

Wait-and-See or Step in? Dynamics of Interventions*

Dana Foarta Takuo Sugaya
Stanford GSB Stanford GSB

September 24, 2018

Abstract

We study when and how intervention to stop a project is optimally used in a repeated relationship between a principal and a policymaker. The policymaker is privately informed about his ability, where a higher ability policymaker has a lower cost of producing a good project. He also privately chooses how much effort to supply on the project. Before the project is completed, the principal receives a signal about its outcome and can intervene to stop it from taking effect. Intervention may prevent a bad outcome, but no intervention leads to better learning about the policymaker's ability. In the benchmark with observable effort, it is optimal to intervene only when the policymaker's reputation is sufficiently low. If effort is not observable, the optimal response features switching between intervention and no intervention on the equilibrium path. The model rationalizes intervention technologies implemented in practice by supranational agreements and governing coalitions.

JEL Classification Codes: D78, D82, H77.

*Foarta: Stanford Graduate School of Business, 655 Knight Way, Stanford, CA, 94305, tel: (650) 723-4058, email: ofoarta@stanford.edu. Sugaya: Stanford Graduate School of Business, 655 Knight Way, Stanford, CA, 94305, tel: (650) 724-3739, email: tsugaya@stanford.edu. A previous version of this paper circulated under the title "Dynamic Responses to Early Warnings." We thank Avidit Acharya, Peter Buisseret, Steve Callander, Ken Shotts, Andy Skrzypacz, Victoria Vanasco, Richard van Weelden, as well as seminar participants at Arizona State, Boston University, Caltech, LSE, Stanford, University of Arizona, University of Warwick, Yale and Washington PECO for useful comments and suggestions on earlier versions of this paper. Foarta thanks the Princeton University JRCPPF and CSDP for financial support.

1 Introduction

Most institutions include specific mechanisms for interventions to overrule a policymaker and stop an ongoing policy project from continuing to completion. For instance, an international organization may pull the funds offered for a development project in response to bad audit results of its intermediate progress, a government's reform proposal may be revised through legislative review in parliament, a regulator's proposed rule may be overruled by an oversight agency, or an investor may force the early liquidation of a project. Such mechanisms are put in place in order to allow for course corrections if information emerges that the original path of action is no longer desirable. Policy change may be undertaken by overruling the policymaker rather than by removing him from his position, even when removal is also an option. In fact, intervention is usually more likely than outright removal of a policymaker.¹

Deciding whether to intervene involves a trade-off between cost saving and information acquisition. Pulling the plug on a project saves resources from being potentially wasted on a faulty project; however, once a project is stopped, its originally intended outcome is never seen. An intervention means policy is changed, reforms are interrupted, funds are pulled, and the observable events differ from what the policymaker set out to do. Therefore, while useful in achieving immediate course corrections, interventions may hamper information acquisition about policy or policymaker effectiveness. Given the ubiquitous use of interventions, an open and fundamental question in the design of institutions is how and when interventions should be used. In this paper, we propose a model to shed light on this question.

We focus on the problem for a principal who repeatedly sponsors projects that are run by an agent. Every period, the agent works on one project, the outcome of which may yield either a benefit or a loss for the principal at the end of that period. Before the project is completed, it produces a signal about its outcome. A good signal indicates a high likelihood of a good outcome. A bad signal indicates a potential roadblock that, unless overcome, will result in a bad outcome. After observing the signal, the principal may intervene by paying

¹For example, one can see this by comparing the data in [Martin and Vanberg \(2004\)](#) on interventions through legislative reviews with data on motions of censure that remove a minister or government.

a cost to stop the current project. If stopped, the project is abandoned, and its outcome is not revealed. Otherwise, the project is completed, and its outcome is revealed.

Our focus on such relationships and intervention technologies is motivated by several applications. A prime example is that of supranational agreements in which transfers are made by a supranational institution (e.g., the World Bank, the European Commission) towards policies implemented by a country's local policymaker. In the European Union, for instance, Structural and Investment funds are provided on a recurring basis by the European Commission (EC), as transfers to support projects run by local governments. The implementation of these projects is verified periodically, and the EC may respond to deficiencies by revising projects or suspending their funding. Another example comes from governing coalitions in parliamentary democracies. There, legislative projects championed by party leaders are implemented by the government minister in that respective policy area.² As shown in studies of coalition governments in parliamentary systems,³ parliamentary committees that correspond to ministerial jurisdictions are set up by the coalition parties as intervention vehicles. These committees may schedule hearings, gather information, and propose amendments to legislation produced by the minister. They are meant to ensure that the minister's course of action is in line with the coalition's agenda. Finally, in the case of private companies, an investor may intervene to stop funding for or to demand restructuring of an entrepreneur's project before it is completed.

The outcome of the project depends on the effort the agent exerts to set it up. More effort increases the likelihood of producing a benefit for the principal. Moreover, only an agent who has exerted effort may overcome a roadblock indicated by the principal's signal. Yet, both the effort supplied and the agent's cost of supplying effort are unobservable to the principal. The agent may be either a "high ability" type, who faces an increasing marginal cost of effort, or an "inept" type who cannot exert effort. This captures, for instance, the ability of a policymaker to successfully adapt a project to local conditions. Costly effort

²See [Laver and Shepsle \(1996\)](#) for an overview of policymaking in parliamentary democracies.

³For instance, [Martin and Vanberg \(2004, 2005\)](#) provide evidence of parliamentary review as an intervention mechanism, with examples from Germany and the Netherlands.

in our model captures, in reduced form, local electoral costs of implementing a potentially unpopular project or the agent's personal cost of working on the principal's project rather than his own ideal project.⁴ The principal would like the agent to be both able to deliver a beneficial project and willing to work on this project. Based on the signal and the information revealed at the end of the period, the principal updates her belief about the agent's ability. At this point, she may start a new project with the same agent or end the relationship with the agent and invest in a new project with another agent (or she may take an outside option).

The repeated nature of the relationship between the principal and the agent captures a key consideration present in each of our motivating examples. The EC repeatedly funds projects in the same policymaker's jurisdiction. If failures on projects go uncorrected, future funding to that policymaker may be suspended indefinitely, and the EC may use those funds for other projects in a different jurisdiction or under the purview of a different policymaker. The conditions for verifying projects, triggering interventions or suspending funding have been a focus of policy changes in recent years⁵ and constitute grounds for broader debates.⁶ In the case of governing coalitions, the relationship between ministers and parliamentary committees (controlled by coalition party leaders) has a repeated nature, and it is instrumental for the governing coalition's stability. Finally, investors and entrepreneurs are oftentimes linked by repeated investments in subsequent ventures.

The model repeats the above sequence of actions each period over an infinite horizon. To deliver results relevant for the motivating examples from the political sphere, we focus on the case in which the principal cannot monetarily compensate the agent. She must instead rely on the intervention decision and on continuation versus termination of the relationship as the only tools for incentives provision.⁷

⁴Politicians must be motivated to implement reforms which might be electorally costly for them. Similarly, investors must motivate entrepreneurs to steer the firm in a direction that will result in profits, rather than letting them pursue a broader scope, with a more uncertain payoff.

⁵The reporting requirements and intervention policies were reformed in 2013, with application to the 2014-2020 period: http://ec.europa.eu/regional_policy/sources/docgener/guides/blue_book/blueguide_en.pdf

⁶Proposals have called for creating broader rules for suspensions of funds (for instance, http://www.cer.eu/sites/default/files/pbrief_structural_funds_nov17.pdf, accessed May 20, 2018)

⁷This assumption also applies in many private sector examples, either due to the use of efficiency wages or in cases where monetary compensation carries too little weight in the agent's utility function for it to be

We characterize the best Perfect Bayesian Equilibrium for the principal (in Section 5, we show that our results are robust to considering the Markov Perfect Equilibrium instead). The core trade-off at the basis of this model is that intervention brings about the benefit of avoiding a bad outcome and the cost of not learning more about the agent’s type from observing the project’s outcome. In examining this trade-off, the principal is facing two unknowns: the agent’s ability to supply effort (a selection problem) and his effort (a control problem). To show the role that each of these two factors plays in the results, we first shut down the control problem and assume that the high ability agent always supplies the highest effort. The problem then becomes a bandit problem for the principal, which delivers a sharp characterization of the optimal policy: it has a cutoff structure. The expected benefit of intervention linearly decreases in the principal’s belief about the agent’s type. The cost of learning as a function of the agent’s type is single peaked and reaches its highest value when the uncertainty about the agent’s type is highest. Combining these two, we find that intervention is optimal below a threshold belief about the agent’s type, and it is not optimal above this threshold. This leads to a simple institutional implementation: intervention is used after a bad signal only while the agent’s reputation is sufficiently low.

This result rationalizes, for instance, observed intervention practices by international organizations for projects over which they have no major preference disagreements with local policymakers. In such cases, the main concern for the international organization is the ability of policymakers to adapt projects to local conditions. International organizations have been documented to start by funding projects that are easily monitored and may be easily stopped in case of negative audit reports, e.g., infrastructure projects. When the local policymaker’s reputation becomes high, international lenders become more likely to offer funds for government projects which are more difficult to monitor.⁸ As predicted by the model, the latter type of projects are likely to be funded when intervention is not optimal.

The simple threshold result is upended when considering the agent’s effort choice. To motivate the agent to exert effort, the principal must use the promise of higher rewards for an effective means of incentive provision.

⁸See, for instance, [Winters \(2010\)](#) for a discussion of the different types of aid offered by the World Bank.

the agent after outcomes which indicate that more effort was exerted. After observing such outcomes, the principal also positively updates her belief about the agent's type. Thus, on the equilibrium path, the principal's belief about the agent and the reward promised to the agent tend to move in the same direction — both increase or both decrease. This insight leads to two main implications. First, if the belief about the agent's type drops to a sufficiently low level, the agent is replaced on the equilibrium path. Second, intervention after bad news is optimal if the belief about the agent is below a low threshold or above a high threshold. When the belief is below the low threshold, intervention is optimal in order to address the high likelihood that the principal is facing a low type agent. When the principal's belief is very high, so is the reward promised to the agent. Intervention becomes optimal because it is too expensive to motivate the agent to exert effort, and there is no large benefit to learning more about the agent's type. In between the low and the high thresholds, there is at least one region where it is optimal not to intervene. This happens because the benefit of learning is high, while the cost of providing rewards, to incentivize effort, is not too high.

An immediate implication of the above results is the emergence of switches between periods of intervention and periods of no intervention after bad news. It also captures the two distinct situations in which intervention is necessary, which correspond to the trade-off between selection and control: either the agent is willing, but likely unable to run the project, or the agent is likely able, but unwilling given the incentives offered. In the intermediate situation, there is sufficient likelihood for both ability and willingness. Learning from the observed outcomes changes the principal's evaluation of the likelihood of these situations, leading to switches in the intervention policy on the equilibrium path. This dynamic implies that simple rules based on reaching a certain reputation no longer implement the optimal policy. Intervention becomes a recurring phenomenon.

This result rationalizes audit and intervention technologies used in practice by international organizations when incentivizing effort by local policymakers is also crucial. This is the case, for example, when considering electorally unpopular reforms projects, over which there are clear preference disagreements between international lenders and local policymak-

ers. The result of optimal switches between intervention and no intervention is consistent with the pattern of ongoing auditing and potential intervention regardless of the policymaker's reputation. Moreover, our dynamics explain the observed decay over time in the implementation of reforms demanded by international lenders. The optimal contract calls for more discretion to be given to local policymakers to maintain the status quo, leading to a slowdown over time in reforms.

For our application to governing coalitions, the results rationalize the empirical evidence that parliamentary intervention through legislative review happens more often when there are larger preference disagreements within the government coalition.⁹ We can map preference disagreements to a higher probability that the coalition breaks up, and thus to more discounting of the future. Performing this comparative static, we show that this increases the value of intervention, making intervention more likely on the equilibrium path.

Related Literature. Our work blends two main strands of the political economy literature, one that has studied dynamic incentive provision under adverse selection¹⁰ and one that has examined the control problem.¹¹ We combine these two strands in the context of institutional relationships with a policymaker. This allows us to study the use of intervention power by institutions in their relationships with politicians who implement policies. Our approach has commonalities with the work of [Banks and Sundaram \(1993, 1998\)](#) on electoral accountability. We augment their setup by considering the case where both the principal and the agent are long-lived, and, central to our focus on institutional relationships, we introduce the action of intervention, which reduces learning and changes the dynamics of the problem. Moreover, we focus on the implications for intervention policy in the Perfect Bayesian Equilibrium, and we also provide results for the Markov Perfect Equilibrium. More broadly, our paper relates to the literature on dynamic incentives provision under both moral hazard and adverse selection ([Strulovici, 2011](#); [Halac et al., 2016](#)). We focus, however, on contracts in

⁹Provided in [Martin and Vanberg \(2004, 2005\)](#).

¹⁰See [Rogoff \(1990\)](#); [Besley and Case \(1995\)](#); [Coate and Morris \(1995\)](#); [Besley and Smart \(2007\)](#)

¹¹See [Barro \(1973\)](#); [Ferejohn \(1986\)](#); [Austen-Smith and Banks \(1989\)](#); [Ray \(2002\)](#); [Acemoglu et al. \(2008\)](#)

which the agent cannot be monetarily compensated, and on a particular signal and action structure motivated by our applications.

The focus on interventions links our paper to the literature on oversight and transparency (Aghion and Tirole, 1997; Prat, 2005; Gavazza and Lizzeri, 2007; Fox and Van Weelden, 2010, 2012; Buisseret, 2016). Levitt and Snyder (1997) examine a similar type of intervention, in a static model with moral hazard. We consider a dynamic context, with moral hazard and adverse selection, both of which are essential for our result of switches between intervention and no intervention on the equilibrium path.

Our paper is also related to the reputation literature,¹² and specifically the application to the reputation of governments (Herrera et al., 2014). We follow a standard approach in this literature and differentiate between an inept agent type, who in our case cannot exert effort to influence the final outcome, and a strategic agent type, who must be motivated to supply effort towards the principal's project. Our results have a similar intuition for the dynamics generated by moral hazard as in Ray (2002) and Acemoglu et al. (2008). The difference in our setup from these models is two-fold: the presence of adverse selection along with moral hazard, and that the only ways to provide rewards to the agent are by not removing him and through the intervention policy.

