

Experimentation and Persuasion in Political Organizations

Alexander V. Hirsch ¹

February 7, 2014

¹Department of Politics and Woodrow Wilson School of Public and International Affairs, Princeton University, 041 Corwin Hall, Princeton, NJ 08544. Email: avhirsch@princeton.edu.

Abstract

There is ample evidence that different beliefs about how to achieve shared goals are common in political organizations like government agencies, campaigns, and NGOs. However, the consequences of such conflicts have not yet been explored. We develop a formal model in which a principal and an agent disagree about the right policy for achieving their shared goals. Disagreement creates a motivational problem, but we show how both observing policy outcomes and experimenting with policies can ameliorate it. We also show that the principal often defers to the agent in order to motivate him, thereby generating more informative policy outcomes and building future consensus. Most surprisingly, she sometimes allows the agent to implement his desired policy even when she is sure it is wrong to persuade him through failure that he is mistaken. Using the model we generate empirical implications about performance measurement and Presidential appointments in U.S. federal agencies.

1 Introduction

Within political organizations, strong and open disagreements about how to achieve shared goals are common. For example, President Obama, his cabinet, and the Joint Chiefs openly held strong and differing opinions about how many troops were necessary to succeed with an Afghanistan “surge” (Gates 2014). Similarly, during the 2006 election cycle there was vicious public disagreement within the Democratic party about the wisdom of Howard Dean’s “50-state strategy” – prominent Democratic strategist Paul Begala famously described it as “just hiring a bunch of staff people to wander around Utah and Mississippi and pick their nose” (Bai 2006). Anecdotes aside, political organizations like public bureaucracies, political campaigns, and NGOs have many structural features that make them particularly vulnerable to belief conflicts. Individuals often choose to work for such organizations precisely because of their zeal for the “public good” and their strong beliefs about how to achieve it (Besley and Ghatak 2006, Perry and Hondeghem 2008). In addition, the manner in which the federal government is staffed – with elected political principals, appointed political agents, and long-term career bureaucrats – all but ensures that individuals with very different personal backgrounds and professional training will be forced to work together (Heclo 1977).

Among scholars of policy studies, the notion that conflict can be driven by different beliefs about “how the world works” is familiar (Sabatier 1988). However, scholars of political agency and bureaucratic politics have yet to analyze the consequences of open “disagreement on beliefs,” and whether such conflicts really “differ significantly from those caused by conflicting objectives” (Bendor and Hammond 1992).¹ In particular, despite the frequency of such disagreements in government agencies, it is not yet known how they affect policy decisions and the delegation of Congressional authority (Bendor and Meirowitz 2004), not to mention internal capacity (Ting 2011), performance (Lewis 2008), and organizational learn-

¹In a rare exception, Bendor and Hammond (1992) briefly conjecture that belief conflicts in organizations are generally less destructive than goal conflicts because they might be resolved through deliberation.

ing (Callander 2011). The canonical model of political agency instead considers a principal-agent relationship in which an agent has both different goals and better information than his principal, and focuses on the tradeoff between the value of the agent’s “greater expertise about the likely effects” of different policies, and the danger that the agent might “pursue goals that diverge from those of” the principal (Stephenson 2007). While such models have yielded deep insights about inter-branch conflict and the design of bureaucratic institutions, they cannot be used to study the consequences of *open* belief disagreement because, by construction, any disagreement cannot be open – if the principal knew the agent’s beliefs, then she could also infer his information and use it to revise her own beliefs.²

More generally, the inability of rational actors to persist in open disagreement is a well-known property of models which assume that differences in beliefs are solely the result of differences in information, which is known in economic theory as the *common priors assumption* (Aumann 1987).³ However, the reality of contemporary politics is that differences in beliefs about “how the world works” often persist even after vigorous deliberation and debate (Mutz 2008). In this paper we therefore develop a formal model to understand the implications of such disagreement for the management of political organizations. Specifically, we consider a model in which a principal and an agent have common goals, but *heterogeneous prior beliefs* about which of two available policy instruments will be effective for achieving them. While assuming heterogeneous priors is unorthodox, there is a small but growing literature in political science and economics that does so to study a variety of topics, including conflict (Smith and Stam 2004), bargaining (Yildiz 2004), expertise acquisition (Che and

²See Gailmard and Patty (2012) for a review.

³Specifically, with common priors there cannot be common knowledge of disagreement; moreover, if beliefs are truthfully communicated back-and-forth then they cannot diverge indefinitely (Geanakoplos and Polemarchakis 1982). Worth noting is that Aumann’s result does not mean that open disagreement cannot result from differences in information; just that it cannot result from differences in information *only*. For example, Geanakoplos (1994) discusses how open disagreement can result from the combination of different information and processing errors.

Kartik 2009), and firm management (Van den Steen 2010).⁴

Our baseline model has a simple structure. The principal and agent share a common goal that they wish to achieve in each of two periods. In each period, the principal can choose from two possible policies to achieve the goal. However, her policy must be implemented by the agent, whose effort is not directly observable to the principal.⁵ The agent’s effort increases the chance that an effective policy succeeds, but only one of the two policies is “in truth” effective, and the principal and agent disagree about which policy this is (in the sense of heterogeneous priors). Finally, the direct effects of the policies cannot be immediately observed; the principal and agent therefore cannot resolve their disagreement through communication, deliberation, or learning.

To understand the logic of the model, consider the following highly stylized example. A city is attempting to combat a wave of violent crime over time. However, the true underlying cause of the crime wave is unknown. One proposed theory is that the city has simply accumulated a large number of violent criminals. The other is that the presence of petty crime and general disorder causes the community to disengage and thereby encourages all crime, i.e. “broken windows” theory (Wilson and Kelling 1982). Consequently, the city’s mayor (the principal) and her police chief (the agent) disagree over whether the city’s approach to controlling the crime wave should focus predominantly on solving violent crimes and apprehending their perpetrators, or focus on maintaining community order by punishing petty crimes. Finally, whichever policy is chosen must also be implemented by the police force with sufficient effort to be successful.

An important feature of our baseline model is that the agent cannot be directly monitored or incentivized to take certain actions. This is a realistic depiction of management in most political organizations including public bureaucracies, where employment contracts

⁴See also Morris (1995) and Gul (1998) for arguments in support of relaxing the common priors assumption.

⁵The assumption that effort is unobservable is appropriate for many public bureaucracies, but especially so of “street level bureaucracies” (Lipsky 1980) such as teachers, firefighters, and policemen.

are tightly constrained by legislation and inter-branch bargaining (Lewis 2008). However, the agent is also intrinsically motivated to work; what motivates him is *believing* that he is working to implement an effective policy for achieving the organization’s goals. As a result, disagreement about the right policy creates a difficult managerial problem for the principal. If she imposes the policy that she believes will be best, the agent be less motivated to implement it because he believes it is unlikely to be effective. Alternatively, if she *defers* to the agent by selecting his desired policy, then he will be more motivated to work; but on a policy that the principal believes to be inferior. Returning to our example, if the chief strongly believes in focusing on violent crimes, then he will be demoralized by being forced to waste energy rounding up pick-pockets and fare-beaters. If the mayor instead believes in “broken windows,” then she only has two bad choices; a demoralized chief, or an ineffective policy.

The model described thus far captures what Wilson’s (1989) seminal study of bureaucracy refers to as “coping organizations” in which neither inputs – the agent’s effort – nor outputs – success or failure – can be observed. The principal copes with disagreement in her organization, but cannot resolve it. Consequently, the model is functionally identical to one in which the principal and the agent simply have different preferences, in that similar patterns of behavior can be derived by instead assuming that the principal and agent both like to achieve the common goal, but simply prefer to achieve it with different policies.

What makes beliefs different, however, is that they can *change*. In many environments, policies can be tried out, their outcomes are observed, and those outcomes can be used to draw lessons about how to approach the policy problem in the future (May 1992, Hall 1993). To explore this fundamental distinction between preferences and beliefs, we next consider a second variant of the model in which the players can observe whether the first period policy succeeded prior to choosing and implementing a policy in the second. Unlike the previous variant, the players can now learn more about which policy is correct, and revise their beliefs and actions in the second period. In our example, this is analogous to the city police force

beginning to collect and analyze real-time crime statistics, allowing it to measure the efficacy of its policy and make changes if necessary.

The second variant of our model captures what Wilson (1989) refers to as “craft organizations” – the agent’s effort cannot be observed, but the results of that effort can be both observed and learned from. Importantly, the ability to observe outcomes in this variant also means that the principal can *experiment* with policies – that is, try one out, and decide whether to keep it based on the results. The idea that governments might actively design “policies as experiments” to learn about their efficacy dates back to the era of social engineering in the late 60’s and 70’s (Campbell 1969), and is currently resurgent in the field of public ecological management (Lee 1993). However, although there is a small and growing literature in political science that analyzes policy experimentation (Callander 2011, Volden, Ting, and Carpenter 2008), there is currently no work analyzing the effects of experimentation in a principal-agent setting. By comparing the variants of our model with and without learning, we can therefore learn both about the distinction between belief conflict and goal conflict, and also about the differences between political organizations that can learn and experiment, and those that cannot.

Our main results from the model with learning and experimentation are as follows. First, the agent is both more motivated, and successes are more likely, relative to the model without learning. Surprisingly, this is true even when the principal does not experiment at all, and instead rigidly implements the same policy in both periods. The reason for this effect is crucial for our subsequent results. When a policy is implemented well, its failure can be better attributed to the policy itself rather than its poor implementation; in other words, failure is *informative*. However, when a policy is implemented poorly, a failure could either be due both to the policy itself and to its poor implementation, i.e. it is uninformative. Consequently, agent can actually *make* the first period policy outcome more informative by simply working harder. In equilibrium, he will choose to do so because he values learning about the right policy in order to avoid wasting his effort on the wrong one in the future. Consequently,

when outcomes are observable the agent works both harder and more efficiently, leading to a greater chance of success in both periods.

Our second result is that the agent is even more motivated when the principal experiments with policies, rather than rigidly implementing a single one she believes to be best. Thus, we find that there is a motivational benefit to experimentation that is independent of “searching” for the best policy (Callander 2011). The reason is that experimenting ties the principal’s future policy decision to the initial outcome, which gives the agent the power to indirectly improve policy decisions by working harder to ensure that those outcomes are informative. Moreover, this effect persists even when the agent strongly disagrees with the principal’s initial policy – he never “sabotages” the policy to ensure failure and a change. An additional implication of this result is that the principal sometimes *underexperiments* relative to what would maximize her utility; this occurs when she believes too strongly in the initial policy to credibly abandon it after a single failure. Consequently, we show that the principal can sometimes benefit from institutional arrangements that force her to experiment, which include institutionalizing trigger mechanisms (Nie and Schultz 2012), and delegating her policy decisions to an individual with more moderate beliefs.

Our third result is that the principal’s decisions about which policy to experiment with are biased toward the agent, in the sense of placing greater weight on his beliefs than her own. As a consequence, the principal often experiments with the policy initially desired agent even when she is *more sure* that it is wrong than the agent is sure that it is right. This additional tendency towards deference in the model with learning arises for two reasons. First, the principal can better motivate the agent by deferring to him. Second, when outcomes are observable this extra motivation is more valuable; it increases the informativeness of policy outcomes and thus learning, which improves both players’ future decisions and also helps to build consensus.

Our final, and potentially most novel, result is that the degree of deference exhibited by the principal in our model far exceeds what can occur in traditional political agency models.

In these models, a principal only defers to an agent to exploit his superior expertise about “how the world works” – thus, when she is certain that she already knows the right policy, she never defers. In our model, however, the principal sometimes initially defers to the agent *even when she is already sure that his policy is wrong*. She does so to try to *persuade* him that he is mistaken by allowing him to implement the policy he believes is best, work hard, and nevertheless watch it fail. Public managers often extol the benefits of a hands off management style that “turn[s] responsibility back onto the worker” (Bratton 1998). Our model illustrates how such a style is useful not only for extracting good ideas from subordinates in political organizations, but also helping them learn more quickly when their ideas are bad.

Overall, our analysis makes the following main contributions. First, it begins to identify the differences between conflict over “goals” and conflict over “beliefs” in political organizations. Second, it provides a fresh perspective on a variety of managerial techniques in real-world public bureaucracies that have been extensively discussed in both the political science and public administration literatures; specifically, we show that there are motivational benefits to measuring performance, encouraging experimentation, appointing ideological moderates to run agencies, and exercising deference to subordinates. Third, it generates a variety of new testable empirical implications about management patterns in political organizations, and especially about the effects of performance measurement in U.S. federal agencies. We discuss these implications extensively after presenting the formal model. Finally, it contributes to the small but growing literature on policy experimentation in political environments.

The paper proceeds as follows. In Section 2 we introduce the model, and in Section 3 we analyze the variant without learning. In Section 4 we introduce the model with learning and analyze the agent’s decisions, and in Section 5 we provide four key results about the principal’s policy choices. In Section 6 we discuss empirical implications, and in Section 7 we conclude.

2 The Model

The model is a two period game of policy choice and implementation between a principal and an agent. Player 1 is the principal, and in each period $t \in \{1, 2\}$ publicly chooses a course of action x^t from the set $X = \{a, b\}$. Henceforth the principal's choice is referred to as a policy. However, it may be thought of more generally as how her organization will pursue a goal. Player 2 is the agent. In each period, after observing the principal's choice, the agent implements the organization's policy with unobservable effort $e^t \in [0, 1]$.

The principal and the agent seek to achieve a shared organizational goal in each period; for simplicity we assume that the outcome in each period y^t can either be to succeed at that goal ($y^t = 1$) or to fail ($y^t = 0$). The model is agnostic as to why the agent shares the organization's goals, and many possible reasons are supported in the literature. It could be that the agent selected into, or was chosen by, an organization whose goals he already intrinsically shared.⁶ Alternatively, the agent may have been actively socialized so that she comes to share the organization's goals, as is common in "mission driven organizations" in both the public and nonprofit sectors (Kaufman 1967, DiIulio 1987, Goodsell 2011). The agent could have an intrinsic preference for "doing his job" and define it as fulfilling his superiors' goals; such "role perception" by bureaucrats is supported by case study research (Wilson 1989, Golden 2000). Or success could align with the agent's career concerns by ensuring his organization remain well funded or enhancing his promotion prospects.

Exerting effort is assumed to be costly to the agent, and this cost takes the form $-(e^t)^2/2\lambda$ where λ is bounded between $\underline{\lambda}$ and $\bar{\lambda}$.⁷ This cost can be thought of as arising from an inherent aversion to work, or perhaps more realistically from the agent's need

⁶See Lipsky (1980), Brehm (1999), Perry and Hondeghem (2008) in political science and public administration, and Heckman, Smith, and Taber (1996), Besley and Ghatak (2006), and Prendergast (2008) in economics.

⁷ $\bar{\lambda} \approx .68466$ is a necessary and sufficient bound to keep the agent's first period effort well-behaved and less than 1 - it is characterized in the proof of Proposition 1. The lower bound $\underline{\lambda} \approx .23505$ is derived in the proof of Proposition 7 - intuitively, it ensures that the effect of the agent's beliefs on his effort is not trivial.

to take time away from other productive activities. The managerial problem that arises in our model is thus bureaucratic *slack*, i.e. the agent pursuing his “mandates with insufficient effort,” rather than bureaucratic *drift*, i.e. the agent pursuing policies that do not achieve principal’s goals (Bueno De Mesquita and Stephenson 2007). The parameter λ determines how willing the agent is to expend effort to achieve the organization’s goals, and thus captures the strength of the agent’s intrinsic motivation.

For simplicity we assume that the players place equal weight on payoffs in each period; this assumption best captures environments where goals must be achieved repeatedly over time, such as minimizing the level of a pollutant, reducing crime, or maximizing the output of a natural resource. However, our insights about how the potential for learning impacts initial choices clearly extend to settings where the second period is more important than the first, such as a trial or pilot project. Summarizing, the player’s payoffs given effort levels and outcomes over the two periods are,

$$\begin{array}{ll} y^1 + y^2 & (Principal) \\ \left(y^1 + \frac{(e^1)^2}{2\lambda}\right) + \left(y^2 + \frac{(e^2)^2}{2\lambda}\right) & (Agent) \end{array}$$

How Successes Occurs Intuitively, the success or failure of a policy depends both on whether it is fundamentally well-suited to achieve the intended goal, and on how well it is implemented. To capture these intuitions, we assume that “nature” initially draws a state $\omega = \{a, b\}$ determining which of the two policies is actually the “correct” one to achieve the goal. In both periods, the probability that the correct policy ($x^t = \omega$) succeeds is equal to the agent’s implementation effort e^t , while the incorrect policy ($x^t \neq \omega$) always fails. Thus, more effort results in a higher chance of success, but only when exerted on the correct policy. Our framework is standard in the economics literature on learning through experimentation with effort.⁸ Its key substantive property is that choosing well and exerting

⁸Discrete and continuous time versions include Bergemann and Hege (2005) and Keller, Rady, and Cripps (2005).

effort are *complementary* to achieving an organizational success, i.e., that both are necessary inputs.

