. 1989. *Democracy and its Critics*. New Haven: Yale University Press.
- DeShazo, J.R., & Jody Freeman. 2005. Public Agencies as Lobbyists. *Columbia L. Rev.* **105**, 2217.
- Dorf, Michael C. 2006. *No Litmus Test: Law Versus Politics in the Twenty-First Century*. Oxford, UK: Rowman & Littlefield Publishers.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Downs, George W., & David M. Rocke. 1994. Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War. *Am. J. Pol. Sci.* **38**, 362.
- Dworkin, Ronald. 1977. *Taking Rights Seriously*. Cambridge: Harvard University Press.
- Effinger, Matthias R., & Matthias K. Polborn. 2001. Herding and Anti-Herding: A Model of Reputational Differentiation. *Eur. Econ. Rev.* **45**, 385.
- Elster, Jon. 1979. *Ulysses and the Sirens*. New York: Cambridge University Press.

- Epstein, Richard A. 2011. *Design for Liberty: Private Property, Public Administration, and the Rule of Law*. Cambridge: Harvard University Press.
- Falaschetti, Dino, & Gary J. Miller. 2001. Constraining Leviathan: Moral Hazard and Credible Commitment in Institutional Design. *J. Theoretical Pol.* **13**, 389.
- Farina, Cynthia R. 1989. Statutory Interpretation and the Balance of Power in the Administrative State. *Columbia L. Rev.* **89**, 452.
- Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types Versus Sanctioning Poor Performance. *Adam Przeworski et al. eds., Democracy, Accountability, and Representation*. New York: Cambridge University Press.
- Feeley, Malcolm, & Edward L. Rubin. 1994. Federalism: Some Notes on a National Neurosis. *UCLA L. Rev.* **41**, 903.
- Fenster, Mark. 2006. Some Opacity about Transparency. *Iowa L. Rev.* **91**, 885.
- Ferejohn, John. 1986. Incumbent Performance and Electoral Control. *Public Choice* **50**, 5.
- Fiorina, Morris. 1996. *Divided Government*, 2dn edn. Needham Heights, MA: Allyn & Bacon.
- Fisher, Louis. 1993. The Legislative Veto: Invalidated, It Survives. *L. & Contemporary Problems* **56**, 273.
- . 2000. *Congressional Abdication of War and Spending*. College Station, TX: Texas A&M University Press.
- Fox, Justin. 2007. Government Transparency and Policymaking. *Public Choice* **131**, 23.
- Fox, Justin, & Matthew C. Stephenson. 2011. Judicial Review as a Response to Political Posturing. *Am. Pol. Sci. Rev.* **105**, 397.
- Fox, Justin. 2014a. The Constraining, Liberating, and Informational Effects of Nonbinding Law. *J. L., Econ. Org.*, doi:10.1093/jleo/ewu013.
- . 2014b. The Welfare Effects of Minority-Protective Judicial Review. *Journal of Theoretical Politics*, doi:10.1177/0951629814547301.
- Fox, Justin, & Richard Van Weelden. 2010. Partisanship and the Effectiveness of Oversight. *J. Pub. Econ.* **94**, 674.
- Friedman, Barry. 1997. Valuing Federalism. *Minnesota L. Rev.* **82**, 317.
- Fu, Qiang, & Ming Li. 2014. Reputation-Concerned Policy Makers and Institutional Design. *J. Pub. Econ.* **110**, 15.
- Garber, Marc N., & Kurt A. Wimmer. 1987. Presidential Signing Statements as Interpretations of Legislative Intent: An Executive Aggrandizement of Power. *Harvard J. Legislation* **24**, 363.
- Garrett, Elizabeth. 1996. Term Limitations and the Myth of the Citizen-Legislator. *Cornell L. Rev.* **81**, 623.