The result of switches between intervention and no intervention after bad news on the equilibrium path also links this work to the literature on policy cycles (Ales et al., 2014; Dovis et al., 2016). Our result of changing responses over time after bad news relies, however, on a distinct mechanism, coming from the interaction between the principal's ability to learn about the agent's type and the agent's incentives to supply effort.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 analyzes the benchmark case with selection only. Section 4 analyzes the full model and provides comparative statics. Section 5 shows the robustness of our result to using a Markov Perfect equilibrium concept instead. Section 6 discusses applications and positive implications of the model, Section 7 concludes, and the Appendix contains the proofs. Additional online

¹²Starting with Kreps and Wilson (1982); Mailath and Samuelson (2001, 2006).

appendices contain a three-period version of the model used to derive analytical results for the comparative statics, and the analysis of the Markov Perfect Equilibrium.

2 The Model

We consider an infinite-horizon discrete time environment with two players: a principal (P) and an agent (A). The agent has a type $\theta \in \{L, H\}$, which is his private information, where type H occurs with commonly known probability μ_H . Every period, the agent works on a project that, if completed, provides an outcome y for the principal. This outcome can be either **Good** or **Bad**. After the project is started and before its outcome is realized, it produces a noisy public signal $s \in \{g, b\}$ about this outcome. It provides information about how likely it is for the project to be completed successfully, but it is not a perfect indicator of the final outcome.

The project's outcome y is a function of the unobservable effort e exerted by the agent. Effort $e \in [0, 1]$ is exerted at the start of the project, and it comes at a cost $c(e)$ to the H -type agent, with $c(0) = 0$, $c'(0) = 0$, $\lim_{e \rightarrow 1} c'(e) = \infty \forall \theta$,¹³ and

$$c'(e) > 0, c''(e) > 0 \forall e > 0.$$

For the L -type agent, following the reputation literature,¹⁴ we assume that this is an inept type, who cannot exert any effort on the principal's project. We make this assumption in order to focus on the optimal incentive structure for the H -type.¹⁵ We further discuss our assumptions about effort in our applications in Section 4.

Each possible combination of a signal and a project outcome, (s, y) , occurs with some probability $\Pr((s, y) | e)$ given effort e . We make the following assumption about the proba-

¹³We assume this limit to ensure that the first order condition is necessary and sufficient for the effort provision. In some of our examples, to simplify calculations, we may consider a quadratic cost function or a binary effort choice. There, we take corner solutions into account.

¹⁴See [Mailath and Samuelson \(2001\)](#).

¹⁵In particular, we do not introduce the payoff function, incentive compatibilities, and other variables/constraints for the L -type.

bility distribution of (s, y) :

Assumption 1 *The probability distribution over (s, y) has the following properties:*

1. $\Pr((b, G) | 0) = 0$ and $\Pr((s, y) | e) > 0$ otherwise, $\forall e \in [0, 1]$;
2. $\Pr((s, B) | e)$ and $\Pr(s = b | e) \equiv [\Pr((b, B) | e) + \Pr((b, G) | e)]$ are decreasing and convex in e ;
3. $\Pr((s, G) | e)$ is increasing and concave in e .

The first property assumes that a G outcome happens after a b signal only if effort is exerted: the b signal indicates a roadblock that can only be potentially overcome if effort was exerted in setting up the project.¹⁶ The monotonicity and concavity/convexity assumptions ensure that a b signal or a B outcome indicate lower effort, while a g signal or a G outcome indicate higher effort. Moreover, observing the final outcome of the project is informative conditional on the signal that is generated before project completion.

After observing signal s , the principal chooses whether to intervene (denoted by ι). If she intervenes ($\iota = 1$), she pays a cost l , the project is stopped, and its final outcome is not reached nor observed. The cost l could be a liquidation cost, or the cost of reversing a policy to its original state. If the principal does not intervene ($\iota = 0$), then she pays no cost at the intermediate stage, the project continues to completion, its outcome $y \in \{G, B\}$ is observed by everyone, and the principal pays a cost C if the outcome is B .

At the end of the period, the principal updates her belief about the agent's type, based on the public history of the game up to that point. To ease exposition, we denote the observable end to the project as $o \in \{I, G, B\}$, where I stands for an intervention having occurred, and hence no outcome being observed. Finally, at the beginning of the following period, before any other actions are taken, the principal decides whether to keep the agent ($\rho = 1$) or to end their relationship ($\rho = 0$). If the relationship is ended, the principal accesses an outside

¹⁶This assumption is only used in Proposition 1 in order to simplify the proof of the result, but we can show the same result for $\Pr((b, G) | 0) = \varepsilon > 0$, where ε is sufficiently small. Details are available upon request.

option. We focus on the case in which the principal's outside option is to start a new contract with another agent selected from a pool of agents where the probability of selecting a type H is μ_H . We consider this to be a natural continuation in many of the applications of this model. For instance, a government minister who is removed through a motion of censure will have to be replaced by another minister. In some applications of our model, however, an exogenous outside option may be a more natural continuation. For instance, an investor may deposit her funds in a bank if the relationship with the entrepreneur is ended, or a supranational institution may keep its funds in a reserve fund. As our analysis will make clear, assuming an exogenous outside option for the principal does not change our qualitative results. We therefore choose to focus our analysis on the more complex problem with agent replacement.

Payoffs. The Principal aims to maximize her expected payoff, where the per-period payoff takes the following form:

$$u = \begin{cases} -l & \text{if } o = I, \\ 0 & \text{if } o = G, \\ -C & \text{if } o = B, \end{cases} \quad (1)$$

where $l, C \in \mathbb{R}$, $0 < l < C$. Replacing the agent and drawing a new agent comes at no cost for the principal. This specification captures situations in which early, preventive intervention is less costly than letting the situation potentially worsen. Yet, an early intervention might stop a project from reaching a successful outcome.

Given a belief μ that the agent is of type H , the effort e exerted by the H -type agent, and the intervention decision $\iota(s) \in \{0, 1\}$ after signal s , the principal's expected utility after observing signal s is given by:

$$u^P(\iota|\mu, e, s) = \begin{cases} -l & , \text{ if } \iota(s) = 1, \\ \Pr(o = B|\mu, e, s) \cdot (-C) & , \text{ if } \iota(s) = 0. \end{cases} \quad (2)$$

By Bayes' Rule,¹⁷

$$\Pr(o = B|\mu, e, s) = \frac{\Pr((s, B)|e) \cdot \mu + \Pr((s, B)|0) \cdot (1 - \mu)}{\Pr(s|e) \cdot \mu + \Pr(s|0) \cdot (1 - \mu)}.$$

Each period, the H -type agent derives a fixed rent based on whether he is kept or not, and he pays the cost of effort as well:

$$u(e) = \begin{cases} 1 - c(e) & \text{if } \rho = 1, \\ 0 & \text{if } \rho = 0, \end{cases} \quad (3)$$

where $\rho = 1$ indicates that the agent is kept, and $\rho = 0$ indicates that the agent is removed. This payoff form corresponds to a purely office-motivated policymaker, who obtains a fixed benefit from being in office (or in the supranational agreement).¹⁸ Outside the political realm, this utility form has the interpretation of an entrepreneur who receives an efficiency wage and a non-pecuniary benefit from working on the project (an empire-building rent). Both the principal and the agent discount the future at rate δ .

Finally, we make the following assumption regarding the principal's costs l and C :

Assumption 2 *The following conditions are satisfied:*

1. *After $s = g$,*

$$\frac{\Pr((g, B)|e)}{\Pr((g, G)|e) + \Pr((g, B)|e)} < \frac{l}{C}, \quad \forall e \in [0, 1]. \quad (4)$$

2. *With $e = 1$,*

$$\frac{\Pr((b, B)|1)}{\Pr((b, G)|1) + \Pr((b, B)|1)} < \frac{l}{C}. \quad (5)$$

¹⁷In general, we write $\Pr(\cdot|\mu, e)$ to denote the conditional probability given that the H -type exerts effort e , while the L -type exerts no effort (by assumption).

¹⁸We assume for simplicity that the agent does not derive a direct benefit from the principal's project. It is worth noting also that alternative forms of utility that include both the fixed rent and the project's payoff yield similar qualitative results. Moreover, similar qualitative results can be obtained even if we allowed for payments to the agent (a rent schedule), as long as the utility function is sufficiently concave. Details available upon request.

Under (4), the probability of $y = B$ occurring after $s = g$ is sufficiently small (regardless of effort), so that it is statically optimal for the principal not to intervene after a g signal. Thus, $\iota(s = g) = 0$, and the intervention decision we are interested in is the decision ι after $s = b$. Similarly, under (5), $\iota(s = g) = 0$ if the principal knows that the agent is an H -type who exerts the highest level of effort.

3 Benchmark without Effort Choice

We begin by analyzing a natural benchmark of our model, given the two unobservable components for the principal: the agent's type and his effort choice. We shut down effort choice and assume effort is fixed for each agent type. We assume that type H supplies effort $e = 1$ at no cost, and type L supplies effort $e = 0$. The benchmark is therefore a selection problem for the principal. It captures the case in which the principal and the H -type agent are aligned in their incentives to achieve a successful project. The difference between agents reduces to the ability to carry through with the project. Later, we will contrast this benchmark case with the full model in which the principal must also worry about the agent's incentives to work on the project.

The game becomes a planning problem for the principal, as the agent does not take any meaningful action. Let $\mu \in [0, 1]$ be the probability that the agent is an H -type given the public history (the principal's *belief*). Given μ , the principal's problem can be expressed recursively, as a function $J(\mu)$ that depends on the principal's current belief. If the principal removes the current agent, the game restarts with a new agent who is expected to be a type H with probability μ_H . Hence, P 's utility after replacing A , denoted by \bar{J} , should satisfy $\bar{J} = J(\mu_H)$.¹⁹

To simplify notation, we denote by $\omega \equiv (s, o)$ the public events observed within each

¹⁹If the outside option were exogenous, it would be represented by a fixed value \bar{J} .

period, where $s \in \{g, b\}$ and $o \in \{I, G, B\}$. The principal's payoff can then be expressed as

$$J(\mu) = \max_{\substack{\rho \in \{0,1\}, \\ (\iota(s))_s \in \{0,1\}^2}} (1 - \rho) \cdot \bar{J} + \rho \cdot \left[\sum_{s \in \{g,b\}} \Pr(s|\mu) \cdot u^P(\iota|\mu, s) + \delta \cdot \sum_{\omega} \Pr(\omega|\mu, \iota) \cdot J(\mu'(\mu, \omega)) \right], \quad (6)$$

where $u^P(\iota|\mu, s)$ is given in (2), but here we omit e since it is always equal to 1; and $\mu'(\mu, \omega)$ is the updated belief given the prior belief μ and the public event $\omega = (s, o)$.

We first establish the basic properties of the value function $J(\mu)$:

Lemma 1 *$J(\mu)$ is increasing and convex in μ .*

The value function $J(\mu)$ is increasing in μ since P can always choose the same continuation strategy when the belief is $\mu' > \mu$ as she does when the belief is μ . This allows P to obtain at least weakly higher welfare with μ' compared to μ . The convexity of $J(\mu)$ follows from considering the case in which P receives additional information to update her belief about A 's type. The belief is martingale, so the updated belief is a mean-preserving spread of the original belief μ . Since P can always ignore this new information, this mean-preserving spread is always (at least weakly) welfare-improving.

Given the linearity of the value function $J(\mu)$ with respect to ρ , Lemma 1 has an immediate implication for the optimal replacement strategy:

Lemma 2 *The optimal replacement strategy for the principal is to remove the agent whenever $\mu \leq \mu_H$ and to continue with the agent otherwise.*

When making the replacement decision, the principal is comparing the payoff from replacement, $\bar{J} = J(\mu_H)$, to the payoff from continuation, $J(\mu)$. The value function $J(\mu)$ is increasing in μ , so the former value is no less than the latter whenever $\mu \leq \mu_H$. Intuitively, the principal continues with the agent as long as the expected type of a replacement agent is lower than the expected type of the current agent.

Assumption 2 implies that a principal maximizing her instantaneous utility would choose the following strategy:

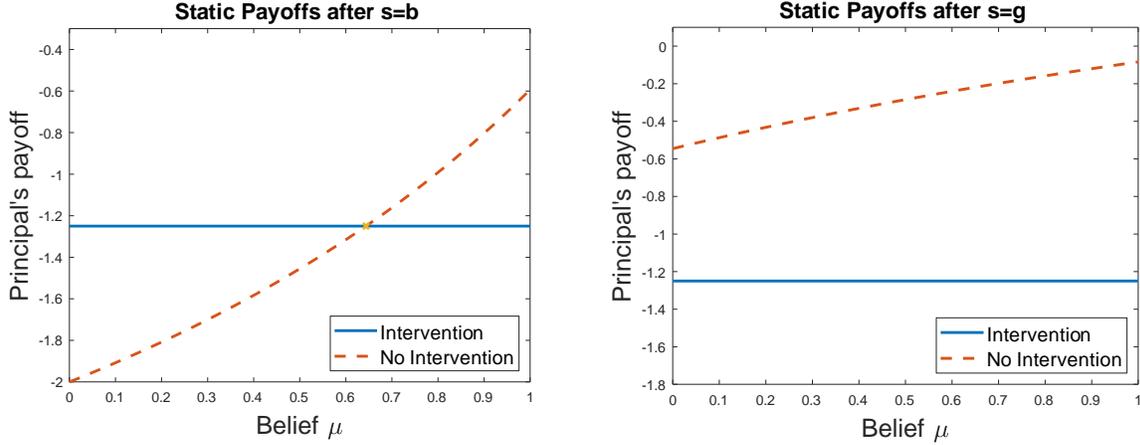


Figure 1: Pricipal's Payoffs when $\delta=0$.