A policy problem that might take this form is identifying the root cause of air pollution in a municipality; city bureaucrats may be uncertain about whether high observed pollution levels are predominantly the result of emissions from refineries or from automobiles. To successfully improve air quality, they must both correctly identify the “root cause,” and to expend effort effectively regulating the true source.⁹

Modeling Good Faith Disagreement To model good faith disagreement, we assume that the players have *heterogeneous prior beliefs* about the probability that $\omega = a$, i.e. that a is the correct policy. Specifically, each player i has their own prior belief θ_i that $\omega = a$. The principal initially believes that a is more likely to be correct ($\theta_1 \geq \frac{1}{2}$), while the agent believes that b is more likely to be correct ($\theta_2 \leq \frac{1}{2}$). They know each others’ prior beliefs, and thus fully understand the nature of their initial disagreement. In the preceding example, the principal could be a city manager who believes that the existing weight of evidence falls on refineries as the root cause of high pollution levels, while the agent is a city bureaucrat charged with implementing the policy who believes that automobiles are more likely to be the root cause. We will use $P_i(\cdot)$ to denote probabilities evaluated with respect to the prior of player i – so for example, $P_2(\omega = b) = 1 - \theta_2$ denotes the agent’s prior belief that b is the correct policy.

Learning and Experimentation To isolate the effects of learning and experimentation we compare two variants of the model. In the first variant, policy outcomes are never observed – thus, disagreement persists unchanged throughout the game. We refer to this variant as the *No Learning Benchmark*. It captures policymaking environments where outcomes cannot be reliably measured and/or circumstances are changing too rapidly for useful learning to

⁹This example roughly captures efforts by Los Angeles to control escalating smog levels in the 1960s (Krier 1977).

occur. In the second variant, both players publicly observe whether the first period policy succeeded or failed (i.e. the value of y^1) prior to making their decisions in the second period. This allows them to *learn* from the initial outcome before making their subsequent policy decisions.

How will this learning occur? This is straightforward to see by considering the players' posterior beliefs from Bayes' rule. Suppose that the principal initially chooses policy $x^1 = a$, and the agent implements it with effort e^1 . If a success is observed ($y^1 = 1$), then updating is simple – since only the correct policy can succeed, both the principal and the agent will infer and agree after success that a is definitely the correct policy.

What if a fails? Then the players could ascribe it *both* to a being incorrect, or to it simply having been implemented with insufficient effort. Specifically, each player i computes a posterior belief that a failed despite being correct equal to,

$$h(e^1, \theta_i) = \frac{\theta_i (1 - e^1)}{\theta_i (1 - e^1) + (1 - \theta_i)} < \theta_i. \quad (1)$$

This is her prior assessment $\theta_i (1 - e^1)$ of the probability that a was correct but still failed, divided by the unconditional prior probability of failure. This posterior belief is lower than the prior θ_i – that is, failure is always a negative signal. However, how much lower depends on how hard e^1 the agent worked to implement a . Intuitively, the harder the agent worked, the more the players will ascribe failure to the policy being fundamentally incorrect, rather than simply poorly implemented. Thus, greater effort by the agent makes failure *more informative* about the correct policy.

This intuitive property has an important implication for our analysis – it means that more implementation effort by the agent can both increase learning, and by implication reduce future disagreement. Returning to the pollution example, suppose that regulations are imposed on refineries, but the effects of this policy on pollution are ultimately observed to be minimal. This policy failure provides some indirect evidence that automobile emissions

may actually be the root cause of the high pollution. However, the indirect evidence is even stronger if the bureaucrat strongly enforced regulations on refineries than if he weakly enforced them.

We now analyze the two variants of the model.

3 The No Learning Benchmark

Without learning, choices and outcomes in the first period do not affect beliefs and incentives in the second. Consequently, the players consider each period in isolation. The agent is a necessary intermediary for policy implementation, and for the organization to succeed he must exert high effort. However, his effort only makes a difference if the principal chose the correct policy. Formally, his subjective expected payoff to exerting effort e^t to implement policy x^t is

$$P_2(\omega = x^t) \cdot e^t - \frac{(e^t)^2}{2\lambda}, \quad (2)$$

where $P_2(\omega = x^t)$ is his subjective prior assessment that the principal chose correctly. His optimal responses to the principal's policy decisions are consequently as follows.

Observation 1. *In a best-response, the agent exerts effort e^t to implement policy x^t equal to*

$$e^*(x^t) = \begin{cases} \lambda\theta_2 & \text{if } x^t = a \\ \lambda(1 - \theta_2) & \text{if } x^t = b \end{cases}.$$

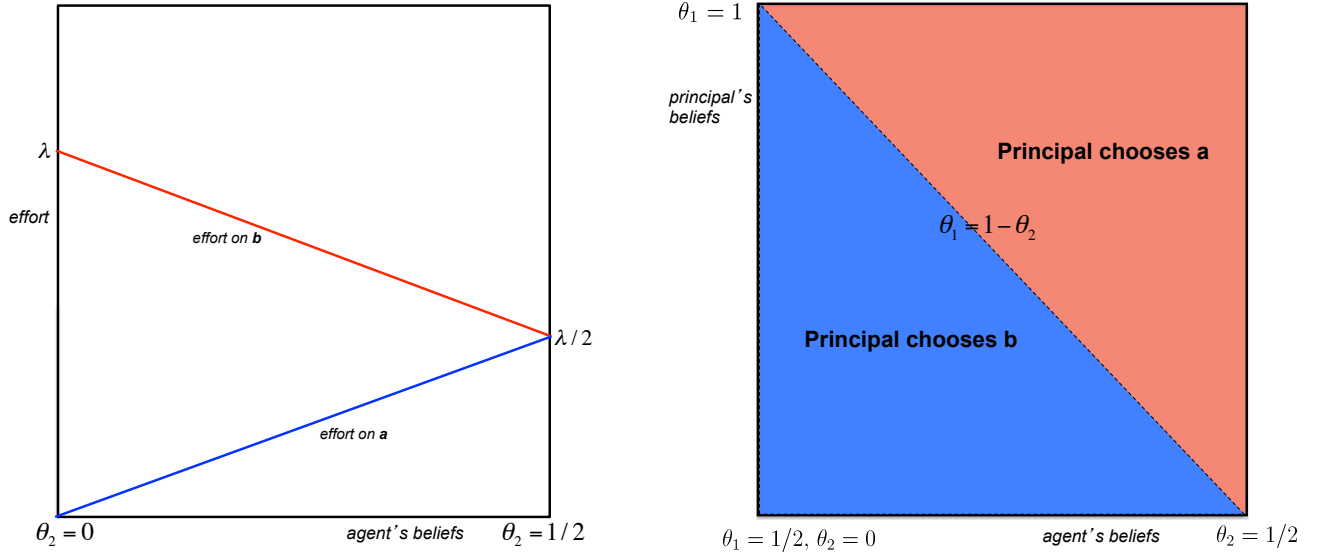
His subjective expected utility is

$$U_2^*(x^t) = \begin{cases} \frac{\lambda}{2}(\theta_2)^2 & \text{if } x^t = a \\ \frac{\lambda}{2}(1 - \theta_2)^2 & \text{if } x^t = b \end{cases}.$$

The left panel of Figure 1 depicts the agent's optimal effort on each policy, which is an increasing (linear) function of his belief that the principal chose correctly. The basic model captures two key features of management in political organizations, where agents are

intrinsically motivated and their superiors lack strong direct mechanisms of control. First, the agent's effort is increasing the strength of his intrinsic motivation, which is captured by the parameter λ (see also Besley and Ghatak (2006)). Second, to motivate their subordinates to work hard, leaders must make decisions that their subordinates *believe* will be effective in achieving the organization's goals.

Figure 1: Effort and Policy Choice in the No Learning Benchmark



Good faith disagreement thus creates a special kind of agency problem. If principal imposes the policy that she thinks is best, the agent will be less motivated, believing that his effort would be wasted on an ineffective policy. Alternatively, if the principal defers to the agent, then this will better motivate him – but from the principal's perspective, there are better than even odds that this extra motivation is useless. Either resolution entails a cost, either in terms of choosing well or in terms of motivating the agent.¹⁰ The following observation states that in the unique equilibrium, the principal resolves this dilemma by deferring to the agent if and only if his belief that *b* is the right policy is stronger than her own belief that *a* is the right policy.

¹⁰See also Van den Steen (2009) for a similar managerial dilemma in a firm setting.

Observation 2. *In the unique equilibrium of the No Learning Benchmark, the principal defers to the agent and chooses policy $x^* = b$ in both periods if and only if the agent’s belief that b is correct is stronger than her own belief that a is correct, i.e. $1 - \theta_2 \geq \theta_1$. Otherwise, she imposes policy $x^* = a$ in every period. Her subjective expected payoff in every period is,*

$$U_1^* = \begin{cases} \lambda \theta_1 \theta_2 & \text{if } x^* = a \\ \lambda (1 - \theta_1) (1 - \theta_2) & \text{if } x^* = b \end{cases}.$$

The right panel of Figure 1 depicts the unique equilibrium policy x^* as a function of the player’s beliefs (θ_1, θ_2) . Although the principal defers to agent when his beliefs in favor of b are sufficiently strong, her rationale is very different than in classic political agency models, where deference is driven by the desire to exploit an agent’s superior expertise. In our model, the agent has the same preferences as the principal, and the principal does not ascribe his difference in beliefs to superior expertise – she simply disagrees with him. When she defers, she does so because she has no other good option. She needs him to work, cannot force him to do, and concludes that more effort on an inferior policy is a better bet overall than less effort on a superior policy. The outcome is therefore inefficient in the sense that the principal would always prefer a scenario where she can select policy a and the agent believes she chose correctly (i.e. $\theta_1 \cdot 1 > \max \{\theta_1 \theta_2, (1 - \theta_1) (1 - \theta_2)\}$ for all θ_2).

4 The Game with Learning

We now analyze the game when policy outcomes can be observed; the players can thus learn more about which policy is correct, as well as experiment with policies. This variant of the model captures political organizations where either the outcomes of interest are straightforward to observe, or where reliable “performance measurement” of the organization’s outputs has been introduced; we later return to this interpretation when we discuss empirical implications.

In the literature on decisionmaking in public organizations, the rationales given for measuring performance have typically been to either help public managers *learn* in order to make better policy decisions, or to *reward and sanction* them for their behavior. The former draws inspiration from theories of organizational learning (Moynihan 2008, Ch. 9), while the latter draws on classical principal-agent theory in economics (Kettl 2005, Ch. 3). Our focus, however, is new; we ask how performance measurement and learning affect the *management* of political organizations with intrinsically motivated subordinates and good faith disagreement. How will the ability to observe policy outcomes change the incentives of the agent? When will the principal choose to experiment, and why? And how can experimentation be exploited to reduce good-faith disagreement?

Formally, a pure strategy for the principal in the two-period game with learning consists of a first period policy x^1 , and a second period policy $x^2(x^1, y^1)$ for each first period policy x^1 and outcome y^1 . In theory the space of possible strategies is large – the principal could always switch policies after the first period regardless of the outcome, or keep the initial policy only if it fails. However, we can rule these strategies because success persuades both players that the initial policy was correct. Thus, the principal’s strategy must therefore satisfy the following conditions.

Observation 3. *In a pure strategy equilibrium, for each possible policy x^1 the principal either*

- ***rigidly implements*** x^1 , *i.e., chooses it in both periods regardless of the first period outcome* ($x^2(x^1, y^1 = 1) = x^2(x^1, y^1 = 0) = x^1$)
- ***experiments with*** x^1 , *i.e. chooses it again in second period if it succeeds* ($x^2(x^1, y^1 = 1) = x^1$) *and switches to the alternative* ($x^2(x^1, y^1 = 0) \neq x^1$) *if it fails.*

In the remainder of this section we characterize how the agent will respond to these two management strategies, i.e. experimenting with vs. rigidly implementing policies. In the subsequent section we complete the equilibrium analysis by characterizing how the principal

will make her policy choices given her own initial beliefs about the right policy and the agent's incentives.

To clarify the exposition, the mathematical details of the characterization are relegated to online appendices. Appendix A provides a general equilibrium characterization, and Appendix B proves the main propositions; Appendix C and an accompanying Mathematica file verify less-critical steps. In the general equilibrium characterization the players may use mixed strategies, but we show in the Online Appendix that the principal is always best off in a pure strategy equilibrium. Since our focus is on what the principal can achieve by managing the agent, we henceforth select this equilibrium for the remainder of the analysis.

Rigid Implementation and Effort

A principal who rigidly implements policies by definition ignores outcomes when making her decisions. However, this does not mean that the ability to observe outcomes has no effect. The reason is that the *agent* will still use it to try to learn which policy is correct, in order to avoid wasting his costly effort on an ineffective policy.

To understand how the ability to observe outcomes affects the agent's incentives, we analyze his optimization problem by working backwards. If the principal's strategy is to rigidly implement policy a (i.e. $x^1 = a$, $x^2(a, 1) = x^2(a, 0) = a$), then in the second and final period the agent anticipates that he will exert effort based on his *posterior* belief a is the right policy after observing the initial outcome, i.e. $e^2 = \lambda P_2(\omega = a | y^1)$. Using this, it is straightforward to show that his two-period expected utility as a function of his *initial* first period effort e^1 to implement a is:

$$\underbrace{\left(\theta_2 e^1 - \frac{(e^1)^2}{2\lambda} \right)}_{\text{first period payoff}} + \underbrace{\frac{\lambda}{2} \theta_2 (e^1 + (1 - e^1) h(e^1, \theta_2))}_{\text{second period payoff}}. \quad (3)$$

Details of the derivation are presented in online Appendix A,¹¹ but the expression can be

¹¹By symmetry, the agent's two-period expected utility when the principal rigidly imple-

intuitively understood as follows. The first term is the agent’s first period expected payoff as a function of his implementation effort e^1 , and is identical to his per-period payoff in the No Learning benchmark. The second term is his expected *future* payoff as a function of his initial effort, and it is increasing in effort for the following reason. More initial effort results in a more informative first period outcome, and thus better learning by the agent. This benefits him in the second period by allowing him to better calibrate his effort; that is, to work hard if the policy is the right one, and avoid wasting effort if it is not. Formally, the expression $e^1 + (1 - e^1) h(e^1, \theta_2)$ is the agent’s average posterior belief that a is correct *when it is actually correct*, and this effectively captures how accurate the agent’s future beliefs will be after observing the initial outcome.

The presence of a “learning premium” to early effort when policy outcomes are observable generates the following result.

Proposition 1. *When the principal plans to rigidly implement some policy $x \in \{a, b\}$, observing outcomes increases the agent’s first period effort, and the probability of success in both periods, relative to the No Learning benchmark.*

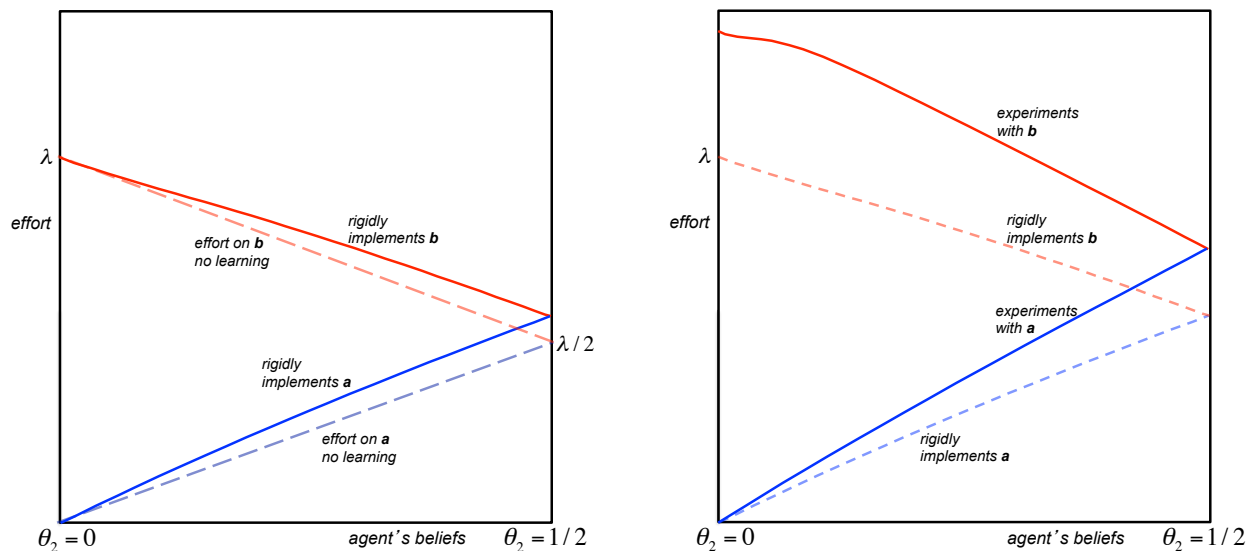
The left panel of Figure 2 illustrates the agent’s first period effort on each policy in both the No Learning benchmark and the game with learning and rigid implementation.