- Gentzkow, Matthew, & Jesse M. Shapiro. 2006. Media Bias and Reputation. *J. Pol. Econ.* **114**, 280.
- Gersen, Jacob E. 2007. Overlapping and Underlapping Jurisdiction in Administrative Law. *The Supreme Court Rev.* **2006**, 201.
- . 2010. Unbundled Powers. *Virginia L. Rev.* **96**, 301.
- Gersen, Jacob E., & Eric A. Posner. 2008. Soft Law: Lessons from Congressional Practice. *Stanford L. Rev.* **61**, 573.
- Glazer, Amihai. 2007. Strategies of the Political Opposition. Unpublished manuscript.
- Goodin, Robert E. 1995. Political Ideas and Political Practice. *British J. Pol. Sci.* **25**, 37.
- Gordon, Sanford C., & Gregory A. Huber. 2004. Accountability and Coercion: Is Justice Blind When It Runs for Office? *Am. J. Pol. Sci.* **48**, 48.
- Gordon, Sanford C. 2007. The Effect of Electoral Competitiveness on Incumbent Behavior. *Quart. J. Pol. Sci.* **2**, 107.
- Gordon, Sanford C. *et al.* 2007. Challenger Entry and Voter Learning. *Am. Pol. Sci. Rev.* **101**, 303.
- Gowrisankaran, Gautam *et al.* 2008. Electoral Design and Voter Welfare from the U.S. Senate: Evidence from a Dynamic Selection Model. *Rev. Econ. Dynamics* **11**, 1.
- Grimmett, Richard F. 2010. *The War Powers Resolution: After Thirty-Six Years*, Cong. Research Serv., R41199.
- Guinier, Lani. 1994. *The Tyranny of the Majority: Fundamental Fairness in Representative Democracy*. New York: The Free Press.
- Hayward, Allison R. 2010. Bentham & Ballots: Tradeoffs between Secrecy and Accountability in How We Vote. *J. L. & Pol.* **26**, 39.
- Hills, Roderick M. 1998. The Political Accountability of Cooperative Federalism: Why State Autonomy Makes Sense and “Dual Sovereignty” Doesn’t. *Michigan L. Rev.* **96**, 813.
- Hinich, Melvin, & Walker Pollard. 1981. A New Approach to the Spatial Theory of Electoral Competition. *Am. J. Pol. Sci.* **25**, 323.
- Holmes, Stephen. 1995. *Passions and Constraint*. Chicago: The University of Chicago Press.
- Holmstrom, BengtH, & Paul Milgrom. 1992. Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design. *J. L., Econ. Organ.* **7**, 24.
- Horwitz, Paul. 2008. Three Faces of Deference. *Notre Dame L. Rev.* **83**, 1061.
- Huber, John D., & Charles R. Shipan. 2002. *Deliberate Discretion: The Institutional Foundations of Bureaucratic Autonomy*. New York: Cambridge University Press.

- Jacobs, Lawrence R., & Robert Y. Shapiro. 2000. *Politicians Don't Pander: Political Manipulation and the Loss of Democratic Responsiveness*. Chicago: The University of Chicago Press.
- Kagan, Elena. 2001. Presidential Administration. *Harvard L. Rev.* **114**, 2245.
- Kennedy, John F. 1955. *Profiles in Courage*. New York: Harper & Brothers.
- Komesar, Neil K. 1994. *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy*. Chicago: The University of Chicago Press.
- Kuran, Timur, & Cass R. Sunstein. 1999. Availability Cascades and Risk Regulation. *Stanford L. Rev.* **51**, 683.
- Kurland, Philip B. 1989. The Supreme Court and Its Judicial Critics. *Utah L. Rev.* **6**, 457.
- Kydland, Finn E., & Edward C. Prescott. 1977. Rules Rather than Discretion: The Inconsistency of Optimal Plans. *J. Pol. Econ.* **85**, 473.
- La Pierre, D. Bruce. 1982. The Political Safeguards of Federalism Redux: Intergovernmental Immunity and the States as Agents of the Nation. *Washington Uni. L. Quart.* **60**, 779.
- . 1985. Political Accountability in the National Political Process—The Alternative to Judicial Review of Federalism Issues. *Northwestern Uni. L. Rev.* **80**, 577.
- Lancaster, Kelvin, & R.G. Lipsey. 1956. The General Theory of the Second Best. *Rev. Econ. Stud.* **24**, 11.
- Landis, James. 1938. *The Administrative Process*. New Haven, CT: Yale University Press.
- Law, David S. 2009. A Theory of Judicial Power and Judicial Review. *Georgetown L. J.* **97**, 723.
- Lawson, Gary, & Christopher D. Moore. 1996. The Executive Power of Constitutional Interpretation. *Iowa L. Rev.* **81**, 1267.
- Lessig, Lawrence, & Cass R. Sunstein. 1994. The President and the Administration. *Columbia L. Rev.* **94**, 1.
- Levinson, Daryl J. 2011. Parchment and Politics: The Positive Puzzle of Constitutional Commitment. *Harvard L. Rev.* **124**, 657.
- . 2012. Rights and Votes. *The Yale L. J.* **121**, 1286.
- Levmore, Saul. 1992. Bicameralism: When Are Two Decisions Better than One? *Intl. Rev. L. Econ.* **12**, 145.
- Levy, Gilat. 2004. Anti-herding and Strategic Communication. *Eur. Econ. Rev.* **48**, 503.
- . 2005. Careerist Judges and the Appeals Process. *Rand J. Econ.* **36**, 275.
- Lewis, David E. 2008. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*. Princeton: Princeton University Press.
- Listokin, Yair. 2008. Learning through Policy Variation. *Yale L. J.* **118**, 480.