Lemma 3 *Consider the principal's problem with $\delta = 0$ (a static problem). The optimal intervention strategy for the principal is to choose no intervention after $s = g$ and to choose intervention after $s = b$ if and only if $\mu \leq \mu^S$, where the threshold $\mu^S < 1$.*

If the signal is bad ($s = b$), then the project is expected to succeed only if there is a sufficiently high probability that the agent is an H -type. Otherwise, intervention comes at a lower cost than the expected loss from the project's outcome. Figure 1 illustrates the static problem faced by the principal after signal $s = b$.

Having established the optimal response in the static case, we move on to show that the optimal policy is still be a threshold strategy when $\delta > 0$:

Proposition 1 *There exists threshold $\bar{\alpha}$, such that if $\Pr(o = B|s = b, e = 1) \leq \bar{\alpha}$, then for each $\delta \in (0, 1)$, there exists $\mu^D < \mu^S$ such that the optimal intervention strategy for the principal is to choose intervention after $s = b$ if and only if $\mu \leq \mu^D$.*

The optimal dynamic policy differs from the static result of Lemma 3 because the principal also takes into account the effect that the current intervention decision has on her belief about the agent's type in the next period. Without intervention, the belief update is based on both the signal and the project outcome, while with intervention, the principal must

update her belief just based on the signal. Thus, not intervening in the current period has a dynamic benefit of better learning of the agent's type. This dynamic consideration changes the trade-off between intervention and no intervention. If the dynamic benefit is sufficiently high, not intervening is optimal even though it does not maximize instantaneous utility.

When $s = g$, no intervention is the best policy in the static context. Since not intervening also produces the benefit of learning, there is no reason for this policy to change in the dynamic game. The intuition explains why no intervention is also dynamically optimal after $s = b$ and $\mu \geq \mu^S$. When $\mu < \mu^S$, however, the static prescription is intervention. This precludes the project's outcome from being revealed. If the principal chooses no intervention instead, and the project's outcome is highly informative about the agent's type, then she obtains a large benefit from learning. In particular, the upper bound on $\Pr(o = B | s = b, e = 1)$ provides a sufficient condition for the outcome after signal b to be highly informative. The condition implies that a B -outcome after a b -signal is unlikely if the agent exerts effort $e = 1$. Hence, observing the event of (b, B) increases the principal's belief that the agent is of type L . Under the threshold $\bar{\alpha}$, the principal's belief about the agent being an H -type following (b, B) drops below μ_H , leading to immediate termination of the contract. Thus, under this sufficient condition, no intervention after $s = b$ has a strong learning benefit: it leads to P either learning that the agent is type H or to removing the agent. Moreover, this set of possible outcomes is the same for all $\mu < \mu^S$. Then, with a decreasing within-period benefit of intervention and a constant benefit of learning as μ increases, there exists a threshold $\mu^D < \mu^S$ above which not intervening is optimal. Figure 2 provides an illustration of the principal's payoff and the optimal dynamic intervention policy as functions of μ .

These results show that the optimal intervention policy under adverse selection has a stark characterization: intervention after bad news is optimal up until the principal is sufficiently confident that the agent is an H -type.

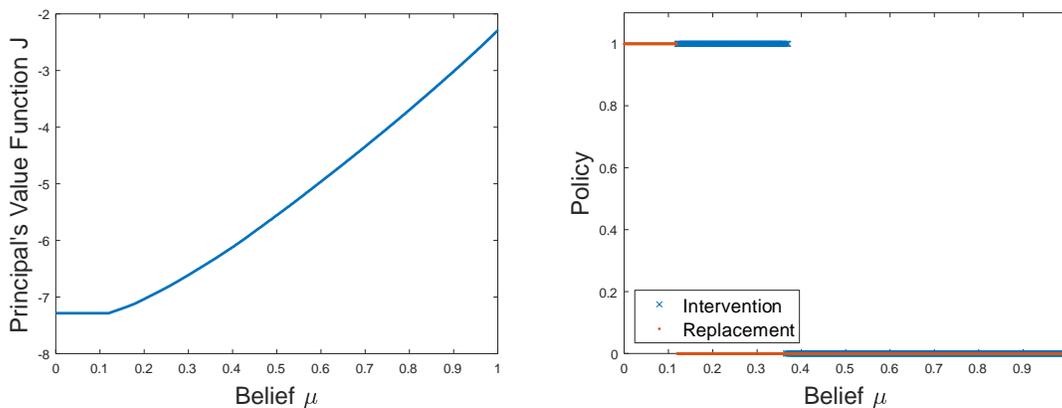


Figure 2: Principal's Payoff and Equilibrium Policies for $\mu=0.12$

4 Intervention in the Full Model

In this section, we analyze the optimal intervention policy for the principal in the full model. Costly effort captures a common problem in the type of principal-agent relationships that motivate this model, be they coalitions in the political realm, or transfers in the supranational context. The costly effort choice made by the agent can be microfounded as an electoral or ideological cost of moving policy away from the status quo. The electoral cost might emerge, for instance, from the mechanism described in [Fernandez and Rodrik \(1991\)](#), in which voters oppose ex-ante welfare-improving reforms due to their uncertain distribution of benefits. This cost is borne by the politician who works on implementing that project. We also assume that effort is unobservable to the principal. This assumption is consistent with our applications, in which the principal is not involved in the operation of the project and only performs an interim evaluation of the project. For instance, a government might assign a team to work on a reform project agreed upon with an international lender, but the de facto effort exerted to implement the reform cannot be directly observed or verified by the international lender. This is especially relevant if the implementation of the project requires specialized knowledge of local conditions, which makes it difficult for an outsider to evaluate the effort involved in an observed implementation. Similarly, our assumption of unverifiable effort applies when the principal is an investor and the agent is an entrepreneur.

4.1 Equilibrium Concept

The principal and the agent are both strategic players. We would like to derive the best Perfect Bayesian Equilibrium for the principal. To construct an equilibrium in this class, we allow for the option of observable mixed strategies. Since players move sequentially, we consider a public randomization device for P , using the technique also employed in [Ales et al. \(2014\)](#). This randomization device is used to determine the (pure strategy) play within the period. At the beginning of each period t , the principal draws a random variable $z_t \sim \text{Uniform}[0, 1]$, which is observed by everyone. Given z_t , the principal either replaces the agent or not, $\rho_t \in \{0, 1\}$, and determines the intervention rule $\iota_t(s_t)$ for each s_t . If not replaced, the agent chooses effort $e_t \in [0, 1]$. After the signal s_t , the principal decides whether to intervene or not, according to the rule $\iota_t(s_t)$.

The public events in period t are z_t , ρ_t , ι_t , and o_t , so the history of public events is $h_t = (z_j, \rho_j, \iota_j, o_j)_{j=1}^{t-1}$. The principal's strategy, σ_P , consists of a mapping from (h_t, z_t) to (ρ_t, ι_t) .²⁰ The agent's strategy, σ_A , is a mapping from $(h_t, z_t, \rho_t, \iota_t, \theta_t)$ to e_t , where θ_t is the private type of the agent in period t . With a continuous action space $e_t \in [0, 1]$ and a strictly convex cost function $c(e_t)$, the agent always chooses a pure strategy. Given the pure strategy, we can assume without loss of generality that the strategy e_t does not depend on the past effort $(e_j)_{j=1}^{t-1}$.

Definition 1 *Given the history h_t of public outcomes up to period t , a Perfect Bayesian Equilibrium consists of continuation strategies $\{\rho_t(h_t, z_t), \iota_t(h_t, z_t)\}_t^\infty$ for the principal, and $\{e_t(h_t, z_t, \rho_t, \iota_t)\}_t^\infty$ for the H -type agent, such that each player maximizes their respective expected continuation utilities and the belief μ_t follows Bayes' Rule whenever possible.*

We derive the equilibrium in two steps. First, we solve a relaxed problem, ignoring P 's incentive to follow the equilibrium. Without P 's incentive compatibility constraint, we can characterize the best on-path outcome for P by the recursive method. In particular,

²⁰As the principal and the agent take actions sequentially and the principal's actions are observable to the agent, it is without loss of generality to focus on pure strategies. In fact, the public randomization device allows us to generate mixed strategies, by associating different pure strategies to different realizations of z_t .

we study the problem of maximizing P 's welfare given her belief μ about the agent and a promised payoff V to the H -type agent. This structure allows P to write a contract for A that promises a certain level of total benefits to A . Second, we show that there exists a punishment equilibrium such that P has the incentive to fulfill this promise given the threat of punishment.

4.2 The Principal's Problem

Let $J(\mu, V)$ be P 's value function given the two state variables. If a replacement occurs, the belief goes back to μ_H , and P chooses what payoff to promise to the new agent in order to maximize her welfare. Hence, P 's utility after replacing A satisfies $\bar{J} = \max_V J(\mu_H, V)$.

The problem for P is to select a vector $\alpha_z = (\rho_z, e_z, \iota_z, (V'_z(\omega))_\omega)$ for each possible realization of z . The effort e_z is the proposed effort for the H -type agent, and $V'_z(\omega)$ is the promised continuation value for the agent of type H given the publicly observable actions and signal-outcome pair $\omega = (s, o)$. Since the on-path outcome has full support,²¹ the principal always believes that the recommended effort e_z is taken. Hence, the next period's belief $\mu'(\mu, e_z, \omega)$ is determined by Bayes' rule, which depends on the prior μ , the recommended effort e_z , and the signal-outcome pair ω .

By Assumption 2, no intervention is statically optimal after $s = g$. Moreover, no intervention allows the principal to have access to more information in the Blackwell sense, and this implies that no intervention is also dynamically optimal after $s = g$. We therefore focus on ι_z after $s = b$, since $\iota_z(s = g) = 0$. The principal's expected instantaneous utility is $u^P(\iota_z|\mu, e_z) := \sum_s \Pr(s|\mu, e_z)u^P(\iota_z(s)|\mu, e_z, s)$ (see (2)) with $\iota_z(s = g) = 0$. Given $u^P(\iota_z|\mu, e_z)$, the principal P chooses $\alpha_z = (\rho_z, e_z, \iota_z, (V'_z(\omega))_\omega)$ to solve the following dynamic

²¹Since P intervenes after $s = b$ with probability one given $\mu = 0$ or $e = 0$, we can pin down the continuation play after these histories arbitrarily.

program. For each $\mu \in [0, 1]$ and $V \in [0, \frac{1}{1-\delta}]$,

$$J(\mu, V) = \max_{\alpha} \int_z \left[(1 - \rho_z) \bar{J} + \rho_z \left\{ u^P(\iota_z | \mu, e_z) + \delta \sum_{\omega} \Pr(\omega | \mu, e_z, \iota_z) J(\mu'(\mu, e_z, \omega), V'_z(\omega)) \right\} \right] dz, \quad (7)$$

subject to the following constraints:

$$V = \int_z \rho_z \left\{ 1 - c(e_z) + \delta \sum_{\omega} \Pr(\omega | e_z, \iota_z) V'_z(\omega) \right\} dz; \quad (8)$$

$$e_z \in \arg \max \left\{ 1 - c(e_z) + \delta \sum_{\omega} \Pr(\omega | e_z, \iota_z) V'_z(\omega) \right\}; \quad (9)$$

$$V'_z(\omega) \in \left[0, \frac{1}{1-\delta} \right] \text{ for each } \omega. \quad (10)$$

Constraint (8) is the promise keeping that P is bound to in equilibrium, and (9) is the incentive compatibility constraint for A .²² Constraint (10) places the upper and lower bounds on the future promised value since 0 is the minimum payoff that the agent receives; and $\frac{1}{1-\delta}$ is the highest feasible utility for the agent, implemented by keeping the agent and allowing him to exert effort $e = 0$ forever.

The next lemma shows that the above maximization problem characterizes the best PBE for P .

Lemma 4 *The best Perfect Bayesian Equilibrium for the Principal is characterized by the mapping from $(\mu, V) \in [0, 1] \times [0, \frac{1}{1-\delta}]$ to a vector $(\rho_z, e_z, \iota_z, (V'_z(\omega))_{\omega})_{z \in [0, 1]}$ which maximizes (7) subject to (8), (9), and (10).*

By definition, objective (7) gives the highest welfare for P given the promise keeping and incentive compatibility constraints. Hence, we are left to construct a punishment equilibrium

²²Notice that, from P 's view point, the distribution of the signal-outcome pair $\Pr(\omega_z | \mu, e_z)$ depends on μ , while for the promise keeping and incentive compatibility, we have $\Pr(\omega | e_z)$ since A knows his own type.

to sustain this outcome on the equilibrium path. Note that there is always an equilibrium in which A exerts no effort, expecting replacement every period; given this, P picks intervention after every bad signal and replaces the agent every period, expecting no effort from the agent. In equilibrium, the lowest value P could obtain is $J\left(\mu, \frac{1}{1-\delta}\right)$, since the promised value $\frac{1}{1-\delta}$ allows A to stay without putting in effort. This value corresponds to the “no effort equilibrium.” Hence, by switching to this worst equilibrium after any deviation from policy α_z , we can sustain the solution for (7) on the equilibrium path.

Remark 1 *Note that in the above Lemma, we allow P to commit to the intervention rule ι_t at the beginning of each period t . This means that A chooses effort at the beginning of the period, given the intervention rule dictated by the equilibrium play. The principal then takes the promised intervention decision and cannot change it once effort is sunk. If instead we wanted to allow P to revise her intervention decision ι_t after observing s_t , all we need is an additional condition on parameters. In particular, notice that if P could revise her intervention decision ι_t after observing s_t , her only potential profitable deviation would be from $\iota_t(s = b) = 0$ to $\iota_t(s = b) = 1$. This is the case because the prospect of no intervention incentivizes A to provide higher effort, but once effort is sunk, intervention maximizes the instantaneous utility. Such a derivation is never profitable if the loss in continuation payoff outweighs the static gain, i.e., $\forall z$ and (μ, V_z) with $\iota_z(s = b) = 0$, the following condition is satisfied:*

$$u^P(1|\mu, e_z) - u^P(0|\mu, e_z) \leq \delta J(\mu'(\mu, e_z, b), V'_z(b)) - \delta \frac{u^P(1|\mu, 0)}{1-\delta}, \quad (11)$$

where $J(\mu'(\mu, e_z, b), V'_z(b))$ is the on-path continuation payoff and $\frac{u^P(1|\mu, 0)}{1-\delta}$ is the payoff from the no effort equilibrium. There are two reasons why P chooses not to intervene despite the static cost: more learning and more efficient use of the continuation payoff to incentivize effort. Both of them suggest that P 's continuation payoff is sufficiently high. This implies that $J(\mu'(\mu, e_z, b), V'_z(b))$ is usually sufficiently high relative to the punishment continuation payoff. In fact, this can be easily ensured given the choice of parameters, and condition (11) holds in all of the numerical examples we present in the rest of the paper.