The managerial implications of Proposition 1 are surprising. When subordinates are intrinsically motivated, the ability to measure and observe organizational outcomes both increases the efficiency of agents’ efforts, and better motivates them overall, *even when superiors do not even use the information in their decisions*. The reason is simple: intrinsically motivated subordinates still value the information for allocating their time and effort, and they can generate more reliable information by working harder. Our results suggest that recent critiques of performance measurement in public organizations based on the observation that policymakers rarely use performance data are missing an important dimension (e.g. Pollitt 2006). Indeed, consistent with this result, an analysis by the U.S. Government

ments policy b is the above expression with $P_2(\omega = b) = 1 - \theta_2$ substituted in for θ_2

Accountability Office (GAO) found that federal agencies value performance measurement to improve the efficiency of their internal decisionmaking (U.S. Government Accountability Office 2005) even when Congress ignores the information in its policy and budget-allocation decisions.

Figure 2: Effect of Observing Outcomes and Experimenting on the Agent's Effort



Experimentation and Effort

A principal who experiments with policies ties her second period policy choice to the outcome of her first period policy; formally, her strategy is to stick with the initial policy if and only if it succeeds. This means that the data contained in policy outcomes directly informs her choices.

A less obvious feature of experimentation, however, is that it indirectly empowers the agent to influence the principal's policy choices. By working hard he can increase the chance of a success, and therefore the chance that the principal will stick with the initial policy. By "shirking" or sabotaging the initial policy, he can decrease the chance of success, and thus also the chance that the principal will switch to the alternative. A priori, however, the effect of this empowerment isn't obvious; it could magnify the agent's willingness to work

by increasing the cost of failing with the right policy (since the principal will then switch to the wrong one), *or* it could give the agent to sabotage a policy with which he disagrees by shirking to induce failure.¹²

To clarify which of these incentives dominates, we now formally analyze the agent's problem when the principal experiments with policy a (i.e. $x^1 = a$, $x^2(a, 1) = a$ and $x^2(a, 0) = b$). The agent's two-period expected utility is:¹³

$$\underbrace{\left(-\frac{(e^1)^2}{2\lambda} + \theta_2 e^1\right)}_{\text{first period payoff}} + \underbrace{\frac{\lambda}{2} (1 - \theta_2) (1 - h(e^1, \theta_2))}_{\text{learning term}} + \underbrace{\frac{\lambda}{2} \theta_2 e^1}_{\text{policy influence term}}. \quad (4)$$

second period payoff

Like the utility function with rigid implementation, the expression is divided into a first and second period term. Moreover, the first period terms are identical. However, because of experimentation the second period term now has two subterms: a *learning term* representing how the agent's initial effort affects his future utility directly through his own learning, and a *policy influence* representing how his initial effort affects his future utility indirectly through the principal's policy choices.

The learning term has properties that are similar to the agent's expected second period utility when the principal rigidly implements a . The key expression in this term is $1 - h(e^1, \theta_2)$, which is the expected value of the agent's posterior belief that a is the wrong policy when it is actually the wrong policy. As before, this captures how accurate the agent's future beliefs will be after observing the first period outcome, and it is increasing in his initial effort e^1 .¹⁴

¹²Sabotage here differs from Brehm (1999), who defines it as “devoting time” to undermine the principal's goals.

¹³Details of the derivation are again in online Appendix A, and again by symmetry the agent's two-period expected utility when the principal experiments with policy b is the above expression with $1 - \theta_2$ substituted in.

¹⁴A useful property in our simplified framework is that the derivatives of the second period payoff in (3) and the learning term in (4) are equal. Intuitively, this means that the agent's incentive to exert effort for his own learning is unaffected by whether the principal rigidly

The policy influence term $\frac{\lambda}{2}\theta_2e^1$, however, only appears when the principal experiments. Our key observation is that this term is always *increasing in effort*. Intuitively, this means that from the agent’s perspective, working harder initially always has a beneficial effect on the principal’s future choices. Experimenting thus never gives the agent an incentive to “sabotage”; it always incentivizes him to work harder. Surprisingly, this is true even when the agent is asked to implement a policy experiment he initially believes to be wrong. This generates the following result.

Proposition 2. *Regardless of the agent’s initial beliefs, a principal who experiments with a policy $x^1 \in \{a, b\}$ elicits more first period effort than a principal who rigidly implements that same policy.*

The right panel of Figure 2 compares the agent’s first period effort on each policy when the principal experiments with vs. rigidly implements it.

Why does experimentation incentivize the agent to work harder, even when he is working on a policy he believes is wrong? The reason is that his conflict with the principal stems from a good faith disagreement about which policy is better for achieving their goals, rather than an intrinsic preference for any one policy. Consider the agent’s incentives when asked to experimentally implement a policy that he believes to be incorrect. In theory, he could reduce his effort to increase the probability of failure and a policy switch. But if he is right about the policy being faulty, then it will fail anyway without his “sabotage.” Alternatively, if he is wrong (a possibility that he entertains), then reducing his effort will raise the chance of an accidental failure, which will induce the principal to abandon a policy that is good for both of them. The main effect of experimentation on the agent’s incentives is thus to raise the stakes of “getting it right” in the first period, which causes him to work harder.¹⁵ Proposition 2 thus demonstrates that there is a benefit to experimentation in political organizations that implements or experiments with policies.

¹⁵The fact that experimentation motivates the agent is what can generate multiple equilibria - experimenting produces more effort, which makes failure more informative, and the principal more willing to experiment.

is independent of “searching” for the best policy (Callander 2011) – it can better motivate the subordinates who are actually tasked with implementing policy choices.

5 Experimentation, Underexperimentation, and Persuasion

We now complete the equilibrium analysis by describing the principal’s equilibrium policy choices given the previously analyzed incentives. To simplify the presentation, we first visually present these choices. We then discuss four properties of these choices in-depth and provide results.

Figure 3 illustrates the principal’s equilibrium policy choices as a function of the players’ initial beliefs (holding fixed the agent’s intrinsic motivation λ).¹⁶ Recall that the agent initially favors policy b ($\theta_2 \leq \frac{1}{2}$), while the principal initially favors policy a ($\theta_1 \geq \frac{1}{2}$). The color of each region indicates the principal’s first period policy choice – in the red regions the principal imposes policy $x^1 = a$, while the blue regions she defers to the agent by choosing policy $x^1 = b$. The darkness of the shading indicates whether the principal experiments with or rigidly implements the first period policy – in the darkly shaded regions she rigidly implements, while in the lightly shaded regions she experiments.

Intrinsically motivated subordinates encourage experimentation

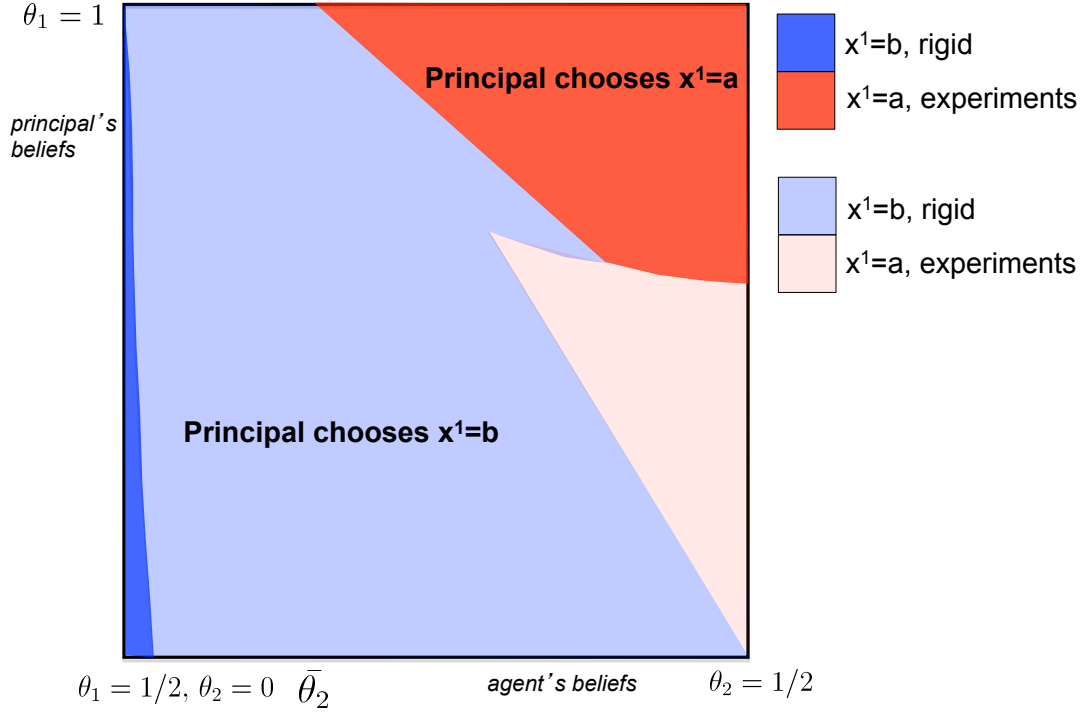
Our first result from the equilibrium analysis is that the principal is more inclined to experiment when the agent is more intrinsically motivated.

Proposition 3. *In equilibrium, the principal experiments for a greater set of initial beliefs (θ_1, θ_2) when the agent is more intrinsically motivated.*

Figure 4 shows how the region of belief space where the principal experiments expands with a more motivated (higher λ) agent; the identity of the first period policy is suppressed for

¹⁶The value $\lambda \approx .653$ is used for this and most subsequent figures.

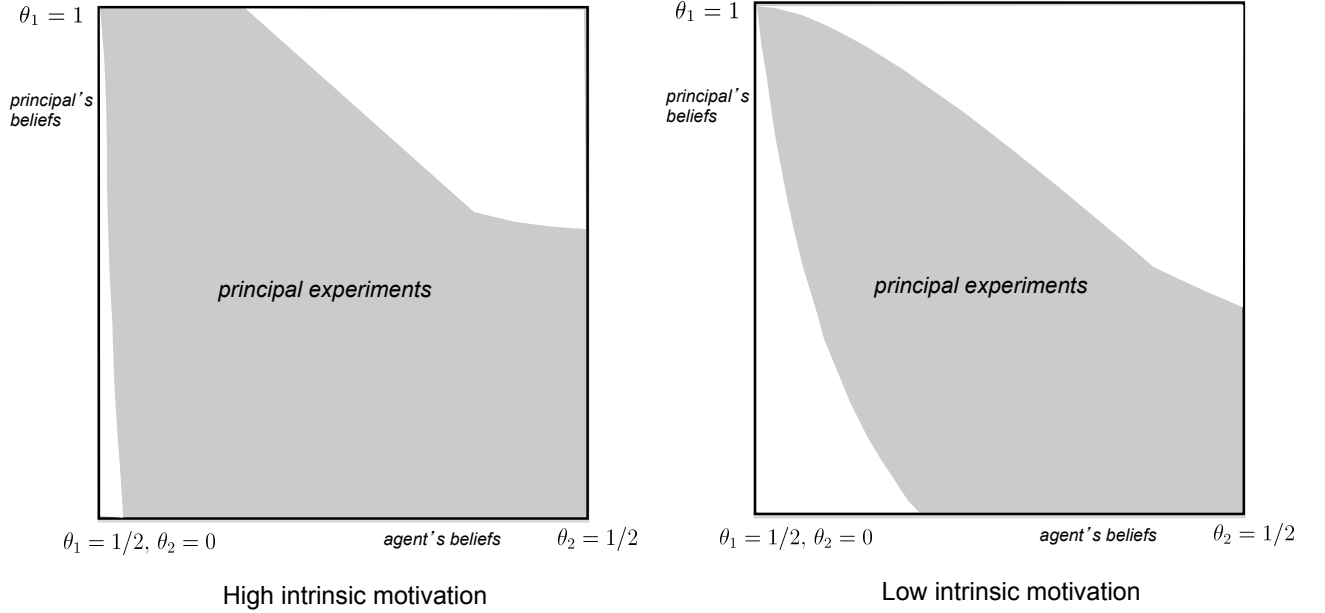
Figure 3: Principal's Equilibrium Policy Choices



clarity. To understand why a more intrinsically-motivated agent encourages experimentation requires developing some intuition for what drives the principal's decision to experiment. For the principal to experiment with a policy in equilibrium, she must not only be willing to try it out – she must also be willing to *abandon it if it fails*. Otherwise, the agent would expect her to ignore failure and rigidly continue with the policy into the second period. What determines the principal's willingness to abandon her initial policy after failure? Formally, she will do so for policy a (i.e. $x^1 = a$, $x^2(a, 1) = a, x^2(a, 0) = b$) if and only if the inequality $h(e^1, \theta_1) < 1 - h(e^1, \theta_2)$ holds. Intuitively, the condition is that failure pushes the principal (posterior) belief $h(e^1, \theta_1)$ that a is correct below the agent's posterior belief $1 - h(e^1, \theta_2)$ that b is correct.

Intuitively, two factors determine whether this will occur: (1) the strength of both players' initial beliefs in favor of a , and (2) the agent's effort to implement it. The latter factor is crucial because the informativeness of failure is *endogenous* to the agent's effort: both players find failure to be relatively more persuasive evidence that a is in wrong when the

Figure 4: Effect of Intrinsic Motivation on Principal's Experimentation



agent worked harder to implement it.¹⁷ Since the agent's initial effort e^1 on any policy is higher when he is more intrinsically motivated (i.e. has higher λ), the principal is more willing to abandon the initial policy after failure, and thus more willing to experiment.

An interesting implication of Proposition 3 is that among organizations where outcomes can be effectively measured, those with more intrinsically motivated subordinates should also exhibit a greater propensity toward experimentation. This comparative static accords with the existing hypothesis that public organizations whose employees have a “shared vision” should be learning and adaptive (Moynihan and Landuyt 2009), but the intuition in our analysis is different. More intrinsically motivated subordinates implement policies more effectively; consequently, the outcomes of those policies are more informative about their underlying quality.

¹⁷Formally, l.h.s. of the necessary inequality is decreasing in e^1 while the r.h.s. is increasing.

The principal underexperiments (unless she has help)

In Proposition 2, we showed that a principal who experiments reaps a motivational benefit from her agent. However, the analysis above demonstrates that the principal's ability to capitalize on this benefit is constrained by her own beliefs. Experimenting with a policy requires her to abandon it if it fails. However, she may be unwilling to do so if her initial beliefs in its favor are too strong.

This constraint on experimentation creates an interesting commitment problem in the model. Sometimes, a principal with relatively strong beliefs in favor of a policy would prefer to experiment with it in order to better motivate the agent, but after actually observing failure will want to renege on the experiment and persist with the initial policy. Anticipating this, the agent will work less and experimentation will collapse in equilibrium. Formally, we have the following result.