- Lohmann, Susanne. 1998. An Information Rationale for the Power of Special Interests. *Am. Pol. Sci. Rev.* **92**, 809.
- Lopez, Ian Hanley. 2011. Freedom, Mass Incarceration, and Racism in the Age of Obama. *Alabama L. Rev.* **62**, 1005.
- Lowi, Theodore. 1969. *The End of Liberalism*. New York: W.W. Norton & Company.
- Lupia, Arthur, & Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.
- Mackie, Gerry. 2003. *Democracy Defended* New York: Cambridge University Press.
- Majumdar, Sumon, & Sharun W. Mukand. 2004. Policy Gambles. *Am. Econ. Rev.* **94**, 1207.
- Marisam, Jason. 2011. Duplicative Delegations. *Admin. L. Rev.* **63**, 181.
- Markovits, Richard S. 1975. A Basic Structure for Microeconomic Policy Analysis in Our Worse-Than-Second-Best World: A Proposal and Related Critique of the Chicago Approach to the Study of Law and Economics. *Wisconsin L. Rev.* **1975**, 950.
- Mashaw, Jerry L. 1985. Prodelegation: Why Administrators Should Make Political Decisions. *J. L., Econ. Organ.* **1**, 81.
- Maskin, Eric, & Jean Tirole. 2004. The Politicians and the Judge: Accountability in Government. *Am. Econ. Rev.* **94**, 1034.
- Merritt, Deborah J. 1988. The Guarantee Clause and State Autonomy: Federalism for a Third Century. *Columbia L. Rev.* **88**, 1.
- Miller, Gary J. 2005. The Political Evolution of Principal-Agent Models. *Annual Rev. Pol. Sci.* **8**, 203.
- Miller III, James C. 2011. The Early Days of Reagan Regulatory Relief and Suggestions for OIRA's Future. *Admin. L. Rev.* **63**, 93.
- Moe, Terry M. 1984. The New Economics of Organization. *Am. J. Pol. Sci.* **28**, 739.
- Morelli, Massimo, & Richard Van Weelden. 2014. Re-Election through Division. Unpublished manuscript.
- Negri, Margherita. 2011. Why Do Good Politicians Take Bad Actions? Unpublished manuscript.
- Nzelibe, Jide O., & Matthew C. Stephenson. 2010a. Complementary Constraints: Separation of Powers, Rational Voting, and Constitutional Design. *Harvard L. Rev.* **123**, 617.
- . 2010b. Political Accountability under Alternative Institutional Regimes. *J. Theoretical Pol.* **22**, 139.
- Patashnik, Eric M. 2000. *Putting Trust in the U.S. Budget: Federal Trust Funds and the Politics of Commitment*. New York: Cambridge University Press.

- Paulsen, Michael Stokes. 1994. The Most Dangerous Branch: Executive Power to Say What the Law is. *Georgetown L. J.* **83**, 217.
- Persson, Torsten, & Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: The MIT Press.
- Pierce, Jr., Richard J. 1989. The Role of the Judiciary in Implementing an Agency Theory of Government. 64 *New York Uni. L. Rev.* 1239.
- Posner, Eric A., & Adrian Vermeule. 2007. The Credible Executive. *Uni. Chicago L. Rev.* **74**, 865.
- Posner, Eric, & Adrian Vermeule. 2013. Inside or Outside the System. *Uni. Chicago L. Rev.* **80**, 1743.
- Prat, Andrea. 2005. The Wrong Kind of Transparency. *Am. Econ. Rev.* **95**, 862.
- Prendergast, Canice. 1993. A Theory of ‘Yes Men’. *Am. Econ. Rev.* **83**, 757.
- Prendergast, Canice, & Lars Stole. 1996. Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning. *J. Pol. Econ.* **104**, 1105.
- Riker, William H. 1982. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. San Francisco: Freeman.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. New York: Oxford University Press.
- Rogers, James R. 2009. Casting the Gimlet Eye on Judicial Review: Can Judicial Review Be Democratically Debilitating? Unpublished manuscript.
- Rogoff, Kenneth. 1985. The Optimal Degree of Commitment to an Intermediate Monetary Target. *Quart. J. Econ.* **100**, 1169.
- Samaha, Adam M. 2006. Government Secrets, Constitutional Law, and Platforms for Judicial Intervention. *UCLA L. Rev.* **53**, 909.
- Scharfstein, David S., & Jeremy C. Stein. 1990. Herd Behavior and Investment. *Am. Econ. Rev.* **80**, 465.
- Scherer, Nancy. 2005. *Scoring Points: Politicians, Activists, and the Lower Federal Court Appointment Process*. Stanford, CA: Stanford University Press.
- Schleicher, David. 2008. Irrational Voters, Rational Voting. *Election L. J.* **7**, 149.
- Schmitz, Gerald. 1988. *The Opposition in a Parliamentary System*. Canada: Library of Parliament, Available at: <http://www.parl.gc.ca/Content/LOP/researchpublications/bp47-e.pdf>.
- Schoenbrod, David. 1993. *Power Without Responsibility: How Congress Abuses the People through Delegation*. New Haven: Yale University Press.
- Shelton, Dinah. 2003. Legal Norms to Promote the Independence and Accountability of International Tribunals. *L. Practice of Intl Courts and Tribunals* **2**, 27.
- Smith, Bradley A. 1996. Faulty Assumptions and Undemocratic Consequences of Campaign Finance Reform. *Yale L. J.* **105**, 1049.
- Sohoni, Mila. 2012. The Idea of “Too Much Law”. *Fordham L. Rev.* **80**, 1585.