4.3 Equilibrium Properties

We first show that P 's value function $J(\mu, V)$ is concave in V , convex and increasing in μ :

Lemma 5 $J(\mu, V)$ is concave in V , convex in μ , and increasing in μ , with this increase strict if $V \in (0, \frac{1}{1-\delta})$.

Since we allow the public randomization, the concavity of $J(\mu, V)$ with respect to V follows from the standard arguments. The properties that $J(\mu, V)$ is increasing and convex with respect to μ follow from the same exercise as in Lemma 1.

We next derive several implications about the shape of $J(\mu, V)$.

Lemma 6 *The value function $J(\mu, V)$ has the following properties:*

1. $J(\mu, 0) = \bar{J} \forall \mu$;
2. For each $\mu \in [0, 1]$, there exists $V(\mu) \in [0, \frac{1}{1-\delta}]$ such that $J(\mu, V)$ is linear for $V \in [0, V(\mu)]$, where $V(\mu) \geq 1$, with strict inequality for $\mu \geq \mu_H$. Moreover, the slope for the linear part, $\frac{d}{dV}J(\mu, V)|_{V \in [0, V(\mu)]}$, is negative for $\mu < \mu_H$, it is zero for $\mu = \mu_H$, and it is positive for $\mu > \mu_H$.
3. There exists $V^*(\mu) \in \arg \max_{V \in [0, \frac{1}{1-\delta}]} J(\mu, V)$ such that $\forall \mu \in (0, 1]$, at $V = V^*(\mu)$, $V'_z(\omega) \leq (\geq) \arg \max_V J(\mu'(\mu, e_z, \omega), V)$ for each event ω with negative (positive) belief update $\mu'(\mu, e_z, \omega) \leq (\geq) \mu$.

The above properties have the following intuition. First, if the value promised to A is 0, only immediate replacement can fulfill this promise. Hence, it must be the case that $J(\mu, 0) = \bar{J}$ for each μ . Second, for a sufficiently small promised value V , replacement must happen with positive probability — otherwise A would obtain a higher payoff than promised. This positive probability is generated by randomizing between keeping and replacing A . We can show that it is optimal for P to promise the same utility $V(\mu)$ if A is kept. Hence, any $V \in [0, V(\mu)]$ can be generated by varying the probability of keeping versus replacing A . These properties and those described in Lemma 5 are illustrated in Figure 3.

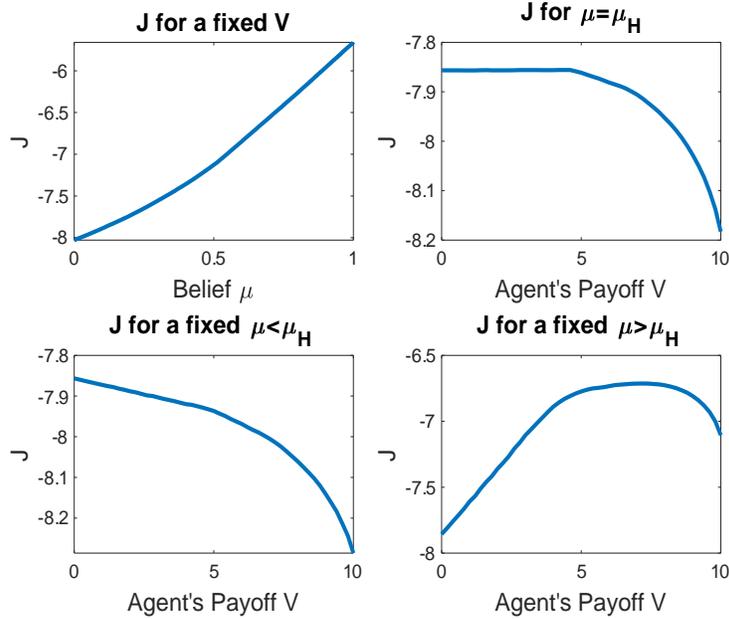


Figure 3: Shape of the Principal's Value Function. For the upper left panel, $V = \frac{0.25}{1-\delta}$. For the upper right panel, $\mu_H = 0.12$, for the lower left panel, $\mu = 0.07$, and the lower right panel, $\mu = 0.6$. We assume quadratic effort cost.

The third property follows from concavity: For each $\mu \in [0, 1]$, the function $J(\mu, V)$ is maximized at some promised value, denoted $V^*(\mu)$. Consider the case in which at $V = V^*(\mu)$ the belief is updated to μ' after the signal and outcome pair ω . A negative (positive) belief update implies that ω is the outcome which happens less (more) often with higher effort. Hence, reducing (increasing) $V'_z(\omega)$ incentivizes A to supply more effort. This increased effort has two benefits: the instantaneous welfare u^P is improved, and the belief update is larger, since higher effort makes type H more distinguishable from type L . If the continuation payoff $V'_z(\omega)$ is higher than the value that maximizes P 's utility ($\arg \max_V J(\mu', V)$), then reducing $V'_z(\omega)$ directly improves P 's continuation welfare as well. Otherwise, P 's continuation welfare decreases, because, intuitively, promising too little continuation payoff to A forces P to replace A even if P believes that the agent is an H -type.

4.4 The Optimal Dynamic Policy

An immediate implication of Lemma 6 is that an agent who starts off by delivering bad news or a bad outcome will be removed.

Corollary 1 *The agent is removed after the first period of his appointment if $\omega = (b, I)$ or $o = B$.*

In deciding the optimal intervention policy, the principal must balance the benefit of learning, the benefit of avoiding a bad outcome, and the cost of incentivizing effort. To shed more light on how this balancing act is resolved, consider first the principal's problem for a fixed effort level. Then, the same reasoning as in the benchmark case implies that not intervening allows the principal to update her belief about the agent more precisely than after an intervention:

Lemma 7 (Learning) *Fixing any targeted effort level e_z , when the principal does not intervene after $s = b$, learning about A 's type happens at a faster pace: the distribution of the updated beliefs $(\mu'(\mu, e_z, \omega))_\omega$ given no intervention is a mean-preserving spread of the distribution of updated beliefs given intervention. Moreover, the increase in the variance of $(\mu'(\mu, e_z, \omega))_\omega$ when P changes the policy from intervention to no intervention is a concave function of μ .*

By Lemma 5, J is convex in μ . Hence, the increased variance of the distribution of $(\mu'(\mu, e_z, \omega))_\omega$ under no intervention implies that no intervention benefits the principal by increasing the continuation payoff through learning. Moreover, this learning effect is stronger if the current belief is intermediate — there is a large room for belief updates.

Amplification of the Learning Benefit.

An implication of the above results is that the benefit of learning is amplified when effort must be incentivized. As discussed above, when she does not intervene, the principal further

increases her expected payoff by better calibrating the agent's continuation payoff. In particular, the outcome y is informative about the agent's effort, and she can use this information to implement higher effort at a lower cost. This increase in effort magnifies the benefit of learning: now the effort exerted by the H -type differs even more from that exerted by the L -type. Thus, the principal learns the agent's type more accurately.

Switches between intervention and no intervention.

Next, we show how balancing the dynamic learning and incentives provision benefits against the static cost of no intervention can lead to switches between intervention and no intervention on the equilibrium path. We formally show this result in the case in which the principal observes a sequence of events which update her belief positively along the path. Afterwards, we discuss other possible sequences of events.

We begin by giving the sufficient conditions for this result. We assume that the initial belief μ_H is sufficiently low, so that intervention is optimal at this starting belief. We derive the conditions on effort under which sufficiently high effort can be incentivized in equilibrium.

Definition 2 *Given $\bar{q} \in (0, 1)$, the problem satisfies the **effort provision condition** if there exists $\bar{e} \in (0, 1)$ such that:*

- (i) $\frac{d}{de} [\Pr(g, G|\bar{e}) \cdot \bar{q} - c(\bar{e})]|_{e=\bar{e}} \geq 0$;
- (ii) $\Pr(g, G|\bar{e}) \geq 1 - \bar{q}$;
- (iii) $\Pr(b, B|\bar{e}) / \Pr(b|\bar{e}) < l/C$.

Part (i) ensures that the H -type agent can be incentivized to exert at least effort \bar{e} . Part (ii) guarantees that, if this effort is exerted, the principal updates her belief μ with sufficiently high probability after observing (g, G) . At the limit where μ converges to 1, Part (iii) guarantees that no intervention is statically optimal. These conditions have the same implication as the condition of Proposition 1: if the principal could observe the agent's type, she would prefer not intervene given an H -type. This type exerts high effort, which makes a bad outcome sufficiently unlikely.

Given this definition, we establish that the simple threshold strategy implied by Proposition 1 no longer holds in the full model, and in fact we obtain switches between intervention and no intervention on the equilibrium path:

Proposition 2 *Consider a path of repeated realizations of the signal-outcome pair $\omega = (g, G)$. There exist upper bounds $\bar{\mu}_H \in (0, 1)$ for the initial belief μ_H , and $\bar{q}(\mu_H) \in (0, 1)$ for the probability of a G outcome given no effort $\Pr(G|0)$, such that if $\mu_H \leq \bar{\mu}_H$, $\Pr(G|0) \leq \bar{q}(\mu_H)$, and the problem satisfies the effort provision condition given $\bar{q}(\mu_H)$, then the optimal intervention policy $\{\iota_t(s = b)\}_t$ exhibits switches between intervention and no intervention. Intervention is optimal in the first period and in the long-run, and no intervention is optimal in some period $t \geq 2$.*

The result in Proposition 2 reflects the dual problem of selection and control faced by the principal. We focus on the simple path of repeated (g, G) realizations that generate a positive belief update each period. With the initial belief μ_H sufficiently small, the optimal policy is to intervene in the initial period. The principal is sufficiently pessimistic about the agent's type, and she intervenes to avoid a potential bad outcome. As the belief is updated positively, the principal becomes less pessimistic about the agent's type. The effort provision condition ensures that the agent can be offered incentives to exert high effort. It also ensures that the cost of intervention is not too high, so that the principal can choose not to intervene and learn more about the agent's type. Hence, the principal switches to no intervention. However, to incentivize high effort, the principal needs to reward the agent after a good outcome. As (g, G) realizations accumulate and the belief is updated higher, the promised reward increases. Eventually, the promised reward becomes so large that incentivizing more effort becomes impossible — given the promised rewards, the principal has to keep the agent without implementing a positive level of effort. Once we reach this phase, intervention becomes optimal again, this time because of the control issue. The principal can no longer control the agent through the promised reward, and therefore, he must intervene because the agent supplies too little effort. Figures 4 illustrates this dynamic in a numerical example of the model.

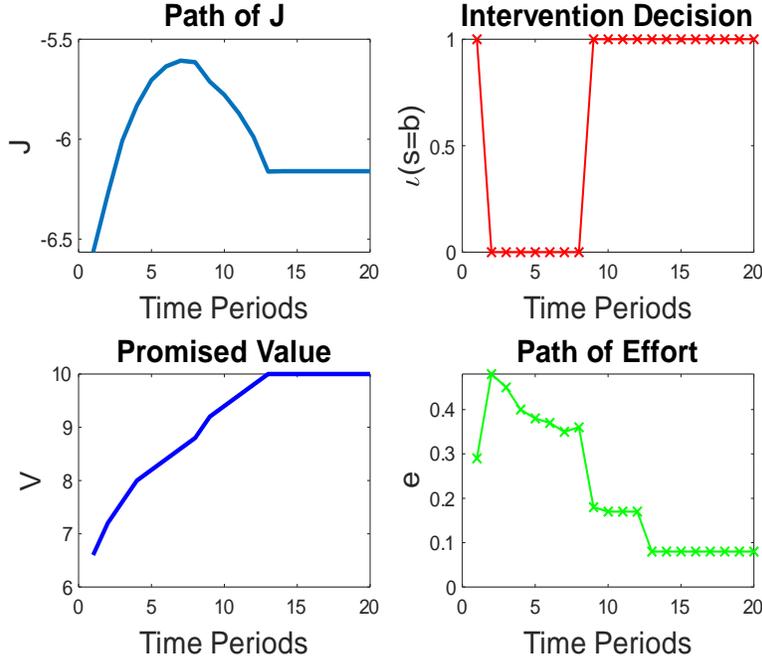


Figure 4: Dynamics after the signal-outcome realization (g,G) over 20 periods.