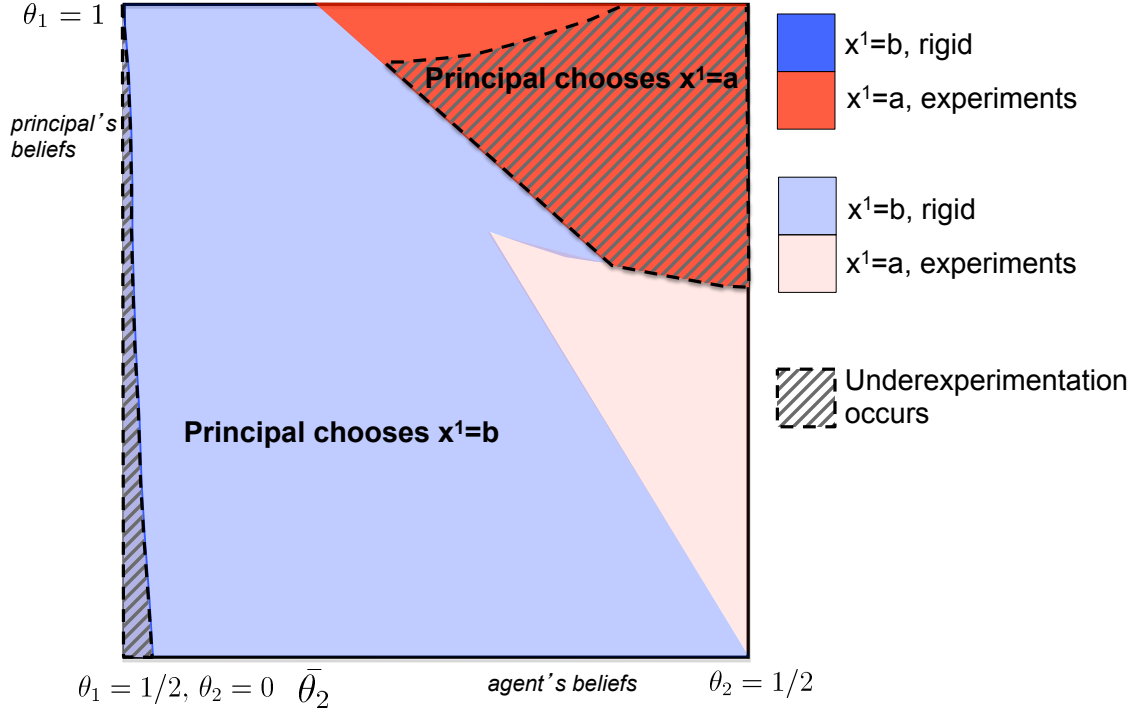
Proposition 4. *For some beliefs (θ_1, θ_2) the principal rigidly implements some policy $x^{*1} \in \{a, b\}$ even though experimenting with it would yield higher ex-ante expected utility. Conversely, the principal never experiments with a policy when rigidly implementing it would be better ex-ante.*

The cross-hatched regions in Figure 5 depict the subset of the belief-space where the principal *underexperiments*. In the upper right region, the principal would be best off experimenting with a , but her initial beliefs in its favor are too strong and she rigidly implements it. In the lower left region, the principal would be best off initially deferring to the agent by experimenting with b , but the agent's beliefs in b 's favor are so strong that the principal will again rigidly implement it.¹⁸

Because underexperimentation results from the principal's power to block policy change *after* a failure occurs, Proposition 4 suggests that there can be managerial rationale for

¹⁸This commitment problem is also the reason for the jagged region in the upper-right quadrant – here it is optimal for the principal to experiment with a , but her beliefs in its favor are so strong that she instead experiments with b .

Figure 5: Equilibrium Underexperimentation



institutional arrangements that help organizational decisionmakers commit ex-ante to make such changes. Below, we explore two common arrangements in political organizations that can achieve this end.

Creating Costs to Rigidity One way underexperimentation can be ameliorated is by creating exogenous *costs* to rigidly maintaining failed policies.

Proposition 5. *For any beliefs (θ_1, θ_2) s.t. the principal underexperiments, there is a cost c of maintaining policy after failure that makes her optimal policy experiment an equilibrium.*

How might political organizations generate such costs in practice? One way is to enact *trigger mechanisms* that make a “pre-negotiated commitment... specifying what actions will be taken if monitoring information shows x or y ” (Nie and Schultz 2012). Because trigger mechanisms can usually be overridden in most political environments, their practical effect is to create costs to policy rigidity rather than eliminate it entirely. Such mechanisms

were recently proposed in debates over health care reform¹⁹ and national security policy (Gates 2014, p375), and are commonly justified on accountability grounds. Our results, however, also suggest a managerial rationale for triggers; public managers may themselves prefer to have such mechanisms in place to ensure their own commitment to experimentation and better motivate their subordinates.

A second arrangement that can create costs to rigidity is requiring written or oral statements from managers that detail past policy outcomes and explicit plans for improvement; such procedures effectively magnify the personal and professional embarrassment of maintaining failing policies. A recent example is the Compstat program in New York City. Although the program is primarily known for analyzing high-frequency crime data to spot trends, its central component is actually a highly formalized weekly meeting where “precinct commanders appear before several of the department’s top brass to report on crime problems in their precincts and what they are doing about them” (Weisburd, Mastrofski, McNally, Greenspan, and Willis 2003). Program designer Jack Maple describes the purpose of meetings as follows: “nobody ever got in trouble because crime numbers on their watch went up... trouble arose only if the commander’s didn’t know why the numbers were up or didn’t have a plan to address the problem” (Maple and Mitchell 2010).

Appointing Moderate Middle Managers An alternative mechanism that political organizations can use to enhance their commitment to experimentation is to simply appoint individuals to make policy decisions whose intrinsic beliefs are amenable to it. In the following proposition, we show that the principal is sometimes better off delegating policy decisions to an individual whose beliefs in favor of the optimal policy are more *moderate* than her own, and is therefore more willing to abandon it after failure.

Proposition 6. *Suppose that a principal with beliefs θ_1 could appoint a player with beliefs $\hat{\theta}_1$ to make policy decisions in her place. If appointing herself is not optimal, then any optimal*

¹⁹http://www.slate.com/articles/news_and_politics/prescriptions/2009/10/public_option_lite.html

appointee $\hat{\theta}_1^*$ believes less strongly in the resulting policy x^{*1} than the principal does.

Because most political organizations are multi-tier hierarchies, elected political principles routinely select the intermediaries who will represent them in the bureaucracy, such as bureau chiefs and cabinet secretaries. Proposition 6 suggests that the best appointees for highly ideological political principles may not be ideological “clones,” because they will demotivate intrinsically-motivated bureaucrats with their unwillingness to respond to negative policy feedback. Instead, the aims of highly ideological principals may actually be better achieved by appointing *moderates*, whose publicly-known beliefs make a commitment to experimentation and “data-driven management” credible.²⁰ This insight is consistent with statistical evidence that Presidents make ideologically divergent appointments to federal agencies (Bertelli and Grose 2011).

The Reagan administration’s experience managing the Environmental Protection Agency furnishes an example of this strategy. Within the EPA, intrinsically motivated bureaucrats must make regulatory policy under considerable technical uncertainty about the consequence of their actions. President Reagan’s first appointee to run the EPA was Anne Gorsuch, a strong conservative ideologue who “shared the president’s environmental policy agenda” and “firm convictions about the negative impact that environmental regulations had on economic growth” (Golden 2000, p. 120). Her brief tenure was marred by a crisis of morale among career bureaucrats, who responded not with “deliberate foot-dragging or sabotage” but a severe waning of “commitment,” “enthusiasm,” and voluntarily effort. After resigning in controversy, Gorsuch was replaced by William Ruckelshaus, also a committed conservative but with an “open” and “data driven” management style that permitted a more responsive and experimental approach (Dobel 1992, p. 251). Despite the opposition of environmental groups and his strong commitment to right-leaning policies like cost-benefit analysis, “Ruckelshaus was regarded as a savior” by EPA bureaucrats who “wanted to work for him.” The

²⁰For a contrasting perspective on the optimal selection of middle managers see Van den Steen (2005).

stark contrast in bureaucratic motivation under the two appointees led Golden (2000) to speculate that “there may be two roads to compliance” in selecting Presidential appointees. President Reagan continued to appoint moderate EPA heads for the remainder of his tenure.

Experimentation is biased toward the agent

In the preceding sections we focused on the principal’s decision about whether to experiment, but have not yet analyzed *how* she will experiment. We now turn to this question. Specifically, when will the principal experiment by forcing the agent to initially try out the policy he believes in (*a*)? When will she experiment by *deferring* to the agent and giving the policy he believes in (*b*) a chance to succeed?

In the No Learning benchmark in Section 3, the principal defers to the agent as a sort of second-best coping strategy given their irresolvable disagreement. The principal needs him to work hard in order to achieve a success, and when his beliefs are stronger than her own she decides that more effort on an inferior policy is better bet than less effort on a superior policy. In the main model, however, there is also the potential for additional learning and persuasion; the principal could be persuaded by the failure of *a* that *b* is the superior policy, and similarly the agent could be persuaded by the failure of *b* that *a* is the superior policy. Crucially, more effort by the agent on the initial policy actually facilitates this process; it makes the initial outcome more informative about the policy’s true underlying quality, resulting in better future decisions by both the principal and the agent (Section 4) and less future disagreement.

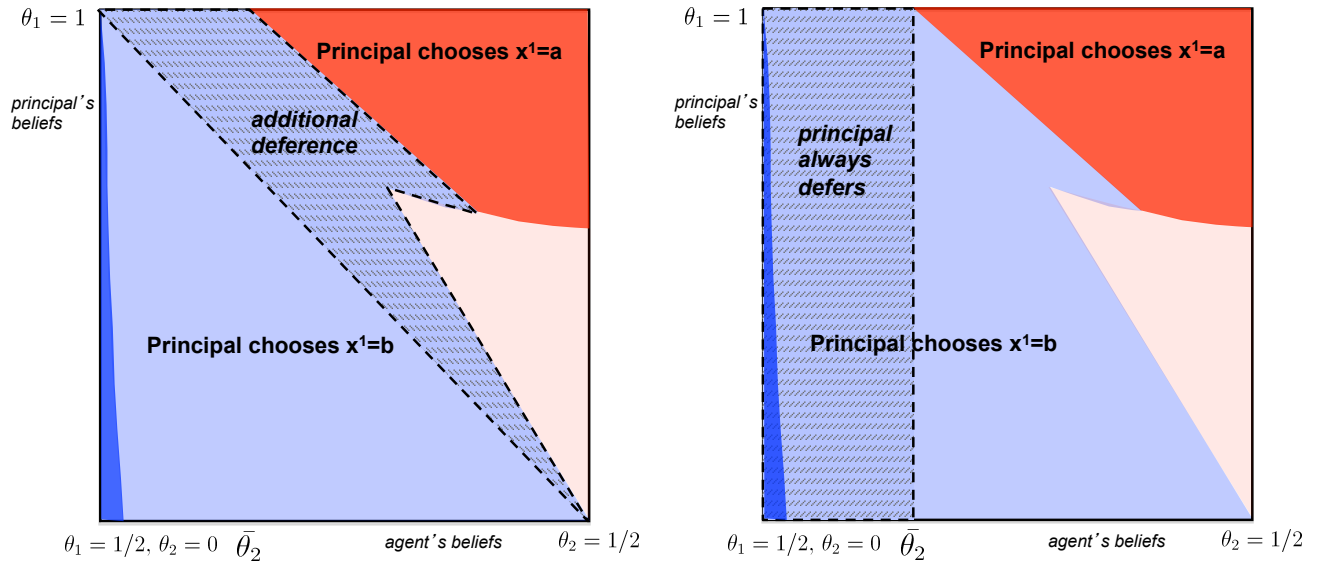
Since the principal can elicit greater initial effort from the agent by simply deferring to him, these informational benefits create an extra *bias* in the principal’s initial policy decisions toward the policy initially desired by the agent. This yields the following result.

Proposition 7. *When the principal experiments, she defers to the agent in the first period for a greater set of initial beliefs than in the No Learning benchmark. When she does not experiment, her policy decisions are unchanged.*

The bias in the principal's initial decisions relative to the No Learning benchmark can be seen in the left panel of Figure 6, which replicates the principal's equilibrium policy choices and demarcates the region of additional deference as compared to the No Learning benchmark in Figure 1. The deference region is strictly larger in the game with learning than in the No Learning benchmark. It also expands to include cases where the principal is essentially certain that policy a is right policy; we soon return to this striking property.

Substantively, Proposition 7 implies that the beliefs and opinions of subordinates will exert greater influence on policy decisions – sometimes significantly greater influence – in organizations where learning and experimentation can occur. The reason is that the benefits of motivated subordinates are greater when this will also facilitate organizational learning.

Figure 6: Bias and Deference in Equilibrium



An additional implication of Proposition 7 is that *introducing* the ability to learn and experiment to political organizations – through new technology or the development of formal procedures to measure outcomes – should result in greater deference to subordinates. Interestingly, this combination of performance measurement and increased deference closely resembles the core principles of the National Performance Review, a high profile management initiative during the Clinton administration that sought to implement assorted “New

Public Management” ideas in U.S. federal agencies. Aberbach (2000) describe how the NPR combined performance measurement with the directive that “decision-making power should be decentralized, giving lower-level employees more authority to make decisions.” We later return to this application in our discussion of empirical implications.

Principals sometimes defer to persuade

Finally, we consider the question of exactly how much the principal’s initial policy decision will be biased toward the agent when learning can occur.

In the No Learning benchmark, the principal will always impose her desired policy when she is sufficiently sure that it is the right one; there is no point in eliciting high effort on the wrong policy. This intuitive property is shared by classical principal-agent models in which the incentive to defer comes from the desire to exploit the agent’s expertise; if the principal is already sure she knows the right policy, then there is no expertise to exploit and no reason to defer.

In our model with learning, however, something quite different happens. If the agent’s beliefs are sufficiently extreme, *and* her intrinsic motivation is sufficiently strong, then the principal will always defer to him initially by selecting policy b – *even when she is sure that he is wrong*.

Proposition 8. *When the agent is both sufficiently motivated ($\lambda > \hat{\lambda} \in (\underline{\lambda}, \bar{\lambda})$) and believes sufficiently strongly in b (i.e. θ_2 close to 0), then the principal **always defers** initially and selects policy $x^1 = b$ regardless of her own beliefs.*

The right panel of Figure 6 demarcates the region of the belief space where the principal always to the agent and selects policy b even when she is effectively certain that a is right. The agent’s intrinsic motivation λ is fixed at some $\lambda > \hat{\lambda}$ where the effect occurs.

If the bias in the principal’s policy decisions is due to the informational benefits of first period effort, why does a principal who is already sure of the right policy defer to realize these benefits? The reason is that the target of the information produced by the

additional effort is the *agent himself*. Specifically, the principal defers in order to *persuade* the agent that his initial beliefs are wrong, and to instead work hard on policy a in the future. This works because allowing the agent an opportunity to *prove* that b is the right policy strongly motivates him, making failure all the more persuasive when it (from the principal’s perspective) inevitably occurs. The alternative – forcing the agent to implement policy a immediately – is less effective for persuading the agent to support a in the future because he will exert low effort, so a failure will likely occur and (slightly) reinforce his initial beliefs.²¹ Our model thus captures a realistic feature of political organizations that is absent from previous political agency models; the incentive to defer can arise not only from the desire to exploit a subordinate’s expertise or better motivate them, but also to *persuade* them and build future consensus that the organization’s strategy is the right one.

Proposition 8 illustrates the benefits of a relatively “hands off” approach to managing political organizations in which learning through experimentation can occur. Such benefits are frequently touted by public managers. For example, successful former NYPD William Bratton characterizes his management style as one in which he “pick[ed] good people and let them do their jobs,” “gave his people room to maneuver” and “turned responsibility back onto the worker.” (Bratton 1998, p. 127). Similarly, Heclo (1977) recounts how an assistant secretary rationalized his failure to discipline an actively insubordinate bureaucrat in the following manner: “the agency won’t get good ideas if you crush the guys running around advocating them.” Our model provides a new microfoundation for this managerial strategy, and generates additional implications about when it will be most effective. For example, deferring to the agent to “teach” him that he is wrong cannot be effective unless the agent is already sufficiently motivated intrinsically to achieve the organization’s goals (i.e. has high λ) – otherwise, he will not be especially motivated by deference, and so not especially

²¹Mathematically, when $\omega = a$ and the principal rigidly implements a , the agent’s expected future belief that a is right $e^1 + (1 - e^1) h(e^1, \theta_2)$, which will be low since θ_2 and consequently e^1 is low. If $\omega = a$ but the principal experiments with b , then the agent’s expected future belief that a is right is $1 - h(e^1, 1 - \theta_2)$, which will be (relatively) high since $1 - \theta_2$ and consequently e^1 are high.

persuaded to revisit his beliefs by the failure that results.

6 Empirical Implications

In the study of principal-agent relationships in political environments, the idea that policy-makers are *uncertain* about how to achieve their policy goals has long been central. However, the previous literature has largely focused on only one aspect of this uncertainty – that an agent may have superior information that he wishes to exploit about the consequences of different policies.

In this paper, we consider a different, but also empirically relevant, aspect of this uncertainty – that principals and agents might *openly disagree* about how to achieve goals that they share. To explore the implications of this disagreement, we develop two models of policy choice and implementation in a political organization – one in which policy outcomes cannot be observed, and another in which they can be observed and learned from. Our main insights are threefold. First, we show that being able to observe and learn from policy outcomes improves the efficiency and motivation of agents, even when principals themselves ignore the information. Second, we identify motivational benefits to experimentation, and demonstrate that principals sometimes benefit from institutions that encourage, force, or select them to be responsive to negative policy feedback. Finally, we demonstrate that there can sometimes be an extraordinary degree of deference to the beliefs of agents driven by the desire to persuade.