- Solum, Lawrence B. 2008. Constitutional Possibilities. *Indiana L. J.* **83**, 307.
- Somin, Ilya. 2004. Political Ignorance and the Countermajoritarian Difficulty: A New Perspective on the Central Obsession of Constitutional Theory. *Iowa L. Rev.* **89**, 1287.
- Stephenson, Matthew C. 2007a. Bureaucratic Decision Costs and Endogenous Agency Expertise. *J. L., Econ. Organ* **23**, 469.
- . 2007b. Judicial Reform in Developing Economies: Constraints and Opportunities. In Francois Bourguignon & Boris Pleskovic, eds., *Annual World Bank Conference on Development Economics: Beyond Transition*. Washington, DC: The World Bank.
- . 2008a. The Price of Public Action: Constitutional Doctrine and the Judicial Manipulation of Legislative Enactment Costs. *Yale L. J.* **118**, 2.
- . 2008b. Optimal Political Control of the Bureaucracy. *Michigan L. Rev.* **107**, 53.
- . 2011. Information Acquisition and Institutional Design. *Harvard L. Rev.* **124**, 1422.
- Stewart, Richard B. 1977. Pyramids of Sacrifice? Problems of Federalism in Mandating State Implementation of National Environmental Policy. *Yale L. J.* **86**, 1196.
- Stiglitz, Joseph E. 2000. *Economics of the Public Sector*, 3rd edn. New York: W.W. Norton & Company.
- Stone, Geoffrey R. 2012. Secrecy and Self-Governance. *NY L. School L. Rev.* **56**, 81.
- Straw, Christopher M. 1996. The Role of Electoral Accountability in the Madisonian Machine. *NY Uni. J. Legislation and Public Policy* **11**, 321, 358–61.
- Thayer, James Bradley. 1893. The Origin and Scope of the American Doctrine of Constitutional Law. *Harvard L. Rev.* **7**, 129.
- Tushnet, Mark. 1999. *Taking the Constitution Away from the Courts*. Princeton: Princeton University Press.
- Ulen, Thomas S. 1998. Courts, Legislatures, and the General Theory of the Second Best in Law and Economics. *Chicago-Kent L. Rev.* **73**, 189.
- Vermeule, Adrian. 2003. Hume's Second-Best Constitutionalism. *Uni. Chicago L. Rev.* **70**, 421.
- . 2006. *Judging under Uncertainty: An Institutional Theory of Legal Interpretation*. Cambridge, MA: Harvard University Press.
- . 2007. *Mechanisms of Democracy: Institutional Design Writ Small*. New York: Oxford University Press.
- . 2009. Forward: System Effects and the Constitution. *Harvard L. Rev.* **123**, 4.

- . 2011. Second Opinions and Institutional Design. *Virginia L. Rev.* 97, 1435.
- Waldron, Jeremy. 1999. *Law and Disagreement*. New York: Oxford University Press.
- Waller, Christopher J. 2000. Policy Boards and Policy Smoothing. *Quart. J. Econ.* 115, 305.
- Wolfers, Justin. 2007. *Are Voters Rational? Evidence from Gubernatorial Elections*. Unpublished manuscript.
- Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.