The above results have two main implications for the implementation of the optimal intervention policy. First, intervention after a bad signal may occur throughout the duration of the relationship with the agent, even if the agent has established a track record of good outcomes. Second, once the agent has built a large enough track record of good outcomes, the principal should intervene after bad news. The pattern suggested by this result has not been empirically examined in settings involving political actors. Yet, examples of this implementation may be found in the private sector. For instance, [Fich and Shivdasani \(2007\)](#) examine what happens to the value of firms which have a director sitting on their board who is also on the board of a firm accused of financial fraud. The study finds that, among the non-accused firms, the investor reaction is more negative for the firms in which the director's tenure on the board is longer. Thinking of the firm's leadership as the agent in our setting and the investors as the principal, the findings suggest the dynamic described by our model: the response to bad news that is correlated to the quality of the firm's leadership is more negative for agents with longer tenures. Our results provide predictions for the examination

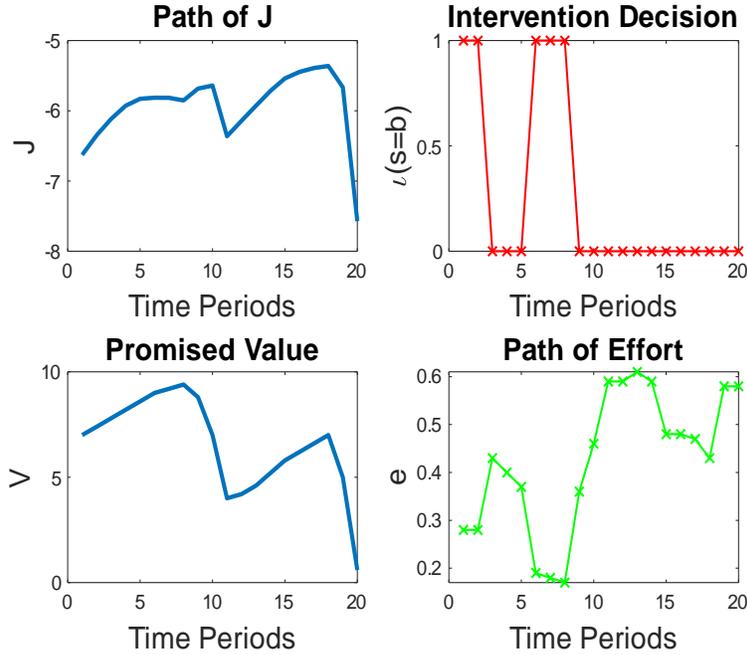


Figure 5: Dynamics after a sequence of alternating 8 periods with the signal-outcome realization (g,G) and 2 periods with the realization (g,B)

on the dynamics of intervention in broader settings, including the audit and intervention policies of international institutions and the functioning of governing coalitions.

Finally, notice that the switches between intervention and no intervention are a function of both the principal's belief μ and the agent's promised value V . Thus, the result in Proposition 2 can be extended to obtain switches between intervention and no intervention under a variety of signal-outcome observations. For instance, consider a sequence of $n > 0$ repeated signal-outcome observations (g, G) followed by $m > 0$ observations (g, B) that lower the principal's belief. The principal chooses to intervene in the periods in which the belief about the agent is sufficiently low or the agent's promised value is sufficiently high. Figure 5 illustrates this dynamic for such an example of outcome realizations.

4.5 Comparative Statics

In this section, we derive several comparative statics to shed more light on the drivers of our results. We are interested in examining how changes in the cost of providing incentives to the agent and changes in the informativeness of the signal affect the dynamics of intervention. In order to arrive at analytical expressions for these comparative statics, we reduce our infinite-horizon model to a simplified three-period version, whereby we preserve the within period trade-offs faced by the principal and the agent, but under less complex continuation payoffs. To further facilitate the analysis, we also assume in this simplified version of the model that effort is binary, $e \in \{0, 1\}$ with costs $c(0) = 0$, $c(1) = \gamma$. To measure the effect of the state variable (μ_1, V_1) in period 1, the principal starts with the initial belief $\mu_1 \in [0, 1]$ and initial promised value $V_1 \in [1, 1 + \delta + \delta^2]$. The full description and analysis of the simplified three-period model is given in Online Appendix B. In what follows, we present its analytical comparative statics results. Except for the comparative statics with respect to V_1 in Proposition 5, we assume that the promise keeping constraint does not bind in the principal's problem, that is, the principal just started a contract with a new agent. We show that the results obtained in this simplified model are consistent with the comparative statics obtained in simulations of the full model.

We first consider variations in the cost of providing incentives to the agent.

Proposition 3 (Cost of Providing Incentives) *An increase in the common discount factor δ increases the marginal benefit of no intervention. An decrease in the marginal cost γ has an ambiguous effect on the marginal benefit of no intervention.*

Higher values of δ increase the value of the future payoffs derived by the principal, for any constant effort level incentivized from the agent. The higher value of future payoffs increases the value of learning about the agent's type, and so, the principal has more to gain from not intervening and observing a more informative signal. Since the agent also discounts the future at rate δ , his continuation payoff also increases, which means that effort can be incentivized with a smaller reward for positive events.

When γ decreases, the marginal cost of effort is lower, which implies that the conflict between the principal and the agent is reduced. As γ decreases, on the one hand, more efficient monitoring becomes less important since incentivizing effort is less costly. On the other hand, using continuation payoffs more efficiently becomes more valuable since the continuation payoff itself is now higher because of the lower cost of incentivizing effort in the future periods.

The above results highlight that, while both δ and γ affect the cost of providing incentives, their effect on the principal's use of intervention is distinct. The discount factor δ captures factors that affect the probability of future projects, and it acts only through the value of future payoffs. It may capture, for instance, in reduced form, the expectation that the principal will have access to funds in the future in order to continue financing projects. In the application in which the principal is a supranational institution, changes in δ may capture the expectation that the institution will survive in future periods. The marginal cost of effort, however, captures the magnitude of the control problem, which exists both in the present and in the future, leading to a trade-off between the current benefits of less monitoring and the future value of more learning. For instance, it may capture a reduction in uncertainty among voters about the benefits of a reform project. This distinction in the comparative statics results highlights once again why the control problem makes the resulting policy markedly different. The difference in the value of no intervention relative to intervention for different values of δ and γ in the full model is illustrated in Figure 6.

Next, we consider changes in the precision of the signal generated by the project.

Proposition 4 (Informativeness of the Signal) *An increase in the probability of a “false-positive” signal, $P(b, G|e = 1)$, without a change in the probability distribution of outcomes $y \in \{G, B\}$, increases the marginal benefit of no intervention.*

As expected, if the information becomes less valuable for the principal, intervention brings a relatively lower gain compared to not intervening. This emphasizes the double cost of fraught audit or oversight systems: not only do they impose more unnecessary cost in

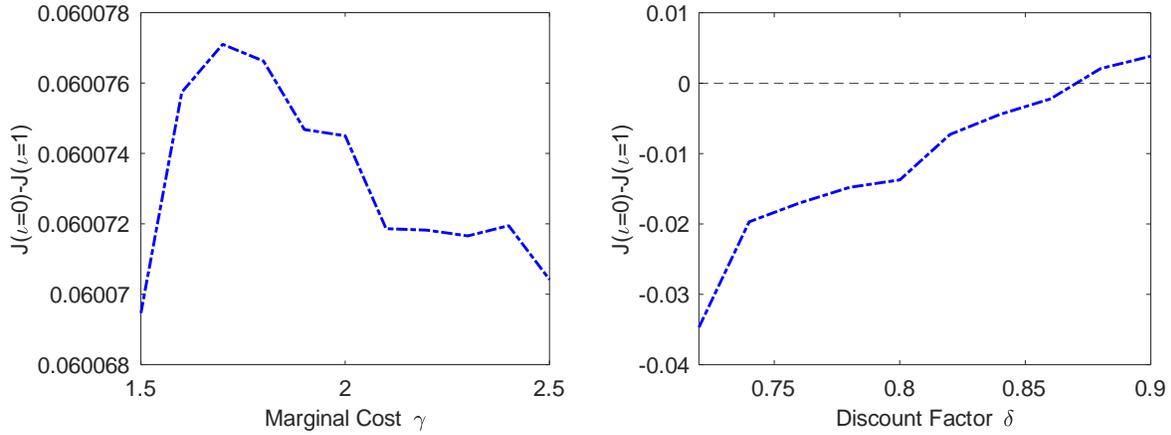


Figure 6: Relative Value of No Intervention as a function of γ (left panel) and as a function of δ (right panel). The graph is shown for $V = ((0.5)/(1 - \delta))$ and $\mu = 0.75$ and quadratic effort cost $\frac{\gamma}{2} \cdot e^2$.

the present, but they reduce future payoffs by slowing down the learning process about the agent’s ability and exacerbating the control problem.

Finally, we discuss two “partial equilibrium” effects: all else equal, we examine that the effect of changes in the value V promised to the agent at the beginning of the period and the effect of changes in the principal’s belief μ that the agent is an H -type.

Proposition 5 *The intervention decision is not monotone in the agent’s promised payoff V . Similarly, the marginal benefit of intervention may increase or decrease in the agent’s reputation μ .*

The non-monotonicity of the intervention decision on the equilibrium path was shown formally in Proposition 2. A similar intuition can be applied when examining the separate roles of V , which captures the reward needed to address the control problem, and μ , which captures the extent of the selection problem. For examining the effect of changes in V , consider a high value of the current’s agent’s reputation μ and a low expected ability of the replacement agent, μ_H . With a small V , the principal has to replace the current agent. After replacement, it is optimal to intervene given a low μ_H . With an intermediate V , it is possible to implement high effort, which makes no intervention optimal given a high μ ; however, if

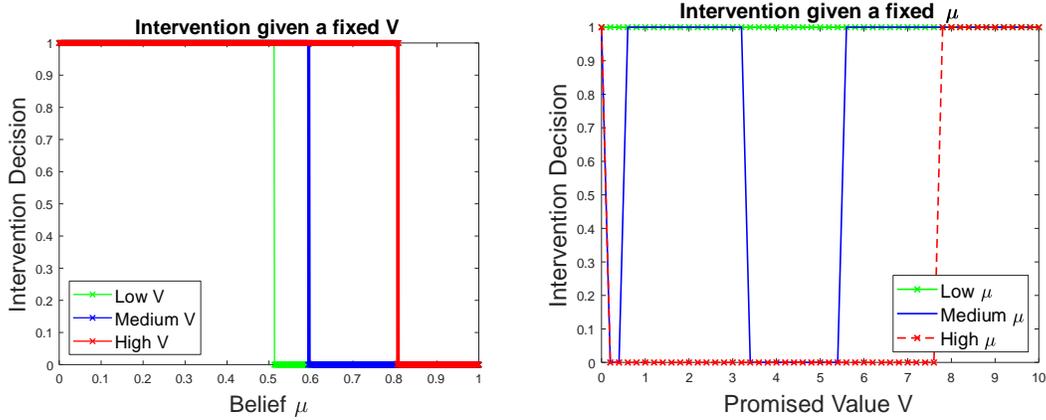


Figure 7: The right panel shows the intervention decision given $V = ((0.2)/(1 - \delta)), ((0.3)/(1 - \delta)), ((0.4)/(1 - \delta))$. The right panel shows the intervention decision given $\mu = 0.3, 0.5, 0.8$

V is very large, then the principal cannot implement high effort and is bound to keep the agent even if he does not supply effort. Hence, intervention again becomes optimal.

The effect of increasing μ on the marginal benefit of not intervening may be either positive or negative, depending on the parameters. On the one hand, no intervention becomes more beneficial since we assume that, with $\mu = 1$, no intervention is optimal. On the other hand, with intervention, the H and L types are less distinguishable, and a principal who wants to incentivize effort must therefore also give more rent to the L -type. As μ increases, the cost to the principal of offering higher rents to the L -type decreases, as the L -type exists with lower probability. If the latter effect dominates, the principal's gain under intervention increases more than her gain under no intervention.

Figure 7 illustrates the relative value of intervention in the full model as a function of μ , for a constant V (in the left panel) and optimal intervention response of as a function of V , for a constant μ (in the right panel).

5 Robustness of the Result

Our result from Proposition 2 of switches between intervention and no intervention on the equilibrium path relies on a specific feature of the PBE. Specifically, the promised reward eventually becomes too high to incentivize effort. Although the principal has incentives to keep such a promise by the threat of the “no effort equilibrium,” in some applications, we may also want to consider an equilibrium that is “renegotiation-proof.” In this section, we show that switches between intervention and no intervention also emerge in a Markov Perfect Equilibrium (MPE). In Online Appendix C, we provide the full definition and analysis of the MPE.²³ Below we summarize the main result, the counterpart to Proposition 2.

Proposition 6 *Consider a path of repeated realizations of the signal-outcome pair $\omega = (g, G)$. There exists a set of parameter values under which the optimal intervention policy $\{\iota_t(s = b)\}_t$ exhibits switches between intervention and no intervention in the optimal regular quasi Markov Perfect Equilibrium. Intervention is optimal if μ is sufficiently low or sufficiently high, and no intervention is optimal for intermediate values of μ .*

In the MPE, the principal’s payoff at the beginning of a period is a function of her belief μ , denoted by $J(\mu)$. We first note that $J(\mu)$ is constant over $\mu \in [\mu_H, 1]$, taking some value $J(\mu) = \bar{J}$, which implies that both the continuation payoff and the per-period payoff must be constant. The intuition is that, if this were not the case, the principal could not replace the agent after reaching a belief μ with $J(\mu) > J(\mu_H)$. Then, the agent would not supply effort, and this would result in a low payoff for the principal.

The intuition behind Proposition 6 can be obtained from the following exercise. Suppose that there are three feasible levels of efforts: $e^H = 1$, $e^M = 1/2$, and $e^L = 0$. We conjecture that e^H is implementable for μ_H , and, at belief μ_H and effort e^H , intervention after $s = b$ maximizes the principal’s instantaneous utility, $(1 - \delta) \cdot \bar{J}$. Since the principal’s payoff $J(\mu)$ is constant, it must be that the expected effort level decreases as μ increases. This

²³Technically speaking, we allow the principal to take period- t actions based on the belief μ and the history within period t . Otherwise, the principal cannot incentivize any effort. In particular, we define a regular quasi Markov Perfect Equilibrium.

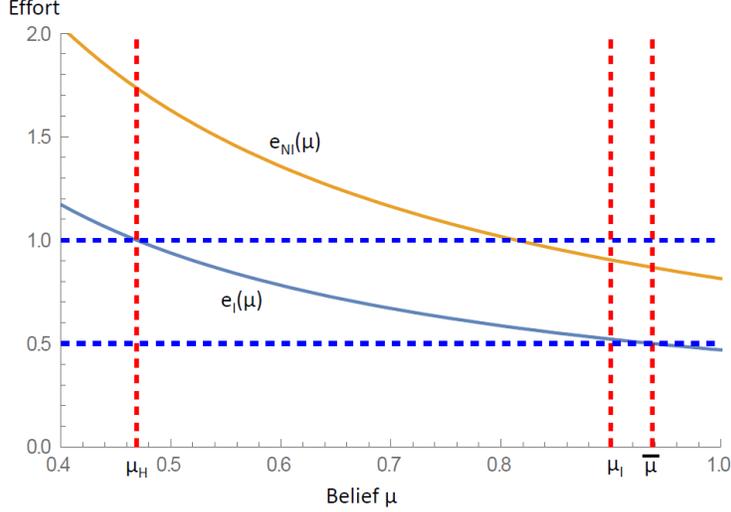


Figure 8: Effort as a function of the belief μ in the MPE

is illustrated in Figure 8. The curve labeled $e_I(\mu)$ represents the level of effort needed to achieve $(1 - \delta) \cdot \bar{J}$ with intervention and the curve labeled $e_{NI}(\mu)$ represented the level of effort needed to achieve the per-period payoff without intervention. The cutoff μ_I represents the threshold belief with which, given $e = e^H$, the principal is indifferent between intervening and not intervening to maximize her instantaneous utility; the cutoff $\bar{\mu}$ represents the belief μ for which $e_I(\mu) = e^M$.