The results of our analysis generate a variety of empirical implications for the study of many types of political organizations, but especially of federal bureaucracies. Existing research in public administration strongly supports the proposition that federal government employees are by-and-large “intrinsically motivated,” like the agents in our model (Perry and Hondeghem 2008). Moreover, frequent disagreements about how to achieve agency goals between political appointees and career bureaucrats are extremely well documented by

case studies of federal management.²² Finally, since the early 1990s the federal government has undertaken over two decades of intensive reform efforts seeking to bring *performance measurement* to federal agencies; that is, to change federal agencies from organizations where outcomes cannot be observed and learned from, into ones where they can. While the rationale and form of these efforts have often shifted, the emphasis on measuring performance has remained constant; the Government Performance and Results Act (1993) sought to generate such data to improve Congressional budgeting and decisionmaking (Long and Franklin 2004); former Vice President Gore’s “National Performance Review” of federal agencies sought to improve performance measurement as part of an ambitious package of managerial reforms to enhance the efficiency of government (Thompson and Riccucci 1998); and the Bush-era Program Assessment Rating Tool (PART) sought to both promote performance data use and measure agencies on a common scale (Lewis 2008).

The federal government has collected voluminous data tracking the efficacy and consequences of these reform efforts, which offers the opportunity to test many of the predictions of our models. Since 2002 the U.S. Office of Personal Management (OPM) has administered annual or biennial surveys of federal employees across agencies and ranks that can be used to track key variables in our model; specifically effort, intrinsic motivation, and managerial styles.²³ The U.S. Government Accountability Office has also conducted three cross-agency surveys of federal managers since 2003 measuring both whether federal agencies have introduced useful performance data, and whether that performance data is being used to guide agency decisionmaking.²⁴ By combining these data, we can paint a picture both across agencies, and across time, about when reliable performance was introduced, and the consequences of doing so for employee effort and managerial practices.

One set of predictions that we can test for is the “basket” of organizational changes that

²²See for example Hecla (1977) and Golden (2000).

²³Federal Employee Viewpoint Survey (FEVS); 2013 summary available at http://www.fedview.opm.gov/2013files/2013_Governmentwide_Management_Report.PDF.

²⁴Summary of 2013 survey is available at <http://www.gao.gov/products/GAO-13-519SP>

we predict when moving from the variant without learning to the variant with learning: effort, experimentation, and deference to subordinates should all increase. This transition can be measured in the data by the introduction of effective performance measures to an agency. We can also test a second set of predictions about patterns specific to organizations where outcomes can be measured: (1) there should be greater experimentation in agencies where subordinates are more intrinsically motivated, and (2) subordinates should be more motivated in agencies where managers actually use performance data in their decisionmaking, i.e. experiment. Preliminary evidence for this latter prediction already exists from the Clinton-era NPR reforms, which explicitly combined performance measurement with a directive to “give lower-level employees more authority to make decisions,” i.e. defer (Aberbach 2000). An important component of these reforms was the introduction of “reinvention labs,” where agencies could receive a special designation and expedited rule waivers for experimental projects spearheaded by low level and front-line bureaucrats. A 1996 GAO study of the reinvention labs found evidence that they had significantly improved agency effort, morale, and performance (U.S. Government Accountability Office 1996).

Finally, our model, combined with these data, can be used to generate new testable predictions about Presidential appointments of federal agency heads, a question that has received extensive attention in the literature. Recent empirical work finds that Presidents do not appoint ideological “clones” to run executive departments (Bertelli and Grose 2011). Our model suggests one reason why. In the model with learning, we find that a highly-ideological principal sometimes benefits from delegating her decisionmaking authority to a more moderate individual whose commitment to experimentation and “data-driven” management is more credible than her own. This generates the empirically testable implication that Presidents should choose less ideological appointees to run agencies that can “learn by doing” with effective performance measurement, because ideologues will demotivate rank-and-file bureaucrats; these agencies can be identified using the aforementioned GAO survey data. This finding is somewhat novel in the Presidential appointments literature, in that it predicts Presidential

appointments as a function of *internal* managerial characteristics of the target agency itself. Previous work on appointments behavior has almost exclusively considered only external political determinants, such as Congressional-Presidential bargaining (McCarty 2004) and the behavior of outside interest groups (Bertelli and Feldmann 2007, Gailmard and Patty 2013). A rare exception is Lewis (2008), who uses the “complexity” of an agency’s tasks as one predictor of Presidential politicization.

More generally, the idea that “management suffers... when government is run by a transient group of strangers” has been familiar since Heclo’s (1977) seminal work on the management of federal agencies. However, the mechanisms through which this occurs are not well understood. Our model provides a possible microfoundation – agency performance suffers when appointees and bureaucrats have different beliefs about how to achieve agency goals, and strategically choose to “go into hiding” or sabotage each other, instead of “educating” each other “irrespective of their politics” (Starobin 1995). Building a richer model of appointments based on our foundation could generate new implications about when Presidential appointments most degrade agency performance, and thus what those appointments will be.

7 Conclusion

Stepping back from the study of principal-agent relationships, our analysis is one of the few to consider the consequences of open differences in beliefs in politics; the literature has instead focused primarily on conflict driven by differing goals or ideologies. This is natural given that citizens’ policy preferences appear closely tied to their economic interests (Gelman 2010) and legislators’ policy positions are strongly predicted by their seemingly stable ideologies (Poole 1997). However, history also furnishes examples of large-scale changes in policy preferences that occurred *across* the ideological spectrum in response to learning through policy experience. For example, Hall (1993) describes how high unemployment and stagnating

growth in 1970s Britain discredited Keynesian policies among politicians on both sides of the ideological spectrum who were “seeking solutions to Britain’s economic problems.” More recently, the Democratic party in the United States evolved significantly toward market-driven policies; President Clinton’s FCC Commissioner Reed Hundt describes this evolution as such – “where the Republicans were right and the old Democrats were wrong, the change in control of Congress could permit us to do what we thought was best. Auction spectrum. Shrink the agency. Fight for free trade in communications services...” (Hundt 2000, pp. 98-99).

Our analysis illustrates some interesting incentives that are specific to belief conflicts; chief among these is the incentive to *persuade* other individuals that their beliefs are mistaken by generating new unbiased information. In our model, this incentive manifests itself as a somewhat “dastardly” form of deference; a principal sometimes deliberately sets up her agent for failure in order to effectively persuade him that his beliefs are wrong. This effect also manifests itself in interesting ways in several other studies of belief conflict and learning; in Yildiz (2004) disagreement in a bargaining game about recognition probabilities creates a persuasion-based incentive for delay; in Che and Kartik (2009) disagreement between a decision-maker and a pool of advisers creates a persuasion-based incentive for costly information acquisition; in Smith and Stam (2004) disagreement between two states about their military strength creates a persuasion-based incentive for war. Given the generality of this effect, it is likely to have implications for policymaking outside of principal-agent environments.

One potential application of this idea is to the political uses of *pilot programs* in federal, state, and local governments. Such projects have historically been justified on technocratic grounds. However, extending our idea that information can be produced strategically to persuade, pilot programs might also be fruitfully thought of as a form of political entrepreneurship. Specifically, individuals and interest groups who are *already* invested in the value of specific policies may try to strategically design such programs to maximize the chance of

generating persuasive evidence of their efficacy to gain political support. This characterization of pilot experimentation dovetails with Carpenter’s (2001) account of the behavior of federal agency heads in early modern America, who engaged in entrepreneurial innovation on sometimes dubious legal grounds to “convince diverse coalitions... of the value of their ideas” (Carpenter 2001, p. 30). Developing a model of pilot experimentation based on our insights could generate important implications for when such programs should be used, and how their results should be evaluated.

Even more generally, such “persuasion”-based incentives might bias the choices of both citizens and politicians in a setting of repeated policy experimentation like Callander (2011) when there is belief disagreement. For example, leaders might allow controlled disasters to occur to persuade other politicians and citizens of the need for action in a crisis; events like the recent failure of Lehman Brothers and the eventual passage of the TARP come to mind. Alternatively, some citizens might be willing to temporarily empower extremist parties in order to “prove” the bankruptcy of their policy ideas to the larger polity and discredit them in the long term. Moreover, these incentives, if present, would naturally be influenced by the rules of policymaking a government; for example, incentives to experiment may be enhanced by supermajority requirements like filibuster rules in the U.S. Senate because it is more essential to persuade and build broad consensus. We hope to explore these and related avenues in future work.

References

- Aberbach, Joel D. 2000. *In the Web of Politics*. Washington, D.C: Brookings Institution Press.
- Aumann, Robert J. 1987. "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica* 55(January): 1.
- Bai, Matt. 2006. "The Inside Agitator." *The New York Times* (October).
- Bendor, Jonathan, and Adam Meirowitz. 2004. "Spatial Models of Delegation." *The American Political Science Review* 98(May): 293–310.
- Bendor, Jonathan, and Thomas H. Hammond. 1992. "Rethinking Allison's Models." *American Political Science Review* 86(June): 301–322.
- Bergemann, Dirk, and Ulrich Hege. 2005. "The Financing of Innovation: Learning and Stopping." *The RAND Journal of Economics* 36(December): 719–752.
- Bertelli, Anthony, and Sven E. Feldmann. 2007. "Strategic Appointments." *Journal of Public Administration Research and Theory* 17(January): 19–38.
- Bertelli, Anthony M., and Christian R. Grose. 2011. "The Lengthened Shadow of Another Institution?" *American Journal of Political Science* 55(4): 767–781.
- Besley, Timothy, and Maitreesh Ghatak. 2006. "Sorting with Motivated Agents." *Journal of the European Economic Association* 4(2-3): 404–414.
- Bratton, William. 1998. *Turnaround*. New York: Random House.
- Brehm, John. 1999. *Working, Shirking, and Sabotage*. Ann Arbor: University of Michigan Press.
- Bueno De Mesquita, Ethan, and Matthew C. Stephenson. 2007. "Regulatory Quality Under Imperfect Oversight." *American Political Science Review* 101(03): 605–620.

- Callander, Steven. 2011. "Searching for Good Policies." *American Political Science Review* 105(4): 643–662.
- Campbell, Donald T. 1969. "Reforms as experiments." *American Psychologist* 24(4): 409–429.
- Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy*. Princeton, N.J.: Princeton University Press.
- Che, Yeon Koo, and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy* 117(October): 815–860.
- DiIulio, John J. 1987. *Governing Prisons*. London: Collier Macmillan.
- Dobel, P. J. 1992. "William D. Ruckelshaus." In *Exemplary Public Administrators*, ed. Terry L. Cooper. San Francisco: Jossey-Bass Publishers pp. 241–269.
- Gailmard, Sean, and John W. Patty. 2012. "Formal Models of Bureaucracy." *Annual Review of Political Science* 15(1): 353–377.
- Gailmard, Sean, and John W. Patty. 2013. *Learning While Governing*. Chicago: The University of Chicago Press.
- Gates, Robert Michael. 2014. *Duty: Memoirs of a Secretary at War*. New York: Alfred A. Knopf.
- Geanakoplos, John. 1994. "Chapter 40: Common Knowledge." In *Handbook of Game Theory with Economic Applications*, ed. Robert Aumann, and Sergiu Hart. Vol. 2 Elsevier pp. 1437–1496.
- Geanakoplos, John D, and Heraklis M Polemarchakis. 1982. "We can't disagree forever." *Journal of Economic Theory* 28(October): 192–200.

- Gelman, Andrew. 2010. *Red State, Blue State, Rich State, Poor State*. Princeton, N.J: Princeton University Press.
- Golden, Marissa. 2000. *What Motivates Bureaucrats?* New York: Columbia University Press.
- Goodsell, Charles T. 2011. *Mission Mystique*. Washington, D.C: CQ Press.
- Gul, Faruk. 1998. "A Comment on Aumann's Bayesian View." *Econometrica* 66(July): 923–927.
- Hall, Peter A. 1993. "Policy Paradigms, Social Learning, and the State." *Comparative Politics* 25(April): 275–296.
- Heckman, James J., Jeffrey A. Smith, and Christopher Taber. 1996. What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program. Working Paper 5535 National Bureau of Economic Research.
- Heclo, Hugh. 1977. *A Government of Strangers*. Washington: Brookings Institution.
- Hundt, Reed E. 2000. *You Say You Want a Revolution*. New Haven: Yale University Press.
- Kaufman, Herbert. 1967. *The Forest Ranger*. Baltimore: Johns Hopkins Press.
- Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica* 73(1): 39–68.
- Kettl, Donald F. 2005. *The Global Public Management Revolution*. Washington, D.C: Brookings Institution Press.
- Krier, James E. 1977. *Pollution and Policy*. Berkeley: University of California Press.
- Lee, Kai N. 1993. *Compass and Gyroscope*. Washington, D.C: Island Press.

- Lewis, David E. 2008. *The Politics of Presidential Appointments*. Princeton: Princeton University Press.
- Lipsky, Michael. 1980. *Street-level Bureaucracy*. New York: Russell Sage Foundation.
- Long, Edward, and Aimee L. Franklin. 2004. "The Paradox of Implementing the Government Performance and Results Act: Top-Down Direction for Bottom-Up Implementation." *Public Administration Review* 64(3): 309–319.
- Maple, Jack, and Chris Mitchell. 2010. *The Crime Fighter*. Random House Digital, Inc.
- May, Peter J. 1992. "Policy Learning and Failure." *Journal of Public Policy* 12(October): 331–354.
- McCarty, Nolan. 2004. "The Appointments Dilemma." *American Journal of Political Science* 48(3): 413–428.
- Morris, Stephen. 1995. "The Common Prior Assumption in Economic Theory." *Economics and Philosophy* 11(02): 227–253.
- Moynihan, Donald P. 2008. *The Dynamics of Performance Management*. Washington, D.C: Georgetown University Press.
- Moynihan, Donald P., and Noel Landuyt. 2009. "How Do Public Organizations Learn? Bridging Cultural and Structural Perspectives." *Public Administration Review* 69(6): 1097–1105.
- Mutz, Diana C. 2008. "Is Deliberative Democracy a Falsifiable Theory?" *Annu. Rev. Polit. Sci.* 11: 521–538.
- Nie, Martin A., and Courtney A. Schultz. 2012. "Decision-Making Triggers in Adaptive Management." *Conservation Biology* 26(6): 1137–1144.

- Perry, James L., and Annie Hondeghem. 2008. *Motivation in Public Management: The Call of Public Service*. Oxford: Oxford University Press.
- Pollitt, Christopher. 2006. "Performance Information for Democracy." *Evaluation* 12(January): 38–55.
- Poole, Keith T. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Prendergast, Canice. 2008. "Intrinsic Motivation and Incentives." *The American Economic Review* 98(May): 201–205.
- Sabatier, Paul A. 1988. "An advocacy coalition framework of policy change and the role of policy-oriented learning therein." *Policy Sciences* 21(June): 129–168.
- Smith, Alastair, and Allan C. Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(December): 783–813.
- Starobin, Paul. 1995. *Surviving at the EPA: Gary Dietrich*. (C16-84-592.0) Cambridge, MA: Kennedy School of Government Case Program.
- Stephenson, Matthew C. 2007. "Bureaucratic Decision Costs and Endogenous Agency Expertise." *Journal of Law, Economics, and Organization* 23(June): 469–498.
- Thompson, Frank J., and Norma M. Riccucci. 1998. "Reinventing Government." *Annual Review of Political Science* 1(1): 231–257.
- Ting, Michael M. 2011. "Organizational Capacity." *Journal of Law, Economics, and Organization* 27(August): 245–271.
- U.S. Government Accountability Office. 1996. *Managing Reform: Status of Agency Reinvention Lab Efforts*. (GAO/GGD-96-69) Washington, D.C.: U.S. Government Printing Office.