For $\mu \in [\mu_H, \mu_I)$, an effort level higher than e^M must be exerted in expectation in order to generate the required per-period payoff. This means that the principal must incentivize e^H and intervene with a positive probability.²⁴ For $\mu \in [\mu_I, \bar{\mu})$, since $e_I(\mu) > e^M$, the principal again has to incentivize e^H ; however, in this region the principal prefers no intervention after e^H . Finally, for $\mu \in [\bar{\mu}, 1]$, the principal has two options: implement e^M and intervention, or e^H and no intervention. There exists a set of parameters such that for $\mu \geq \bar{\mu}$, the conjectured value $(1 - \delta) \cdot \bar{J}$ is supportable in the MPE with e^M and intervention, but it is not supportable with e^H and no intervention. Intuitively, if the cost of exerting effort e^H is very high, then implementing e^H and no intervention provides a low payoff to the agent at a high μ . This in

²⁴She also incentivizes $e = e^M$ to exactly achieve the instantaneous utility of $(1 - \delta) \cdot \bar{J}$, depending on the realization of the public randomization. Instead, we could consider a smooth cost function $c(e)$ for $e \in [0, 1]$. In such a case, the principal exactly incentivizes $e_I(\mu)$ with probability one given the convex cost.

turn limits the principal’s ability to reward the agent sufficiently after he generates a good outcome at a low μ , as generating a positive belief update leads the agent to a low payoff in the next period.

6 Discussion of Applications

In this section, we focus on two applications that are at the center of policy and academic debates, and that have motivated our model: intervention policies in supranational agreements and monitoring in government coalitions. Our model is stylized and meant to capture the general dynamics of intervention applicable in a broad class of cases. This naturally limits its ability to fully capture the richness of the institutional details of each application. Nevertheless, we describe in detail below how to map our model to each of application, and we discuss the resulting positive implications.

6.1 Supranational Agreements

The issues of selection and intervention are central to supranational agreements in which transfers are made by a supranational institution for projects implemented by local policymakers. For instance, since 1969, a stated objective of the World Bank aid program has been that “increased allocation of aid should be primarily linked to performance,”²⁵ and since 1996, the World Bank program rules allow it to perform audits on contractors working on the its funded projects.²⁶ In our model, the principal is the supranational institution that funds projects. The agent is the local policymaker who implements a project — the minister or the head of the agency who is in charge of running the project locally. The project may be anything from a building rehabilitation to a broad program aimed at achieving a development objective. The relationship between the Bank and the local policymaker generally has a repeated nature, as projects are proposed over time. The projects may be audited —

²⁵Pearson Commission on International Development, 1969, as cited in [Winters \(2010\)](#).

²⁶See [Rose-Ackerman \(1997\)](#).

which maps to the signal in our model — and, if the audit results are negative, e.g., the signal is bad, then funding for the project may be stopped — an intervention may occur. Finally, the Bank may decide to stop funding future projects ran by that policymaker, by cutting aid to programs ran by that ministry or agency.

A similar mapping can be made from our model to the Structural and Investment funds provided by the European Commission (EC) for projects in EU member countries. In this case, the EC is the principal and the local policymaker — for instance, a minister, head of local agency or a local mayor — is the agent. The projects funded include, for instance, infrastructure projects, job training programs, research and innovation initiatives.²⁷ The progress made on projects is verified periodically — the signal in our model —, and the EC may intervene by revising projects or by suspending funding. For failures that go uncorrected, future funding may be suspended indefinitely, and the EC may use those funds for other projects in a different country or under the purview of a different policymaker.

The Cost of Delegation. In the context of supranational agreements like the ones described above, our model highlights the complexity added to these agreements whenever a control problem exists — stemming from a concern over corruption as described in [Rose-Ackerman \(1997\)](#), electoral costs to implementing reforms as in [Fernandez and Rodrik \(1991\)](#) or simply the need to motivate local policymakers to work on the priorities of the supranational institution. In particular, in the benchmark model, the agent’s type is isomorphic to underlying conditions for project success — a good or a bad environment for that project. Thus, this case is akin to the supranational institution directly running a project and learning whether the underlying environment is good for the project — for instance whether small business training programs are effective in that local market. The optimal intervention policy then takes the form of a simple threshold policy: the program is monitored while information is gathered about its effectiveness in that local market, and monitoring stops once there is enough evidence that it is working well. When the supranational institution cannot conduct

²⁷More details at https://ec.europa.eu/info/funding-tenders/funding-opportunities/funding-programmes/overview-funding-programmes/european-structural-and-investment-funds_en

the project and must instead delegate to a local policymaker, it must ensure that sufficient effort is exerted on the project, not only that the project is adaptable to local conditions. This requires ongoing monitoring and potential intervention, which increases the cost of the project and the complexity of the institutional structure.

The Use of Intervention. Our model shows that intervention is used when the belief about the policymaker is low, and, as reputation increases, intervention may or may not be used, depending on the path of observed outcomes. This can be mapped to the choice of the type of project a supranational institution decides to fund in a given period. [Winters \(2010\)](#) examines World Bank funding decisions and shows that when the local policymaker’s reputation is low, as measured by a good governance score, the World Bank is more likely to offer *project* aid — funding aimed at projects that are easily monitored and may be easily stopped in case of negative audit reports, e.g. local infrastructure projects; however, when the local policymaker’s reputation is high, the World Bank is more likely to offer *programmatic* aid — funds aimed at budgetary support for government programs —, which is more difficult to monitor and intervene in. Programmatic aid is thus more likely to be used when intervention is not going to be used.

Dynamics of Reforms. The model also has implications regarding the evolution of projects over time. In particular, our results show that supranational agreements might not become better over time at implementing the supranational institution’s priorities. Instead, over time more discretion must be given to local policymakers to pursue their preferred projects (or maintain the status quo). This dynamic provides another rationalization of the phenomenon described as “reform fatigue” ([Bowen et al., 2016](#)), the slowdown of reforms after several reforms have already been implemented. It provides an mechanism through which this fatigue may be observed specifically for reforms demanded by supranational agreements.

6.2 Governing Coalitions

In parliamentary democracies, governments are oftentimes formed by coalitions made up of two or more political parties. Policy originates in a ministry of the government, which maps that minister to the agent in our model. The parliamentary committee that corresponds to that ministerial jurisdiction maps to the principal in the model. The parliamentary committee's membership reflects the composition of the ruling coalition and seeks to implement the projects agreed upon inside the coalition. These committees may schedule hearings, gather information — the signal in our model — and they can engage in legislative review by proposing amendments to the legislation produced by the minister, delaying or blocking legislation — actions which map into intervention in our model. The relationship between ministers and parliamentary committees has a repeated nature, and a minister may be removed by losing the support of the ruling coalition.

Patterns of legislative review. [Martin and Vanberg \(2004, 2005\)](#) provide empirical evidence on the patterns of parliamentary intervention through legislative review. They show that intervention is observed more often when there are larger preference disagreements within the government coalition. Our model rationalizes this pattern when preference disagreements imply a higher discount of the future, i.e., a coalition in which there is more disagreement between members will exogenously collapse with higher probability. Our comparative statics results show that this increases the value of intervention, making intervention more likely on the equilibrium path.

Our model also generates additional predictions, given the role that moral hazard plays in generating preference disagreements. Preference disagreements within the coalition also imply the need for costly effort on the part of ministers. Then, in policy areas with large preference disagreements, our model predicts legislative review to be recurring over time. In policy areas with no preference disagreements within the government coalition, legislative review should be observed less often after a number of policies are passed successfully.

Dynamics of legislative proposals. Our model also predicts changes in policy proposals as the tenure of the coalition increases and it registers several legislative successes. Specifically, it predicts policy proposals moving away from the compromise that weighs the ideological preferences of individual coalition members. Each minister can no longer be motivated to implement projects that depart from his own preferences (as effort on the principal’s project decreases). Thus, policy moves towards the ideological preference of the minister. As this happens, coalition members make more use of legislative review (intervention), which leads to less effective policymaking.

7 Conclusion

In this paper, we presented a model that captures several key features of environments in which we encounter interventions. First, an intervention changes the course of action and leads to a new outcome than the one that would have been observed had the intervention not happened. Second, the post-intervention outcome is uninformative about the final outcome that would have happened without intervention. Third, intervention is usually one of the very few levers that can be pulled to influence a policymaker’s actions. This is usually the case in the political realm, where the policymaker cannot be paid a wage contingent on the final outcome of his chosen policy. We show that, if the principal faces a selection problem only, the optimal policy involves intervention only as long as the belief about the agent’s ability is sufficiently low. If, in addition, there is also a control problem, the optimal policy can exhibit switches between intervention and no intervention on the equilibrium path.

Our model considers the case in which the principal runs only one project each period, and the agent cannot choose among multiple projects that can benefit the principal. In this setting, we could provide a rigorous dynamic analysis of the intervention problem, on which more complexity can be later built. Our insights can be used to understand the necessary deviations from the single project optimal policy once multiple projects are introduced. In particular, our results highlight the rich structure of the optimal policy that emerges due to

the control problem in political environments, in which the agent cannot be offered a wage schedule. Having an environment with multiple projects could add an additional instrument for the principal. The principal could link decisions across projects in order to provide better incentives. This could offer additional insights into how to sustain effort over time. It would also be a natural extension for the applications discussed above.

References

- Acemoglu, D., Golosov, M., Tsyvinski, A., 2008. Political economy of mechanisms. *Econometrica* 76 (3), 619–641.
- Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. *Journal of Political Economy* 105 (1), 1–29.
- Ales, L., Maziero, P., Yared, P., 2014. A theory of political and economic cycles. *Journal of Economic Theory* 153, 224–251.
- Austen-Smith, D., Banks, J. S., 1989. Electoral accountability and incumbency. In: Ordeshook, P. C. (Ed.), *Models of Strategic Choice in Politics*. University of Michigan Press.
- Banks, J. S., Sundaram, R. K., 1993. Adverse selection and moral hazard in a repeated elections mode. In: William A. Barnett, Melvin J. Hinich, N. J. S. (Ed.), *Political Economy: Institutions, Competition and Representation*. Cambridge University Press, pp. 295–311.
- Banks, J. S., Sundaram, R. K., 1998. Optimal retention in agency problems. *Journal of Economic Theory* 82 (2), 293–323.
- Barro, R. J., 1973. The control of politicians: an economic model. *Public Choice* 14 (1), 19–42.
- Besley, T., Case, A., 1995. Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition. *American Economic Review* 85 (1), 25–45.

- Besley, T., Smart, M., 2007. Fiscal restraints and voter welfare. *Journal of Public Economics* 91 (3-4), 755–773.
- Bowen, T. R., Chan, J., Dube, O., Lambert, N., 2016. Reform fatigue. Mimeo, Stanford University.
- Buisseret, P., 2016. Together or Apart? On Joint versus Separate Electoral Accountability. *Journal of Politics* 78 (2), 542–556.
- Coate, S., Morris, S., 1995. On the form of transfers to special interests. *Journal of political Economy* 103 (6), 1210–1235.
- Dovis, A., Golosov, M., Shourideh, A., 2016. Political economy of sovereign debt: A theory of cycles of populism and austerity. NBER Working Paper.
- Ferejohn, J., 1986. Incumbent performance and electoral control. *Public Choice* 50 (1-3), 5–25.
- Fernandez, R., Rodrik, D., 1991. Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review* 81 (5), 1146–1155.
- Fich, E. M., Shivdasani, A., 2007. Financial fraud, director reputation, and shareholder wealth. *Journal of Financial Economics* 86 (2), 306–336.
- Fox, J., Van Weelden, R., 2010. Partisanship and the effectiveness of oversight. *Journal of Public Economics* 94 (9), 674–687.
- Fox, J., Van Weelden, R., 2012. Costly transparency. *Journal of Public Economics* 96 (1), 142–150.
- Gavazza, A., Lizzeri, A., 2007. The perils of transparency in bureaucracies. *American Economic Review* 97 (2), 300–305.
- Halac, M., Kartik, N., Liu, Q., 2016. Optimal contracts for experimentation. *The Review of Economic Studies* 83 (3), 1040–1091.

- Herrera, H., Ordoñez, G., Trebesch, C., 2014. Political booms, financial crises. NBER Working Paper 20346.
- Kreps, D. M., Wilson, R., 1982. Reputation and imperfect information. *Journal of Economic Theory* 27 (2), 253–279.
- Laver, M., Shepsle, K. A., 1996. Making and breaking governments: Cabinets and legislatures in parliamentary democracies. Cambridge University Press.
- Levitt, S. D., Snyder, C. M., 1997. Is no news bad news? Information transmission and the role of "early warning" in the principal-agent model. *The RAND Journal of Economics*, 641–661.
- Mailath, G. J., Samuelson, L., 2001. Who wants a good reputation? *Review of Economic Studies* 68 (2), 415–441.
- Mailath, G. J., Samuelson, L., 2006. Repeated games and reputations: long-run relationships. Oxford University Press.
- Martin, L. W., Vanberg, G., 2004. Policing the bargain: Coalition government and parliamentary scrutiny. *American Journal of Political Science* 48 (1), 13–27.
- Martin, L. W., Vanberg, G., 2005. Coalition policymaking and legislative review. *American Political Science Review* 99 (1), 93–106.
- Prat, A., 2005. The wrong kind of transparency. *American Economic Review* 95 (3), 862–877.
- Ray, D., 2002. The time structure of self-enforcing agreements. *Econometrica* 70 (2), 547–582.
- Rogoff, K., 1990. Equilibrium political budget cycles. *American Economic Review* 80 (1), 21.
- Rose-Ackerman, S., 1997. The role of the world bank in controlling corruption. *Law and Policy in International Business* 29, 93.