- U.S. Government Accountability Office. 2005. *Performance Budgeting*. (GAO-05-117SP)
Washington, D.C.: U.S. Government Printing Office.
- Van den Steen, Eric. 2005. "Organizational Beliefs and Managerial Vision." *Journal of Law, Economics, and Organization* 21(April): 256–283.
- Van den Steen, Eric. 2009. "Authority versus Persuasion." *The American Economic Review* 99(May): 448–453.
- Van den Steen, Eric. 2010. "Disagreement and the Allocation of Control." *Journal of Law, Economics, and Organization* 26(August): 385–426.
- Volden, Craig, Michael M. Ting, and Daniel P. Carpenter. 2008. "A Formal Model of Learning and Policy Diffusion." *American Political Science Review* 102(August): 319–332.
- Weisburd, David, Stephen D. Mastrofski, Ann Marie McNally, Rosann Greenspan, and James J. Willis. 2003. "Reforming to Preserve." *Criminology & Public Policy* 2(3): 421–456.
- Wilson, James Q. 1989. *Bureaucracy*. New York: Basic Books.
- Wilson, James Q., and George L. Kelling. 1982. "Broken Windows." *Atlantic Monthly* 249(3): 29–38.
- Yildiz, Muhamet. 2004. "Waiting to Persuade." *The Quarterly Journal of Economics* 119(February): 223–248.

Online Appendix

A Equilibrium Characterization

We now introduce additional additional notation and provide general equilibrium conditions for the baseline model. A strategy for the agent consists of two functions $e^1(x^1), e^2(x^1, e^1, y^1, x^2)$ to $[0, 1]$ mapping histories to effort. Since the principal cannot observe the agent's effort, her strategy is a probability p^1 of initially choosing policy a , and probabilities $p_{y^1}^{x^1}$ of sticking with the initial policy x^1 after outcome y^1 for every (x^1, y^1) .

The Agent's Problem In the second period the agent exerts effort $e^2(x^1, e^1, y^1, x^2) = \lambda P_2(\omega = x^2 | x^1, e^1, y^1)$ and his expected utility is $\frac{\lambda}{2} [P_2(\omega = x^2 | x^1, e^1, y^1)]^2$. Denoting his prior that $\omega = a$ as θ and his initial effort as e for simplicity, his expected two-period utility when $x^1 = a$ as a function of first period effort is

$$\begin{aligned} U_2(e, \theta, p_s^a, p_f^a) &= \left(\theta e - \frac{e^2}{2\lambda} \right) + \frac{\lambda}{2} (p_f^a \cdot \theta (e + (1 - e) h(e, \theta)) + (1 - p_f^a) \cdot (1 - \theta) (1 - h(e, \theta))) \\ &\quad + (p_s^a - p_f^a) \cdot \frac{\lambda}{2} \theta e. \end{aligned} \tag{5}$$

By symmetry his expected utility from effort e^1 on some policy x is $U_2(e^1, P_2(\omega = x), p_s^x, p_f^x)$. To derive equations (3) and (4) observe that “rigidly implementing x ” is equivalent to $(p_s^x = p_f^x = 1)$ and “experimenting with x ” is equivalent to $(p_s^x = 1, p_f^x = 0)$; thus, eqn. (3) is $U_2(e^1, \theta_2, 1, 1)$ and eqn. (4) is $U_2(e^1, \theta_2, 1, 0)$.

The Principal's Problem The principal's second period policy choices must be interim-optimal given $e^2(x^1, e^1, y^1, x^2)$ and her posteriors computed with the agent's equilibrium strategy. Thus, she must stay with the initial policy x if it succeeds ($p_s^x = 1$), and can only stay with the initial policy if it fails ($p_f^x > 0$) i.f.f. $h(e^1(x), P_1(\omega = x)) \geq 1 - h(e^1(x), P_2(\omega = x))$. If the inequality is strict then $p_f^x = 1$. In period 1 the principal's expected utility from se-

lecting policy a when she expects effort e^1 and future equilibrium behavior is,

$$\begin{aligned} U_1(e^1, \theta_1, \theta_2, p_s^a, p_f^a) &= \theta_1(e^1 + \lambda(e^1 p_s^a + (1 - e^1) p_f^a \cdot h(e^1, \theta_2))) \\ &\quad + (1 - \theta_1)(1 - p_f^a) \lambda(1 - h(e^1, \theta_2)) \end{aligned} \quad (6)$$

By symmetry her expected utility from some x is $U_1(e^1, P_1(\omega = x), P_2(\omega = x), p_s^x, p_f^x)$.

Equilibrium Conditions The following conditions are necessary and sufficient for equilibrium.

Lemma 1. *Strategies $(x^1, p_s^a, p_f^a, p_s^b, p_f^b)$ and $(e^1(x^1), e^2(x^1, e^1, y^1, x^2))$ are an equilibrium i.f.f.*

(Agent Optimality)

1. $e^2(x^1, e^1, y^1, x^2) = \lambda P_2(\omega = x^2 | x^1, e^1, y^1)$ (the agent optimizes in the second period)
2. $e^1(x) \in \arg \max_{e^1 \in [0,1]} \{U_2(e^1, P_2(\omega = x), p_s^x, p_f^x)\} \ \forall x \in \{a, b\}$ (the agent optimizes in the first period given the principal's strategy and expectations about his own future effort)

(Principal Optimality)

1. $p_s^x = 1 \ \forall x \in \{a, b\}$ (the principal always stays after success)
2. $p_f^x \geq 0 \iff h(e^1(x), P_1(\omega = x)) \geq 1 - h(e^1(x), P_2(\omega = x))$ and $= 1$ if satisfied with strict inequality (the principal only stays after failure if it is interim-optimal given on-path posteriors)
3. $x^1 \in \arg \max_{x \in \{a, b\}} \{U_1(e^1(x), P_1(\omega = x), P_2(\omega = x), p_s^x, p_f^x)\}$ (the principal's initial policy choice maximizes her expected continuation value). ■

B Proofs

Proof of Propositions 1 and 2

The agent's objective function in first period effort e^1 given $x^1 = a$ is $U_2(e^1, \theta_2, p_s^a, p_f^a)$; say that it is *well-behaved* when there is a unique interior optimum characterized by the lowest (and possibly unique) solution to the FOC. This property is not straightforward because the objective function is not generically concave. In Lemma 2 in the Supplemental Appendix we show the set of λ s.t. the problem is well-behaved for all feasible parameters is an interval $\lambda \in [0, \bar{\lambda})$ where $\bar{\lambda} \approx .68466$, and henceforth proceed assuming this property holds. The proof has 3 parts; first we provide a first-order characterization of the agent's optimal effort, second we show that effort is higher with rigid implementation than in the no learning benchmark, and higher with experimentation than rigid implementation, and third we show that when the principal rigidly implements a policy, observing outcomes generates a higher probability of success in both periods.

Part 1 We temporarily write e^1 as e and θ_2 as θ for simplicity, and before taking the derivative we note a few useful properties. First, the derivative w.r.t. e of the agent's posterior belief after failure is $\frac{\partial h(e, \theta)}{\partial e} = -\frac{\theta(1-\theta)}{(1-\theta e)^2}$. Next, the derivatives of both $\theta(e + (1-e)h(e, \theta))$ and $(1-\theta)(1-h(e, \theta))$ w.r.t. e are,

$$\frac{\lambda}{2}\theta \cdot k(e, \theta) > 0 \text{ for } \theta \in (0, 1), \text{ where } k(e, \theta) = (1-h(e, \theta))^2 \quad (7)$$

To show this use that $(1-e)\frac{\partial h}{\partial e} = -h(\cdot)(1-h(\cdot))$ and $-(1-\theta)\frac{\partial h}{\partial e} = \theta(1-h(\cdot))^2$. It is also easy to verify that $k(e, \theta)$ is increasing and convex in e . Now applying these observations and rearranging, we have

$$\frac{\partial U_2}{\partial e} = \frac{1}{\lambda} \left(-e + \lambda \theta \left(1 + \frac{\lambda}{2} [k(e, \theta) + \delta^a] \right) \right), \quad (8)$$

where $\delta^a = p_s^a - p_f^a$ is the *increase* in the probability of staying with a after success – recall that $p_s^a = 1$ (the principal always stays after success in equilibrium) so $\delta^a \in [0, 1]$. Thus, when the problem is well behaved we may write the unique interior solution as $e^*(\theta_2, \delta^a)$ where $\delta^a = p_s^a - p_f^a$ and $e^*(\theta_2, \delta^a)$ satisfies

$$e^*(\theta_2, \delta^a) = \lambda\theta \left(1 + \frac{\lambda}{2} [k(e^*(\theta_2, \delta^a), \theta_2) + \delta^a] \right)$$

Applying symmetry, the agent's unique best-response to either first period policy $x \in \{a, b\}$ is $e^*(P_2(\omega = x), p_s^x - p_f^x)$.

Part 2 Rigid implementation corresponds to $p_s^a = p_f^a = 1 \rightarrow \delta^a = 0$, and effort in the no learning benchmark is $\lambda\theta$. Since $\left. \frac{\partial U_2(e^1, \theta, 1, 1)}{\partial e} \right|_{e=\lambda\theta} = \frac{\lambda\theta}{2} k(e, \theta) > 0 \forall \theta \in (0, 1)$, $e^*(\theta, 0) > \lambda\theta$. Next, $\frac{\partial U_2}{\partial e \partial \delta^a} = \frac{\lambda\theta}{2} > 0$; thus, $e^*(\theta, \delta^a)$ is increasing in δ^a . Consequently, experimenting ($p_s^a = 1$ and $p_f^a = 0 \rightarrow \delta^a = 1$) elicits more effort for all θ than rigid implementation ($\delta^a = 0$). Finally, because the preceding results are shown for all $\theta \in [0, 1]$, they apply to both $x^1 \in \{a, b\}$ by symmetry.

Part 3 The probability of success being higher in the first period follows immediately from the agent's greater effort. In the second period, when the agent exerts first period effort e^1 and optimizes in the second period the expected second-period probability of success equal to,

$$\theta \cdot (\lambda(e^1 + (1 - e^1)h(e^1, \theta_2))),$$

where θ_2 is the agent's prior, and θ is the true probability. This is increasing in e^1 and equal to the probability of success in the no learning benchmark $\theta \cdot \lambda\theta_2$ at $e^1 = 0$, proving the result. ■

Proof of Proposition 3

We proceed in two parts. In Part 1 we make an equilibrium selection and introduce additional notation employed in the subsequent proofs. In Part 2, we prove the proposition with a mixture of analytical and computational arguments.

Part 1 Lemma 4 in the Supplemental Appendix proves that whenever experimenting is an equilibrium in a subgame x^1 , it is the optimal strategy for the principal even if she could precommit to her future decisions. Thus, absent commitment it must also be the best possible equilibrium for the principal. Furthermore, if experimenting with $x^1 = a$ is *not* an equilibrium of the subgame following $x^1 = a$, then the unique equilibrium is rigid implementation. If experimentation is not an equilibrium then $h(e^*(\theta_2, 1 - p_f^a), \theta_1) > 1 - h(e^*(\theta_2, 1 - p_f^a), \theta_2)$ at $p_f^a = 0$. Another equilibrium with $p_f^a > 0$ would require that the l.h.s. \leq r.h.s. – but this cannot be since $e^*(\theta_2, 1 - p_f^a)$ is decreasing in p_f^a , so the r.h.s. is increasing and the l.h.s. is increasing.

We can therefore select the equilibria that are best for the principal by considering only pure strategy equilibria, and choosing experimentation in a subgame x^1 whenever it is an equilibrium. To do so we introduce additional notation. First let $e^s(\theta_2)$ denote the agent's first-period best response when the principal's pure strategy is $(x^1 = a, s)$, where $s \in \{R, E\}$ denotes whether the principal (**R**)igidly implements or (**E**)xperiments with the initial policy; so $e^R(\theta_2) = e^*(\theta_2, 0)$ and $e^E(\theta_2) = e^*(\theta_2, 1)$. Second, let $V_1^s(\theta_1, \theta_2)$ denote the principal's two-period expected utility when her pure strategy is $(x^1 = a, s)$ and the agent best-responds, so $V_1^R(\theta_1, \theta_2) = U_1(e^R(\theta_2), \theta_1, \theta_2, 1, 1)$ and $V_1^E(\theta_1, \theta_2) = U_1(e^E(\theta_2), \theta_1, \theta_2, 1, 0)$. Third, let $\bar{\theta}(\theta_2)$ be the unique solution to

$$h(e^R(\theta_2), \bar{\theta}(\theta_2)) = 1 - h(e^R(\theta_2), \theta_2), \quad (9)$$

By Lemma 1, experimenting with $x^1 = a$ is an equilibrium of that subgame i.f.f. $\theta_1 \leq \bar{\theta}(\theta_2)$,

and it is easy verified that $\bar{\theta}(\theta_2) > 1 - \theta_2$. Finally, by symmetry the agent's effort on $x^1 = b$ is $e^s(1 - \theta_2)$, the principal's utility is $V_1^s(1 - \theta_1, 1 - \theta_2)$, and the threshold for experimentation with b is $1 - \theta_1 < \bar{\theta}(1 - \theta_2) \iff \theta_1 > 1 - \bar{\theta}(1 - \theta_2)$.

Part 2 In Lemma 3 in the Supplemental Appendix we show that the agent's best-response $e^s(\theta_2)$ is increasing in his motivation $\lambda \forall s \in \{R, E\}$. Since $h(e, \theta)$ is decreasing in e and increasing in θ , it is straightforward to observe from the definition in (9) that $\bar{\theta}(\theta_2)$ is increasing in λ (and consequently $1 - \theta(1 - \theta_2)$ is decreasing in λ). Thus, the region of the parameter space where the principal would experiment down *both* paths of play is increasing in the agent's motivation λ . Now to show that experimentation expands overall, we must *also* show that within the regions where the principal rigidly implements one of the two policies x (and therefore experiments with the other $\neg x$), increasing λ cannot induce her to switch from initially selecting $x^1 = \neg x$ to $x^1 = x$; we prove this property in Lemma 5 in the Supplemental Appendix, and the proof relies on the supplemental Mathematica code. This completes the proof. ■

Proof of Proposition 4

In Lemma 6 of the Supplemental Appendix, we prove that fixing the principal's experimentation decisions down each path of play, if she prefers $x^1 = a$ given beliefs $\hat{\theta}_1$ then she also prefers it for all higher beliefs. We call this property “preference monotonicity” and employ it in the subsequent proofs. First, there is never *over*experimentation in equilibrium, i.e., the principal never experiments with x^1 when rigidly implementing either x^1 or $\neg x^1$ would be better. The former is immediately ruled out Lemma 4. The latter also ruled out – if rigidly implementing $\neg x^1$ were optimal with commitment then it must be better than experimenting with $\neg x^1$, and by implication the unique equilibrium of the subgame following $\neg x^1$ without commitment; thus, the principal failing to choose it would be a contradiction.