Stokey, N. L., 1989. Recursive methods in economic dynamics. Harvard University Press.

Strulovici, B., 2011. Renegotiation-proof contracts with moral hazard and persistent private information. SSRN Working Paper No. 1755081.

Winters, M. S., 2010. Choosing to target: What types of countries get different types of world bank projects. *World Politics* 62 (3), 422–458.

A Appendix

A.1 Proof of Lemma 1

Monotonicity with respect to μ .

Suppose that $J(\mu) = J$ for some μ . Then, for a higher value $\mu' > \mu$, we have $J(\mu') \geq J$. To see why, if (ρ, ι) is a feasible policy when the belief is μ , then it is also feasible when the belief is μ' . Recall that the instantaneous utility for P given ι is

$$\Pr(s = g|\mu) u^P(0|\mu, e, s) + \Pr(s = b|\mu) u^P(\iota(s = b)|\mu, e, s),$$

which is (weakly) increasing in μ . Hence, by the standard arguments — see [Stokey \(1989\)](#) — $J(\mu)$ is (weakly) increasing in μ .

Convexity with respect to μ .

Let $J(\mu, \theta)$ be the payoff when the principal follows the optimal strategy given μ , and the current type is $\theta \in \{H, L\}$. We have

$$J(\mu) = \mu J(\mu, H) + (1 - \mu) J(\mu, L) = J(\mu, L) + \mu [J(\mu, H) - J(\mu, L)].$$

Take μ, μ_1, μ_2 and $\beta \in [0, 1]$ such that $\mu = \beta\mu_1 + (1 - \beta)\mu_2$. For $n \in \{1, 2\}$, by taking the strategy given μ when the belief is μ_n , P obtains

$$J(\mu, L) + \mu_n [J(\mu, H) - J(\mu, L)] \leq J(\mu_n).$$

Hence,

$$\begin{aligned} & \beta J(\mu_1) + (1 - \beta) J(\mu_2) \\ \geq & \beta J(\mu, L) + \beta\mu_1 [J(\mu, H) - J(\mu, L)] \\ & + (1 - \beta) J(\mu, L) + (1 - \beta)\mu_2 [J(\mu, H) - J(\mu, L)] \\ = & J(\mu, L) + \mu [J(\mu, H) - J(\mu, L)] = J(\mu). \end{aligned}$$

A.2 Proof of Lemma 3

After $s = b$, the principal does not intervene if

$$\frac{l}{C} \leq \Pr(H|\mu, s = b) \cdot \Pr(B|H, b) + \Pr(L|\mu, s = b) \cdot \Pr(B|L, b),$$

where $\theta = H$ implies $e = 1$ and $\theta = L$ implies $e = 0$. This inequality reduces to

$$\mu \geq \mu_S := \frac{\Pr(b, B|L) \left(1 - \frac{l}{C}\right)}{(\Pr(b, B|L) - \Pr(b, B|H)) \left(1 - \frac{l}{C}\right) + \frac{l}{C} \Pr(b, G|H)}. \quad (12)$$

A.3 Proof of Proposition 1

After $s = b$, by Bayes' rule, the belief is updated to

$$\mu_b = \frac{\mu}{\mu + (1 - \mu) \frac{\Pr(b, B|L) + \Pr(b, G|L)}{\Pr(b, B|H) + \Pr(b, G|H)}}.$$

Writing $u^\theta(\iota)$ as the expected payoff from $\iota \in \{0, 1\}$ if $s = b$ and the agent is of type θ , the principal's problem given $s = b$ is

$$\begin{aligned} & \mu_b u^H(\iota) + (1 - \mu_b) u^L(\iota) \\ & + \delta \iota J(\mu_b) + \delta (1 - \iota) (\alpha \mu_b J(1) + (1 - \alpha \mu_b) J(\mu')), \end{aligned} \quad (13)$$

where μ' is the belief after $s = b$ and $y = B$:

$$\mu' = \frac{\mu \Pr(b, B|H)}{\mu \Pr(b, B|H) + (1 - \mu) \Pr(b, B|L)}.$$

Notice that when $\mu \geq \mu^S$, $\iota = 0$ is optimal by Lemma 3. When $\mu = \Delta$, where $\Delta \rightarrow 0$, we have $\mu' \rightarrow 0$. Thus, $\iota = 1$ is optimal for $\mu \rightarrow 0$. Therefore, we can establish that intervention is optimal for μ sufficiently small, and no intervention is optimal for μ sufficiently large.

Since the prior μ and the interim belief μ_b have a monotone relationship, it suffices to show that there exists $\mu_b^* \in (0, 1)$ such that the intervention after $s = b$ is optimal if and only if $\mu_b \leq \mu_b^*$ for some $\mu_b^* \in (0, 1)$.

Notice also that μ' is increasing in μ . Hence, $\mu' \leq \mu_H$ for all $\mu \leq \mu^S$ whenever the following condition is satisfied:

$$\frac{1 - \mu^S}{\mu^S} \geq \frac{1 - \mu_H}{\mu_H} \frac{\Pr(b, B|H)}{\Pr(b, B|L)}, \quad (14)$$

where μ^S is derived in (12). Thus, substituting for μ^S , the above condition reduces to:

$$\frac{\Pr(b, B|H)}{\Pr(b, G|H)} \leq \frac{\frac{\iota}{C}}{1 - \frac{\iota}{C}} \mu_H.$$

Therefore, with the upper bound $\frac{\iota}{C} \left(1 - \frac{\iota}{C}\right)^{-1} \mu_H$ on $\frac{\Pr(b, B|H)}{\Pr(b, G|H)}$, we have $J(\mu') = J(\mu_H)$ for each $\mu \leq \mu^S$. Hence, for each $\mu \leq \mu^S$, in (13), the terms

$$\mu_b u^H(\iota) + (1 - \mu_b) u^L(\iota) + \delta (1 - \iota) (\alpha \mu_b J(1) + (1 - \alpha \mu_b) J(\mu'))$$

are linear in μ_b , while $J(\mu_b)$ is convex in μ_b . Together with the facts that (i) at $\mu \geq \mu^S$, no intervention is optimal and (ii) at $\mu_b = 0$, intervention is optimal, there exists a unique μ_b^* such that, conditional on $s = b$, intervention is optimal if and only if $\mu_b \leq \mu_b^*$.

A.4 Proof of Lemma 4

Let \underline{SW} be P 's value under the no effort equilibrium (per-period utility is $(1 - \delta) \underline{SW}$). In this equilibrium, after each public history, P replaces A with probability one, and A exerts zero effort. This is a sequential equilibrium since (i) it is optimal for P to replace the agent given that the future agent supplies zero effort and (ii) it is optimal for the agent to supply zero effort since he is replaced for sure.

By feasibility, $V \in [0, \frac{1}{1-\delta}]$ since $\frac{1}{1-\delta}$ is the maximum average payoff that the principal can deliver to the agent by letting $\rho_z = e_z = 0$ for each z . If $V = \frac{1}{1-\delta}$, then A does not work and is kept forever, so we have $J(\mu, \frac{1}{1-\delta}) = \underline{SW}$ for each μ . If $V = 0$, then P replaces A right away, and so $J(\mu, 0) = \bar{J} \geq \underline{SW}$ for each μ . Since $J(\mu, V)$ is concave in V by Lemma 5,²⁸ for each μ and $V \in [0, \frac{1}{1-\delta}]$, we have $J(\mu, V) \geq \underline{SW}$.

Consider the following strategy: On the equilibrium path, the public history h^t decides μ and V . Given z , as long as P chooses ρ_z and ι_z corresponding to the solution to the dynamic program, A supplies e_z . A 's deviation is ignored. If P deviates from this equilibrium path, then P chooses $\rho_z = 1$ and $\iota_z = 1$ forever, and A chooses $e_z = 0$ forever (switching to the no effort equilibrium). Incentive compatibility (9) ensures the agent's incentive to choose e_z , and $J(\mu, V) \geq \underline{SW}$ ensures the principal's incentive.

It will be useful to verify that the payoff at the arrival of a new agent is higher than this no effort equilibrium: $\bar{J} > \underline{SW}$. To see why, the principal can improve upon \underline{SW} as follows: In this equilibrium, for each z , the principal always takes $\iota_z = 1$ as in the no effort equilibrium. If $\omega = gG$, then P keeps the agent forever. Otherwise, P replaces the agent (and goes back to the no effort equilibrium). That is, P rewards the agent after a good outcome in the first period, which incentivizes the high-type agent to supply a positive effort. Hence, the principal can obtain a payoff greater than \underline{SW} in the first period, and then obtain the continuation payoff of $\delta \underline{SW}$. In total, P obtains a payoff higher than \underline{SW} .

Given $\bar{J} > \underline{SW}$, for each (μ, V) with $V \in (0, \frac{1}{1-\delta})$, by concavity of $J(\mu, \cdot)$, we have

$$J(\mu, V) \geq \frac{\frac{1}{1-\delta} - V}{\frac{1}{1-\delta}} \bar{J} + \frac{V}{\frac{1}{1-\delta}} J\left(\mu, \frac{1}{1-\delta}\right) > \underline{SW}. \quad (15)$$

A.5 Proof of Lemma 5

Part 1. Concavity with respect to V .

We show that $J(\mu, V)$ is concave in V for a fixed μ . Suppose $V = \beta V_1 + (1 - \beta) V_2$ for $V_1, V_2, \beta \in [0, 1]$; and let $\alpha[V_1]$ and $\alpha[V_2]$ be the optimal policies for (μ, V_1) and (μ, V_2) , respectively.

Suppose P chooses $\alpha[V_1]$ with probability β and $\alpha[V_2]$ with probability $1 - \beta$, according to the realization of the public randomization device.

²⁸Note that proof of Lemma 5 does not depend on Lemma 4.

1. Since $\alpha[V_1]$ delivers V_1 to the agent and $\alpha[V_2]$ delivers V_2 , the agent's expected payoff is $V = \beta V_1 + (1 - \beta) V_2$. Hence, promise keeping is satisfied.
2. Conditional on the realization of the public randomization device, since both $\alpha[V_1]$ and $\alpha[V_2]$ are incentive compatible, the agent's incentive compatibility is satisfied.
3. With probability β , the principal achieves $J(\mu, V_1)$, and with probability $1 - \beta$, she achieves $J(\mu, V_2)$, since we fixed μ . Hence she achieves $\beta J(\mu, V_1) + (1 - \beta) J(\mu, V_2)$.

Hence, the principal with (μ, V) achieves at least $\beta J(\mu, V_1) + (1 - \beta) J(\mu, V_2)$.

Part 2. Convexity with respect to μ .

Since V is fixed, the proof is the same as Lemma 1.

Part 3. Monotonicity with respect to μ .

Suppose that $J(\mu, V) = J$ for some μ and V . Then, for a higher value $\mu' > \mu$ and the same promised utility V , we have $J(\mu', V) \geq J$. Since V is fixed, the proof is the same as Lemma 1.

We now show it is strictly increasing for $V \in (0, \frac{1}{1-\delta})$. Fix public history h^t with (μ, V) with $V \in (0, \frac{1}{1-\delta})$ arbitrarily, and let $\alpha[\mu]$ be the principal's optimal strategy from this history. Given the starting belief $\mu' > \mu$, suppose the principal in period $\tau \geq t$ takes the same strategy $\alpha[\mu'] = \alpha[\mu]$ as long as $e_{z_\tau} = 0$ for each z_τ given $\alpha[\mu]$. Then, as long as $e_{z_\tau} = 0$ for z_τ given $\alpha[\mu]$, the payoff is exactly the same between μ and μ' (and the belief stays the same unless replacement happens); and once the current agent exerts a positive effort (if he is of H type), the principal's expected payoff is higher with μ' than with μ . Hence, we have $J(\mu', V) > J(\mu, V)$ if there exist $\tilde{t} \geq t$ and $z_{\tilde{t}}$ such that, given $\alpha[\mu]$, (i) $h^{\tilde{t}}$ happens with a positive probability, (ii) the same agent stays until period \tilde{t} given $h^{\tilde{t}}$, and (iii) $e_{z_{\tilde{t}}} > 0$.

We now show that there exists such $(h^{\tilde{t}}, z_{\tilde{t}})$. Suppose otherwise. Then, the principal's payoff is equal to $J(\mu, V) = \alpha J(\mu, \tilde{V}) + (1 - \alpha) \bar{J}$, where $1 - \alpha$ is the probability of immediate replacement and the promise keeping constraint implies $V = \alpha \tilde{V}$. That is,

$$J(\mu, V) = \frac{V}{\tilde{V}} J(\mu, \tilde{V}) + \left(1 - \frac{V}{\tilde{V}}\right) \bar{J}.$$

Suppose $\tilde{V} = 1$. Then, since $e = 0$, we have $J(\mu, \tilde{V}) = (1 - \delta) \underline{SW} + \delta \bar{J}$, and so $\bar{J} - J(\mu, \tilde{V}) = (1 - \delta) (\bar{J} - \underline{SW})$. Suppose next that $\tilde{V} = 1 + \Delta$. Then, the principal can implement $e = 0$ in period t , which makes the next-period promised value equal to $\frac{\tilde{V}-1}{\delta} = \frac{\Delta}{\delta}$. Hence, the principal can achieve the payoff at least

$$(1 - \delta) \underline{SW} + \delta \left(\frac{\Delta}{\delta} ((1 - \delta) \underline{SW} + \delta \bar{J}) + \left(1 - \frac{\Delta}{\delta}\right) \bar{J} \right).$$

Hence,

$$\begin{aligned}
& \frac{J(\mu, \tilde{V} + \Delta) - J(\mu, \tilde{V})}{\Delta} \\
& \geq \frac{(1 - \delta) \underline{SW} + \delta \left(\frac{\Delta}{\delta} \left((1 - \delta) \underline{SW} + \delta \bar{J} \right) + \left(1 - \frac{\Delta}{\delta} \right) \bar{J} \right) - (1 - \delta) \underline{SW} - \delta \bar{J}}{\Delta} \\
& \geq -(1 - \delta) (\bar{J} - \underline{SW}).
\end{aligned}$$

In total, we have

$$\frac{d}{d\tilde{V}} \left[\frac{V}{\tilde{V}} J(\mu, \tilde{V}) + \left(1 - \frac{V}{\tilde{V}} \right) \bar{J} \right] \Big|_{\tilde{V}=1} \geq 0.$$

Hence, the first order effect of increasing \tilde{V} by Δ keeping e fixed is no less than 0. Suppose that the principal increases V'_{gG} in the problem to maximize $J(\mu, \tilde{V})$, keeping all the other continuation payoffs fixed. This increases e and \tilde{V} . Since the first order effect of changing \tilde{V} given e is 0, the principal is strictly better off by implementing $e > 0$, as desired.