Next, we there $\exists(\theta_1, \theta_2)$ s.t. underexperimentation occurs, i.e. the principal rigidly im-

plements x^1 when experimenting would be better. In part 3 of Lemma 9 in the Supplemental Appendix we show there exists a nonempty interval of θ_2 s.t. the principal prefers rigidly implementing b to experimenting with a when $\theta_1 = 1 - \bar{\theta}(1 - \theta_2)$. By continuity, rigidly implementing b is thus the equilibrium outcome for $\theta_1 = 1 - \bar{\theta}(1 - \theta_2) - \varepsilon$ when $\varepsilon > 0$ is sufficiently close to 0, and it is also worse than experimenting with b since the agent's effort drops discretely. Formally,

$$\begin{aligned}
V_1^E(\bar{\theta}(1 - \theta_2), 1 - \theta_2) &= U_1(e^E(1 - \theta_2), \bar{\theta}(1 - \theta_2), 1 - \theta_2, 1, 0) \\
&= U_1(e^E(1 - \theta_2), \bar{\theta}(1 - \theta_2), 1 - \theta_2, 1, 1) > U_1(e^R(1 - \theta_2), \bar{\theta}(1 - \theta_2), 1 - \theta_2, 1, 1) \\
&= V_1^R(\bar{\theta}(1 - \theta_2), 1 - \theta_2)
\end{aligned}$$

The first and last equalities follow from the definitions, the second follows from the definition of $\bar{\theta}(\cdot)$, and the inequality from $U_1(\cdot)$ increasing in e^1 and $e^E(1 - \theta_2) > e^R(1 - \theta_2)$. ■

Proof of Proposition 5

If there were an exogenous cost $c > 0$ of maintaining policy after failure, then for experimentation to fail to be an equilibrium of the subgame following policy x requires that the principal prefer to reselect x given initial effort $e^E(P_2(\omega = x))$, the players' resulting posterior beliefs, *and* the cost c . This condition is,

$$h(e^E(P_2(\omega = x)), P_1(\omega = x)) - \frac{c}{\lambda} > 1 - h(e^E(P_2(\omega = x)), P_2(\omega = x))$$

Now suppose the principal underexperiments with policy x^{1*} when $c = 0$, and also suppose w.l.o.g. that $x^{1*} = a$. Experimenting must then be the unique equilibrium of the subgame following b , it will remain so with any higher cost $c > 0$, and also rigidly implementing a is better than experimenting with b . So experimenting with a is the optimal pure strategy policy sequence with commitment, but not an equilibrium of the subgame following a without

commitment. However, it will become one when

$$c \geq \lambda \left(h \left(e^E (\theta_2), \theta_1 \right) - \left(1 - h \left(e^E (\theta_2), \theta_2 \right) \right) \right),$$

and so the principal will select it in equilibrium. A symmetric argument holds when the principal underexperiments with b . ■

Proof of Proposition 6

Suppose appointing herself is not optimal; then for a principal with beliefs θ_1 , the resulting equilibrium (x^*, s^*) is strictly worse than the equilibrium that would result if the optimal appointee with beliefs $\hat{\theta}_1$ were making policy decisions. Denote this equilibrium (\hat{x}, \hat{s}) . First note \hat{s} must not be an equilibrium experimentation decision in the subgame following \hat{x} for θ_1 (since otherwise the principal would choose it). We next argue that $\hat{s} = E$. If $\hat{s} = R$ then by implication experimenting must be an equilibrium of the subgame following $x^1 = \hat{x}$ for θ_1 ; but then by Lemma 4 it is also strictly better than rigidly implementing \hat{s} and we have a contradiction. Finally, since $(\hat{x}, \hat{s} = E)$ is an equilibrium of the subgame following \hat{x} for $\hat{\theta}_1$ but not θ_1 , it follows that

$$\begin{aligned} h \left(e^E (P_2 (\omega = \hat{x})), \hat{P}_1 (\omega = \hat{x}) \right) &< 1 - h \left(e^E (P_2 (\omega = \hat{x})), P_2 (\omega = \hat{x}) \right) \\ &< h \left(e^E (P_2 (\omega = \hat{x})), P_1 (\omega = \hat{x}) \right), \end{aligned}$$

implying $\hat{P}_1 (\omega = \hat{x}) < P_1 (\omega = \hat{x})$. ■

Proof of Proposition 7

First, whenever the principal rigidly implements a policy x^1 in equilibrium, she must also choose that same policy in the no learning benchmark, since $P_1 (x^1 = \omega) \geq \bar{\theta} (P_2 (x^1 = \omega)) > 1 - P_2 (x^1 = \omega)$. We next show the set of beliefs where the principal chooses $x^1 = b$ expands by arguing that the following two conditions proved in the Supplemental Appendix are jointly

sufficient; 1) for all θ_2 s.t. $1 - \bar{\theta}(1 - \theta_2) > \frac{1}{2}$, a principal with beliefs $\theta_1 = 1 - \bar{\theta}(1 - \theta_2)$ weakly prefers rigidly implementing b to experimenting with a , and 2) a principal with beliefs $\theta_1 = 1 - \theta_2$ strictly prefers experimenting with b to experimenting with a . If Property 1 holds, then a principal with beliefs $\theta_1 \in [\frac{1}{2}, 1 - \bar{\theta}(1 - \theta_2)]$ who would rigidly implement b if chosen also prefers that to experimenting with a and so selects b initially. If Property 2 holds, then the principal prefers experimenting with b to experimenting with a when θ_1 is $<$ some $\hat{\theta}_1$ that is $> 1 - \theta_2$, and thus experiments with b in equilibrium when $\theta_1 \in [1 - \bar{\theta}(1 - \theta_2), \min\{\bar{\theta}(\theta_2), \hat{\theta}\}]$ where $\min\{\bar{\theta}(\theta_2), \hat{\theta}\} > 1 - \theta_2$. (She may also experiment with b when $\theta_1 > \bar{\theta}(\theta_2)$). This completes the argument.

Property 1 is proved with the aid of Mathematica in Lemma 7 in the Supplemental Appendix, and requires $\lambda > \underline{\lambda} \approx .23505$. Property 2 is proved analytically in Lemma 8 in the Supplemental Appendix, and holds for all $\lambda < 1$. Note that property 2 alone is sufficient for the desired result if the principal could commit ex-ante to her strategy (pure or mixed); for $\theta_1 \leq 1 - \theta_2 < \bar{\theta}(\theta_2)$ experimenting with a is better than any other strategy with a by Lemma 4, and so if experimenting with b is better than experimenting with a then it is also better than any other strategy with a . ■

Proof of Proposition 8

We first argue that following three conditions on θ_2 are jointly *sufficient* for the principal to always defer in the first period regardless of her own beliefs; 1) $V_1^E(0, 1 - \theta_2) \geq V_1^R(1, \theta_2)$, 2) $V_1^R(1, \theta_2) > V_1^E(1, \theta_2)$, 3) $V_1^R(\bar{\theta}(1 - \theta_2), 1 - \theta_2) > V_1^E(1 - \bar{\theta}(1 - \theta_2), \theta_2)$. Conditions (1) and (2) jointly imply that experimenting with b is better than both experimenting with or rigidly implementing a when $\theta_1 = 1$; by preference monotonicity (proved in Lemma 6 in the Supplemental Appendix) this also implies that experimenting with b is better $\forall \theta_1 \in [0, 1]$. Thus, whenever experimenting with b is an equilibrium strategy ($\theta_1 \geq 1 - \bar{\theta}(1 - \theta_2)$) it is chosen. Now whenever experimenting with b is not an equilibrium strategy ($\theta_1 < 1 - \bar{\theta}(1 - \theta_2)$), the principal compares rigidly implementing b to experimenting with a ; again applying pref-

erence monotonicity, condition (3) implies that she prefers the former $\forall \theta_1 \leq 1 - \bar{\theta}(1 - \theta_2)$. All possible principal beliefs are covered, which completes the argument. We next argue that condition (1) is *necessary* for the principal to always defer regardless of her own beliefs; if it fails then $V_1^R(1, \theta_2) > V_1^E(0, 1 - \theta_2)$. For $\theta_1 > \bar{\theta}(\theta_2)$ the principal would rigidly implement a and experiment with b , and by continuity she also prefers rigidly implementing a to experimenting with b for θ_1 sufficiently close to 1. Thus for such θ_1 she selects $x^1 = a$ in equilibrium and does not defer.

Finally, Lemma 9 in the Supplemental Appendix proves analytically that when $\lambda > \hat{\lambda}$ (where $\hat{\lambda}$ is the unique solution to $\lambda(1 + \lambda)(1 + \frac{\lambda}{2}) = 1$ and $\approx .5214$), each of the three conditions $k \in \{1, 2, 3\}$ holds for θ_2 in a nonempty interval $(0, \varepsilon_k)$.²⁵ Since they then all hold for $\theta_2 \in (0, \varepsilon_k)$, $\lambda > \hat{\lambda}$ is therefore sufficient for existence of a range of θ_2 where the principal always defers. In addition, the supplemental mathematica code verifies that condition 1 fails $\forall \theta_2 > 0$ when $\lambda \leq \hat{\lambda}$, which is equivalent to

$$\frac{V_1^E(0, 1 - \theta_2) - V_1^R(1, \theta_2)}{\lambda \theta_2} < 0 \quad \forall \theta_2 \in \left[0, \frac{1}{2}\right] \quad \text{when } \lambda < \hat{\lambda}.$$

$\lambda > \hat{\lambda}$ is thus also *necessary* for existence of a range of θ_2 where the principal always defers.

■

²⁵Note that we have already shown that property (3) holds $\forall \theta_2$ when $\lambda > \underline{\lambda}$ in Lemma 7; however, the proof is computational. In Lemma 9 we prove analytically that property (3) holds for sufficiently low θ_2 for any $\lambda \in [0, 1]$.

C Supplemental Appendix

Lemma 2 (Good Behavior). *The set of λ s.t. the agent's problem is well-behaved for all feasible parameters is an interval $\lambda \in [0, \bar{\lambda})$, where $\bar{\lambda} \in \left(\frac{\sqrt{5}-1}{2}, \sqrt{3}-1\right)$ and is $\approx .68466$.*

Proof: From the proof of Propositions 1 and 2, the derivative of the agent's objective function is

$$\frac{\partial U_2}{\partial e} = \frac{1}{\lambda} \left(-e + \lambda \theta \left(1 + \frac{\lambda}{2} [k(e, \theta) + \delta] \right) \right),$$

Now observe the following. First, $\frac{\partial U_2}{\partial e} > 0$ at $e = 0$. Second, $\frac{\partial U_2}{\partial e}$ is convex in e , which follows from the convexity of $k(e, \theta)$. Combining the previous two observations, there are at most two solutions to the first order condition. When there are two solutions the optimum is either the lower solution or maximum effort $e = 1$, when there is one solution it is exactly the optimum, and when there are no solutions the optimum is $e = 1$. Thus, whenever $e = 1$ is not the optimum, the problem is well-behaved. Third, observe that if the problem is well behaved for λ' , then holding the other parameters fixed it is also well behaved for $\lambda'' < \lambda'$. The cross partial of the objective function in (e, λ) is $\frac{e}{\lambda^2} + \frac{\theta}{2} (k(e, \theta) + \delta) > 0$, so the set of optima must decrease when λ decreases. If the problem were well behaved at λ' but not $\lambda'' < \lambda'$ then $e = 1$ would be optimal for λ'' but not λ' , a contradiction.

Thus, for any profile of parameters the set of λ s.t. the problem is well-behaved is an interval; consequently, the set of λ s.t. the problem is well-behaved for *all* feasible parameters is also an interval $[0, \bar{\lambda})$. Moreover, $\bar{\lambda} > \frac{\sqrt{5}-1}{2}$ since the problem is strictly concave for all feasible parameters (and by implication well behaved) when $\frac{\partial U_2}{\partial e}|_{e=1} = -\frac{1}{\lambda} + \theta \left(1 + \frac{\lambda}{2} (1 + \delta) \right) \leq -\frac{1}{\lambda} + (1 + \lambda) < 0$, which holds i.f.f. $\lambda < \frac{\sqrt{5}-1}{2}$. It is also $< \sqrt{3}-1$ since best-response effort at $\theta_2 = 1$ when the principal experiments is $e(1, 1) = \lambda \left(1 + \frac{\lambda}{2} \right) < 1$ i.f.f. $\lambda < \sqrt{3}-1$.

Finally, in the supplemental mathematica code to this document, we verify that $\bar{\lambda} \approx .68466$. Since the set of optima is increasing in δ^a (by $\frac{\partial^2 U_2}{\partial e \partial \delta} = \lambda \theta > 0$), to find $\bar{\lambda}$ it suffices to check that the problem is well-behaved $\forall \theta$ when $\delta = 1$ (i.e. the principal experiments).

We thus identify $\bar{\lambda}$ by compute the highest λ s.t. $\forall \theta \in [0, 1]$, the agent's utility at the lowest solution to the first-order condition is greater than his utility from $e = 1$. ■

Lemma 3. *The agent's optimal effort $e^*(\theta, \delta)$ is increasing in λ and θ .*

Proof: In the proof of Lemma 2 we showed that $\frac{\partial^2 U_2}{\partial e \partial \lambda} > 0$ – the first comparative static is thus immediate. The second comparative static is less straightforward because it is not universally the case that $\frac{\partial^2 U_2}{\partial e \partial \theta} > 0$. However, it suffices to show that $\frac{\partial^2 U_2}{\partial e \partial \theta} > 0 \forall e \in [0, e^*(1, \delta)]$ for the following reason; if this condition holds, then it implies that every $e \in [0, e^*(1, \delta)]$ is optimal for at most one θ . Consequently $e^*(\theta, \delta)$ must be strictly monotone in θ – otherwise some $e \in [0, e^*(1, \delta)]$ would be optimal for more than one θ . Finally since $e^*(0, \delta) = 0 < e^*(1, \delta) = \lambda(1 + \frac{\lambda}{2}\delta)$, it must be strictly increasing.

We now prove that $\frac{\partial^2 U_2}{\partial e \partial \theta} > 0 \forall e \in [0, e^*(1, \delta)]$ when $\lambda \leq \bar{\lambda}$ by deriving some $\lambda' > \bar{\lambda}$ that is both necessary and sufficient for the desired property. First,

$$\frac{\partial^2 U_2}{\partial e \partial \theta} = \left(1 + \frac{\lambda}{2}\delta\right) + \frac{\lambda}{2} \left(k(e, \theta) + \theta \frac{\partial k}{\partial \theta}\right) \quad (10)$$

Second,

$$\frac{\partial^3 U_2}{\partial e \partial \theta^2} = \frac{\lambda(1-e)}{(1-\theta e)^4} ((3-e)\theta - 2),$$

whose sign is determined by $(3-e)\theta - 2$ which is strictly increasing in θ . Thus, holding e fixed $\frac{\partial^2 U_2}{\partial e \partial \theta}$ is minimized at $\theta = \frac{2}{3-e}$. Substituting this into eqn. (10), the minimum value of the cross partial for a given (e, δ, λ) over all θ must therefore be,

$$\left(1 + \frac{\lambda}{2}\delta\right) - \frac{\lambda}{2} \left(\frac{8 + (1-e)}{27(1-e)}\right) \quad (11)$$

This is decreasing in e and therefore achieves its minimum over the range of interest at $e^*(1, \delta) = \lambda(1 + \frac{\lambda}{2}\delta)$. Rearranging and substituting in, it is therefore ≥ 0 i.f.f.

$$-18\lambda + 2\lambda^2 + \delta\lambda^3 - 27(2 + \delta\lambda)(\delta\lambda^2 + 2\lambda - 2) \geq 0 \quad (12)$$

Now we conjecture that if the condition is satisfied for some δ then it is also satisfied for all $\delta' < \delta$; if true it then suffices to check at $\delta = 1$. Substituting this in, the condition reduces to

$$54 - 36\lambda - 53\lambda^2 - 13\lambda^3 \geq 0$$

The l.h.s. is a cubic polynomial with a unique real root at $\lambda' \approx .687302 > \bar{\lambda}$. Finally, we check that our final conjecture is correct in the supplemental mathematica code to this document by verifying that eqn. (12) is $\geq 0 \forall \lambda < \lambda'$ and $\delta \in [0, 1]$; thus, λ' is both a necessary and sufficient bound for the cross partial to be positive $\forall (e, \theta, \delta)$ s.t. $\forall e \in [0, e^*(\theta, \delta)]$. ■

Lemma 4. *Whenever experimenting is an equilibrium of the subgame following initial policy $x^1 \in \{a, b\}$, then it is also the optimal (pure or mixed) strategy for the principal if she could precommit to her responses to success and failure.*

Proof: Because of symmetry we can restrict attention to the subgame following $x^1 = a$. Applying the characterization of the principal's utility in Appendix A and the analysis of the agent's best-response in the proof of Propositions 1 and 2, the principal's utility if she could precommit to her responses to success and failure (p_s^a, p_f^a) would be

$$U_1(e^*(\theta_2, p_s^a - p_f^a), \theta_1, \theta_2, p_s^a, p_f^a) \quad (13)$$

Recall that $U_1(\cdot)$ is strictly increasing in both first period effort e^1 and the probability of staying after success p_s^a holding the other parameters fixed. Since the agent's best response effort $e^*(\theta_2, p_s^a, p_f^a)$ is also increasing in p_s^a eqn. (13) must therefore be increasing in p_s^a , and the optimal choice after success is therefore to always stay, i.e. $p_s^{a*} = 1$. This feature is shared with the baseline model without commitment. Intuitively, the reason is that a higher probability of staying with the initial policy after success is both interim-better for the principal, and also better motivates the agent ex-ante.

The principal's optimal choice after failure p_f^{a*} satisfies

$$p_f^{a*} \in \arg \max_{p_f^a \in [0,1]} \{U_1(e^*(\theta_2, 1 - p_f^a), \theta_1, \theta_2, 1, p_f^a)\}$$

Now the principal's original utility function $U_1(\cdot)$ can be rewritten as,

$$\begin{aligned} & \theta_1 e^1 \cdot (1 + \lambda p_s^a) \\ & + (1 - \theta_1 e^1) \cdot \lambda (p_f^a \cdot h(e^1, \theta_1) h(e^1, \theta_2) + (1 - p_f^a) \cdot (1 - h(e^1, \theta_1)) (1 - h(e^1, \theta_2))) , \end{aligned}$$

and is easy to verify that this is decreasing in p_f^a whenever $h(e^1, \theta_1) \leq 1 - h(e^1, \theta_2)$, i.e., if posteriors after failure are s.t. it is better to switch. If experimenting is an equilibrium of the subgame $x^1 = a$ then by definition this property holds for $e^*(\theta_2, 1)$. In addition, recall that $U_1(\cdot)$ is increasing in e^1 and that $e^*(\theta_2, 1 - p_f^a)$ is decreasing in p_f^a . Combining these observations,

$$U_1(e^*(\theta_2, 1), \theta_1, \theta_2, 1, 0) > U_1(e^*(\theta_2, 1), \theta_1, \theta_2, 1, p_f^a) > U_1(e^*(\theta_2, 1 - p_f^a), \theta_1, \theta_2, 1, p_f^a)$$

whenever experimenting is an equilibrium. Consequently $(p_s^{a*} = 1, p_f^{a*} = 0)$, i.e. experimenting, is the optimal strategy with commitment. ■

Lemma 5. *If the principal prefers experimenting with x^1 to rigidly implementing $\neg x^1$ for λ , she also prefers it for $\lambda' > \lambda$.*

The principal's utility from rigidly implementing a is $V_1^R(\theta_1, \theta_2) = \theta_1 \cdot V_1^R(1, \theta_2)$ and from experimenting with b is $V_1^E(1 - \theta_1, 1 - \theta_2) = \theta_1 \cdot V_1^E(0, 1 - \theta_2) + (1 - \theta_1) \cdot V_1^E(1, 1 - \theta_2)$. She thus prefers experimenting with b to rigidly implementing a i.f.f.

$$\theta_1 < \frac{V_1^E(1, 1 - \theta_2)}{V_1^E(1, 1 - \theta_2) + (V_1^R(1, \theta_2) - V_1^E(0, 1 - \theta_2))}. \quad (14)$$

and symmetry, she prefers experimenting with a to rigidly implementing b i.f.f.

$$1 - \theta_1 < \frac{V_1^E(1, \theta_2)}{V_1^E(1, \theta_2) + (V_1^R(1, 1 - \theta_2) - V_1^E(0, \theta_2))}.$$

Thus, the desired property holds if the r.h.s. of eqn. (14) is increasing in $\lambda \in [0, \bar{\lambda}]$ for all $\forall \theta_2 \in [0, 1]$; we verify this property in the supplemental mathematica code. ■

Lemma 6 (Preference Monotonicity). *For all $(s, s') \in \{R, E\}^2$,*

$$V_1^s(\hat{\theta}_1, \theta_2) > V_1^{s'}(1 - \hat{\theta}_1, 1 - \theta_2) \rightarrow V_1^s(\theta_1, \theta_2) > V_1^{s'}(1 - \theta_1, 1 - \theta_2) \text{ for all } \theta_1 > \hat{\theta}_1.$$

In words, fixing the principal's experimentation decisions down each path of play, if she prefers $x^1 = a$ given beliefs $\hat{\theta}_1$ then she also prefers it for any higher belief.

Proof: Because the principal's expected utility for each first period policy is linear in her prior beliefs (holding her future experimentation decisions fixed), a nonmonotonicity would imply that $x^1 = b$ is better when $\omega = a$ and $x^1 = a$ is better when $\omega = b$. The former could not be true if she is rigidly implementing b since it would always fail, and the latter could not be true if she is rigidly implementing a . Thus, a nonmonotonicity requires that she be experimenting down both paths of play. To rule it out, it therefore suffices to show that experimenting with b is better than experimenting with a when $\omega = b$ (given the agent is predisposed to b , i.e. $\theta_2 \leq \frac{1}{2}$). This is,

$$(1 + \lambda) e^E(1 - \theta_2) + \lambda(1 - e^E(1 - \theta_2)) h(e^E(1 - \theta_2), 1 - \theta_2) > \lambda(1 - h(e^E(\theta_2), \theta_2))$$

Applying the definition, we know $e^E(1 - \theta) > \lambda(1 - \theta) \rightarrow$ the l.h.s. is $> (1 + \lambda) \lambda(1 - \theta_2)$. Also, $e^E(\theta_2) < \lambda\theta_2(1 + \lambda) \rightarrow$ the r.h.s. $< \frac{\lambda(1 - \theta_2)}{1 - \theta_2^2 \lambda(1 + \lambda)}$. The above inequality will thus hold

when

$$(1 + \lambda) \lambda (1 - \theta_2) > \frac{\lambda (1 - \theta_2)}{1 - \theta_2^2 \lambda (1 + \lambda)} \iff (1 + \lambda)^2 < \frac{1}{\theta_2^2} \leq 4 \iff \lambda < 1,$$

and therefore always holds for $\lambda < \bar{\lambda} < 1$. ■

Lemma 7. *For all θ_2 s.t. $1 - \bar{\theta}(1 - \theta_2) > \frac{1}{2}$, the principal weakly prefers rigidly implementing b to experimenting with a when $\theta_1 = 1 - \bar{\theta}(1 - \theta_2)$.*

Proof: Since $1 - \bar{\theta}(1) = 1$ and $1 - \bar{\theta}(\frac{1}{2}) < \frac{1}{2}$, there exists some unique $\hat{\theta}_2$ satisfying $1 - \bar{\theta}(\hat{\theta}_2) = \frac{1}{2}$ s.t. the set of beliefs $[\frac{1}{2}, 1 - \bar{\theta}(1 - \theta_2)]$ where the principal would rigidly implement b in that subgame is nonempty. Note that $\hat{\theta}_2$ is a function of λ so we henceforth write $\hat{\theta}_2(\lambda)$ for clarity. Now applying the definitions and rearranging, we wish to show that

$$\forall \lambda \in [\underline{\lambda}, \bar{\lambda}],$$

$$\bar{\theta}(1 - \theta_2) \cdot (V_1^R(1, 1 - \theta_2) - V_1^E(0, \theta_2)) > (1 - \bar{\theta}(1 - \theta_2)) \cdot V_1^E(1, \theta_2) \quad \forall \theta_2 \in [0, \hat{\theta}_2(\lambda)],$$

where $\underline{\lambda} \approx .23505$. We verify this step in the supplemental mathematical code. ■

Lemma 8. *The principal strictly prefers experimenting with b to experimenting with a when $\theta_1 = 1 - \theta_2$ and $\theta_2 < \frac{1}{2}$.*

Proof: Let $\phi(\theta_2) = \lambda \theta_2 (1 + \frac{\lambda}{2} (k(e^E(\theta_2), \theta_2) + 1)) (1 + \lambda) - \lambda (1 - h(e^E(1 - \theta_2), 1 - \theta_2))$ be the principal's utility difference between experimenting with a and experimenting with b when $\omega = a$. Applying symmetry, her expected utility difference between experimenting with a and experimenting with b with prior θ_1 is $\theta_1 \phi(\theta_2) - (1 - \theta_1) \phi(1 - \theta_2)$. We now wish to show that this is < 0 when $\theta_1 = 1 - \theta_2$ given that $\theta_2 < \frac{1}{2}$, i.e.

$$(1 - \theta_2) \phi(\theta_2) - \theta_2 \phi(1 - \theta_2) < 0 \iff \frac{\phi(\theta_2)}{\lambda \theta_2} - \frac{\phi(1 - \theta_2)}{\lambda (1 - \theta_2)} < 0 \text{ when } \theta_2 < \frac{1}{2}.$$

It is simple to verify that

$$\frac{\phi(\theta_2)}{\lambda\theta_2} = \left(1 + \frac{\lambda}{2} (k(e^E(\theta_2), \theta_2) + 1)\right) (1 + \lambda) - \frac{1}{1 - e^E(1 - \theta_2) \cdot (1 - \theta_2)},$$

and by symmetry it suffices to show that $\frac{\phi(\theta_2)}{\lambda\theta_2} - \frac{\phi(1-\theta_2)}{\lambda(1-\theta_2)} > 0$ when $\theta_2 > \frac{1}{2}$. Using substitution and rearranging, we then have that $\frac{\phi(\theta_2)}{\lambda\theta_2} - \frac{\phi(1-\theta_2)}{\lambda(1-\theta_2)} > 0 \iff$

$$\frac{\theta_2 e^E(\theta_2) - (1 - \theta_2) e^E(1 - \theta_2)}{(1 - \theta_2 e^E(\theta_2)) \cdot (1 - (1 - \theta_2) e^E(1 - \theta_2))} > \frac{\lambda}{2} (1 + \lambda) (k(e^E(1 - \theta_2), 1 - \theta_2) - k(e^E(\theta_2), \theta_2)) \quad (15)$$

Since the denominator of the l.h.s. is < 1 , and $1 + \lambda < 2$, the above inequality holds if the following yet stronger inequality holds,

$$\theta_2 e^E(\theta_2) - (1 - \theta_2) e^E(1 - \theta_2) > \lambda (k(e^E(1 - \theta_2), 1 - \theta_2) - k(e^E(\theta_2), \theta_2)). \quad (16)$$

Now again substituting in the definition of $e^E(\cdot)$, the l.h.s. can be rewritten

$$\lambda \left((2\theta_2 - 1) \left(1 + \frac{\lambda}{2}\right) + \frac{\lambda}{2} (\theta_2^2 k(e^E(\theta_2), \theta_2) - (1 - \theta_2)^2 k(e^E(1 - \theta_2), 1 - \theta_2)) \right),$$

implying that the desired inequality holds i.f.f.

$$(2\theta_2 - 1) \left(1 + \frac{\lambda}{2}\right) > \left(1 + \frac{\lambda}{2} (1 - \theta_2)^2\right) k(e^E(1 - \theta_2), 1 - \theta_2) - \left(1 + \frac{\lambda}{2} \theta_2^2\right) k(e^E(\theta_2), \theta_2)$$

It is easily verified that the r.h.s. $< (1 + \frac{\lambda}{2}) (k(e^E(1 - \theta_2), 1 - \theta_2) - k(e^E(\theta_2), \theta_2))$ when $\theta_2 > \frac{1}{2}$, implying that the above inequality holds if the following yet stronger inequality holds:

$$2\theta_2 - 1 > k(e^E(1 - \theta_2), 1 - \theta_2) - k(e^E(\theta_2), \theta_2) \quad (17)$$

We now show that eqn. (17) holds. Substituting in the definition of $k(\cdot)$, observing that

$2\theta_2 - 1 = \theta_2^2 - (1 - \theta_2^2)$, and rearranging, we have that the inequality is equivalent to

$$\theta_2^2 \left(1 - \frac{1}{(1 - (1 - \theta_2) e^E (1 - \theta_2))^2} \right) > (1 - \theta_2)^2 \left(1 - \frac{1}{(1 - \theta_2 e^E (\theta_2))^2} \right)$$

Since $e^E(\theta_2)$ is increasing in θ_2 , this clearly holds because $(1 - \theta_2) e^E (1 - \theta_2) < \theta_2 e^E (\theta_2)$ when $\theta_2 > \frac{1}{2}$. The desired property is hence shown. ■

Lemma 9. *The following three properties hold.*

1. When $\lambda > \hat{\lambda}$, $V_1^E(0, 1 - \theta_2) - V_1^R(1, \theta_2) > 0$ for θ_2 in a nonempty interval $[0, \varepsilon_1]$.
2. When $\lambda \in [0, 1]$, $V_1^R(1, \theta_2) - V_1^E(1, \theta_2) > 0$ for θ_2 in a nonempty interval $[0, \varepsilon_2]$.
3. When $\lambda \in [0, 1]$, $V_1^R(\bar{\theta}(1 - \theta_2), 1 - \theta_2) - V_1^E(1 - \bar{\theta}(1 - \theta_2), \theta_2) > 0$ for θ_2 in a nonempty interval $[0, \varepsilon_3]$.

To show that some function $f(\theta_2)$ satisfying $f(0) = 0$ is > 0 for $\theta_2 \in (0, \varepsilon)$ where $\varepsilon > 0$, it suffices to show by continuity that $\left. \frac{f(\theta_2)}{\theta_2} \right|_{\theta_2=0} > 0$ provided that this quantity is finite. We now show this for each of the desired expressions.

Property 1: When $\omega = a$, the principal's utility from experimenting with b is,

$$\lambda(1 - h(e^E(1 - \theta_2), 1 - \theta_2)) = \frac{\lambda\theta_2}{1 - (1 - \theta_2)e^E(1 - \theta_2)},$$

and from rigidly implementing a is,

$$e^R(\theta_2)(1 + \lambda) + (1 - e^R(\theta_2))h(e^R(\theta_2), \theta_2) = \lambda\theta_2 \left((1 + \lambda) \left(1 + \frac{\lambda}{2} k(e^R(\theta_2), \theta_2) \right) + \frac{(1 - e^R(\theta_2))^2}{1 - \theta_2 e^R(\theta_2)} \right)$$

Since $e^R(0) = 0$, $e^E(1) = \lambda(1 + \frac{\lambda}{2})$, and $k(e^R(0), 0) = 1$,

$$\left. \frac{1}{\lambda\theta_2} (V_1^E(0, 1 - \theta_2) - V_1^R(1, \theta_2)) \right|_{\theta_2=0} = \left(\frac{1}{1 - \lambda(1 + \frac{\lambda}{2})} \right) - \left((1 + \lambda) \left(1 + \frac{\lambda}{2} \right) + 1 \right).$$

Manipulating the above expression demonstrates that it is ≥ 0 i.f.f.

$$\lambda(1 + \lambda) \left(1 + \frac{\lambda}{2}\right) > 1. \quad (18)$$

which holds i.f.f. $\lambda > \hat{\lambda}$ by definition.

Property 2: When $\omega = a$, the principal's utility from experimenting with a is,

$$e^E(\theta_2)(1 + \lambda) = \lambda\theta_2 \left(1 + \frac{\lambda}{2} (k(e^E(\theta_2), \theta_2) + 1)\right) (1 + \lambda),$$

Since $k(e^E(0), 0) = 1$ and applying Part 1, the desired condition is equivalent to

$$(1 + \lambda) \left(1 + \frac{\lambda}{2}\right) + 1 > (1 + \lambda)^2$$

which holds if $\lambda(1 + \lambda) < 2 \iff \lambda < 1$.

Property 3: Using the definition from the proof of Proposition 3, the threshold function $\bar{\theta}(\theta_2)$ in closed form is $\bar{\theta}(\theta_2) = \frac{1-\theta_2}{(1-\theta_2)+\theta_2(1-e^E(\theta_2))^2}$. Thus, when $\theta_1 = 1 - \bar{\theta}(1 - \theta_2)$ the difference in the principal's utility between rigidly implementing b and experimenting with a is,

$$\bar{\theta}(1 - \theta_2) \cdot (V_1^R(1, 1 - \theta_2) - V_1^E(0, \theta_2)) - (1 - \bar{\theta}(1 - \theta_2)) \cdot V_1^E(1, \theta_2) \quad (19)$$

The first term is the subjective probability that $\omega = b$ times the utility difference between rigidly implementing b and experimenting with a when $\omega = b$. The second term is the subjective probability that $\omega = a$ times the utility of experimenting with a when $\omega = a$ (the payoff from rigidly implementing b when the state is a is 0).

Now to get the desired expression we divide through by θ_2 and then evaluate at $\theta_2 = 0$ – we do so by dividing $\bar{\theta}(1 - \theta_2)$ by θ_2 in the first term, then $V_1^E(1, \theta_2)$ by θ_2 in the second

term, and then evaluating all expressions at $\theta_2 = 0$. First,

$$\left(\frac{1}{\theta_2}\right) \cdot \bar{\theta}(1 - \theta_2) = \frac{1}{\theta_2 + (1 - \theta_2)(1 - e^E(1 - \theta_2))^2}$$

which is equal to $(1 - \lambda(1 + \frac{\lambda}{2}))^{-2}$ at $\theta_2 = 0$. Second, $V_1^R(1, 1 - \theta_2) = 2\lambda$, $V_1^R(0, \theta_2) = \lambda$, and $1 - \bar{\theta}(1 - \theta_2) = 1$ at $\theta_2 = 0$. Third, applying part 2 we know $\left(\frac{1}{\theta_2}\right) \cdot V_1^E(1, \theta_2) = \lambda(1 + \frac{\lambda}{2}(k(e^E(\theta_2), \theta_2) + 1))(1 + \lambda)$ which is $= \lambda(1 + \lambda)^2$ evaluated at $\theta_2 = 0$. Assembling these observations, the desired expression is

$$\left(1 - \lambda\left(1 + \frac{\lambda}{2}\right)\right)^{-2} \cdot (2\lambda - \lambda) - \lambda(1 + \lambda)^2 > 0 \iff \left(1 - \lambda\left(1 + \frac{\lambda}{2}\right)\right)^2 \cdot (1 + \lambda)^2 < 1$$

Now the l.h.s. is $< (1 - \lambda)^2 \cdot (1 + \lambda)^2 = (1 - \lambda^2)^2 \leq 1 \forall \lambda \in [0, 1]$, so the result is shown. ■