A.6 Proof of Lemma 6

We have $J(\mu, 0) = \bar{J}$ for each μ since P has to replace A right away. Hence we are left to prove the other four properties:

A.6.1 Part 1. There exists $V(\mu)$ such that $J(\mu, V)$ is linear for $V \in [0, V(\mu)]$.

Suppose such $V(\mu)$ does not exist. By Lemma 5, this means that $J(\mu, V)$ is strictly concave near $V = 0$.

Take $V \in (0, 1)$. This means that P needs to stochastically replace A , since otherwise A receives 1 by not working. Let β be the probability of a replacement. The promise keeping condition implies

$$\beta \times 0 + (1 - \beta) \times \hat{V} = V, \tag{16}$$

where $\hat{V} \geq 1$ is the promised utility conditional on A not being replaced.

P maximizes

$$\max_{\beta \in [0, 1], \hat{V} \in [0, \frac{1}{1-\delta}]} \beta J(\mu, 0) + (1 - \beta) J(\mu, \hat{V}) \tag{17}$$

subject to

$$\beta \times 0 + (1 - \beta) \times \hat{V} = V \text{ and } \hat{V} \geq 1.$$

Substituting the constraint, P 's payoff is

$$J(\mu, 0) + \frac{V}{\hat{V}} \left[J(\mu, \hat{V}) - J(\mu, 0) \right]. \tag{18}$$

Taking the derivative with respect to \hat{V} (the differentiability of $J(\mu, \hat{V})$ follows from the

Envelope Theorem), we obtain

$$V \frac{J(\mu, 0) + [J_2(\mu, \hat{V})\hat{V} - J(\mu, \hat{V})]}{\hat{V}^2},$$

where J_n is the derivative of J with respect to its n^{th} argument.

We show that the numerator is always negative for each $\hat{V} \geq 0$. With $\hat{V} = 0$, the numerator is 0. Taking the derivative of the numerator,

$$\frac{d}{d\hat{V}} \left\{ J(\mu, 0) + [J_2(\mu, \hat{V})\hat{V} - J(\mu, \hat{V})] \right\} = \hat{V} \frac{d^2}{d\hat{V}^2} J(\mu, \hat{V}).$$

Since we assumed $J(\mu, \cdot)$ is strictly concave, this is negative for each $\hat{V} \geq 0$. Therefore, the numerator is globally negative.

Hence, the smallest $\hat{V} = 1$ is optimal. Given $\hat{V} = 1$, by (18), we have

$$J(\mu, V) = J(\mu, 0) + V \times [J(\mu, 1) - J(\mu, 0)],$$

for $V \in [0, 1]$, which is linear in V .

A.6.2 Part 2. For $\mu \geq \mu_H$, we have $V(\mu) > 1$.

Suppose $\mu \geq \mu_H$. For the sake of contradiction, assume that $V \leq 1$ for each $V \in \arg \max_V J(\mu, V)$. Then, in the above problem (17), $\hat{V} = 1$ — the smallest continuation payoff without immediate replacement — is the unique optimum. Recall that β is defined as the probability of immediate replacement in (16). Hence P cannot replace A in the current period after P picks \hat{V} with probability $1 - \beta$. If P promised a positive continuation payoff from the next period, then since $c(0) = \lim_{e \rightarrow 0} c'(e) = 0$, A could obtain a payoff greater than 1 with providing a sufficiently small e . We therefore have to make sure that $V'_z[\hat{V}](\omega) = 0$ for each z and ω , and so $e_z = 0$ for each z . Therefore, the effort has to be equal to 0. Then, P 's instantaneous payoff is $(1 - \delta)\underline{SW}$. Moreover, since $V'_z[\hat{V}](\omega) = 0$ for each z and ω , the agent will be replaced in the next period with probability one. Hence, the continuation payoff is $\delta\bar{J}$. Since $\beta = 0$ if the current promised value is 1 and $\hat{V} = 1$, we have

$$J(\mu, 1) = (1 - \delta)\underline{SW} + \delta\bar{J}. \quad (19)$$

For $\mu = \mu_H$, (19) together with (15) implies that $J(\mu_H, 0) = \bar{J}$ and $J(\mu_H, V)$ is linear and less than \bar{J} for each $V \in (0, 1]$. By concavity, this means that $J(\mu_H, V) < \bar{J}$ for each $V > 0$. Thus, $\arg \max_V J(\mu_H, V) = 0$. This means that \bar{J} is uniquely obtained by always replacing A ; however, this implies that A exerts no effort, which is a contradiction. Hence, $V(\mu_H) > 1$. Moreover, since $\bar{J} = \max_V J(\mu_H, V)$, it follows that

$$J(\mu_H, V) = \bar{J} \text{ for } V \in [0, V(\mu_H)]. \quad (20)$$

For $\mu > \mu_H$, by Lemma 5, we have $J(\mu, 1) > J(\mu_H, 1) \geq \bar{J}$, which contradicts (19). Hence, $V(\mu) > 1$ as well.

A.6.3 Part 3. The Slope of the Linear Part.

Since $J(\mu, V)$ is strictly increasing in $\mu \in (0, 1)$, and $J(\mu, 0) = \bar{J}$ for each μ , (20) implies the slope of the linear part is negative for $\mu < \mu_H$ and positive for $\mu > \mu_H$.

A.6.4 Part 4. Property of $V \in \arg \max_{\hat{V}} J(\mu, \hat{V})$

Without loss of generality, we can take $V \in \arg \max_{\hat{V}} J(\mu, \hat{V})$ such that V is the extreme point of the graph $\{\hat{V}, J(\mu, \hat{V})\}_{\hat{V}}$. This means that no mixture can implement $(V, J(\mu, V))$. Hence, P 's payoff $J(\mu, V)$ at $V \in \arg \max_{\hat{V}} J(\mu, \hat{V})$, denoted by $J(\mu)$, is determined by the dynamic program without mixture:

$$J(\mu) = \max_{(e, \iota, V')} \{u^P(\iota|\mu, e) + \delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega))\},$$

subject to incentive compatibility constraint:

$$e \in \arg \max 1 - c(e) + \delta \sum_{\omega} \Pr(\omega|e, \iota) V'(\omega).$$

Note that we do not impose the promise keeping constraint since we are free to choose \hat{V} to maximize $J(\mu, \hat{V})$. Moreover, since the first-order condition for e is always necessary and sufficient by the assumption of the cost function c , we can see the above dynamic program as deciding $(V'(\omega))_{\omega}$, and then e is determined by the first-order condition.

In this problem, we first show that $V'(\omega) \leq \arg \max_{\hat{V}} J(\mu'(\mu, e, \omega), \hat{V})$ after $\mu'(\mu, e, \omega) \leq \mu$. Suppose otherwise: There exists $\bar{\omega}$ such that $V'(\bar{\omega}) > \arg \max_{\hat{V}} J(\mu'(\mu, e, \bar{\omega}), \hat{V})$ after $\mu'(\mu, e, \bar{\omega}) \leq \mu$.

Since

$$\mu'(\mu, e, \bar{\omega}) = \frac{\mu \Pr(\bar{\omega}|e, \iota)}{\mu \Pr(\bar{\omega}|e, \iota) + (1 - \mu) \Pr(\bar{\omega}|0, \iota)} \leq \mu, \quad (21)$$

we have $\Pr(\bar{\omega}|0, \iota) \geq \Pr(\bar{\omega}|e, \iota)$. We assume $\Pr(\omega|e, \iota)$ is monotone in e for each ω and ι , so the probability $\Pr(\bar{\omega}|e, \iota)$ is decreasing in e .

Then, the first-order condition for the optimality of $V'(\bar{\omega})$ is

$$\begin{aligned}
0 &= \frac{d}{dV'(\bar{\omega})} \{u^P(\iota|\mu, e) + \delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega))\} \\
&= \{u_e^P(\iota|\mu, e) \\
&\quad + \delta \sum_{\omega} \Pr_e(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) \\
&\quad + \delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J_1(\mu'(\mu, e, \omega), V'(\omega)) \mu'_e(\mu, e, \omega)\} \frac{de}{dV'(\bar{\omega})} \\
&\quad + \delta \Pr(\bar{\omega}|\mu, e, \iota) J_2(\mu'(\mu, e, \bar{\omega}), V'(\bar{\omega})),
\end{aligned}$$

where J_n is the derivative of J with respect to its n^{th} argument; and $u_e^P \geq 0$, \Pr_e , and μ'_e are the derivatives of u^P , \Pr , and μ' with respect to e , respectively. Since $\Pr(\bar{\omega}|e, \iota)$ is decreasing in e , it follows that $\frac{de}{dV'(\bar{\omega})} < 0$. Moreover, $J_2(\mu'(\mu, e, \bar{\omega}), V'(\bar{\omega})) < 0$, since $V'(\bar{\omega}) > \arg \max_{\hat{V}} J(\mu'(\mu, e, \bar{\omega}), \hat{V})$ and J is concave. Hence, we have

$$\begin{aligned}
&\{u_e^P(\iota|\mu, e) + \delta \sum_{\omega} \Pr_e(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) \\
&\quad + \delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J_1(\mu'(\mu, e, \omega), V'(\omega)) \mu'_e(\mu, e, \omega)\} < 0.
\end{aligned} \tag{22}$$

Similarly, if there exists $\hat{\omega}$ such that $\Pr(\hat{\omega}|e, \iota)$ is decreasing in e but $V'(\hat{\omega}) \leq \arg \max_{\hat{V}} J(\mu'(\mu, e, \hat{\omega}), \hat{V})$, then the symmetric argument implies that

$$\begin{aligned}
&\{u_e^P(\iota|\mu, e) + \delta \sum_{\omega} \Pr_e(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) \\
&\quad + \delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J_1(\mu'(\mu, e, \omega), V'(\omega)) \mu'_e(\mu, e, \omega)\} \geq 0,
\end{aligned}$$

which is a contradiction.

Therefore, letting Ω_- be the set of signal-outcome pairs ω such that $\Pr(\omega|e, \iota)$ is decreasing in e , for each $\omega \in \Omega_-$, we have $V'(\omega) > \arg \max_{\hat{V}} J(\mu'(\mu, e, \omega), \hat{V})$. Symmetrically, letting Ω_+ be the set of ω such that $\Pr(\omega|e, \iota)$ is increasing in e , for each $\omega \in \Omega_+$, we have $V'(\omega) < \arg \max_{\hat{V}} J(\mu'(\mu, e, \omega), \hat{V})$.

Now we set $V^*(\omega) = \arg \max_{\hat{V}} J(\mu'(\mu, e, \omega), \hat{V})$ for each ω , and let e^* be the new optimal effort (fixing ι throughout). Since $V^*(\omega) < V'(\omega)$ for $\omega \in \Omega_-$ and $V^*(\omega) > V'(\omega)$ for $\omega \in \Omega_+$, we have $e^* > e$ (here, e is the original effort). Hence, we have

$$u^P(\iota|\mu, e^*) > u^P(\iota|\mu, e). \tag{23}$$

In addition, we adjust $V^*(\omega)$ so that the continuation payoff increases with fixed e :

$$\sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) < \sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)). \quad (24)$$

Moreover, since $\max_{\hat{V}} J(\mu', \hat{V})$ is increasing in μ' , we have

$$J(\mu'(\mu, e, \omega), V^*(\omega)) < J(\mu'(\mu, e, \hat{\omega}), V^*(\hat{\omega}))$$

for each $\omega \in \Omega_-$ and $\hat{\omega} \in \Omega_+$. Since increase in e increases the probability of event ω if and only if $\omega \in \Omega_+$, we have

$$\sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)) < \sum_{\omega} \Pr(\omega|\mu, e^*, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)). \quad (25)$$

Finally, learning (the difference between $\mu'(\mu, e, \omega)$ and $\mu'(\mu, e^*, \omega)$) further increases the continuation payoff. To show this, we first make the following claim:

Claim 1 For $\mu_1 < \mu_2$, $V^*(\mu_1) \in \arg \max_{\hat{V}} J(\mu_1, \hat{V})$ and $V^*(\mu_2) \in \arg \max_{\hat{V}} J(\mu_2, \hat{V})$, we have $J_1(\mu_1, V^*(\mu_1)) \leq J_1(\mu_2, V^*(\mu_2))$.

Proof. We have

$$J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \leq J(\mu_2, V^*(\mu_1)),$$

since J is convex in μ , and

$$J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \leq J(\mu_2, V^*(\mu_2)),$$

since $V^*(\mu_2)$ maximizes $J(\mu_2, V)$ at μ_2 . At the same time,

$$J(\mu_1, V^*(\mu_1)) \geq J(\mu_2, V^*(\mu_2)) - J_1(\mu_2, V^*(\mu_2))[\mu_2 - \mu_1],$$

since J is convex in μ , and from the first inequality of the proof,

$$\begin{aligned} J(\mu_1, V^*(\mu_1)) &\geq J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \\ &\quad - J_1(\mu_2, V^*(\mu_2))[\mu_2 - \mu_1]. \end{aligned}$$

Hence,

$$0 \geq [J_1(\mu_1, V^*(\mu_1)) - J_1(\mu_2, V^*(\mu_2))](\mu_2 - \mu_1).$$

■

Given this claim, $J_1(\mu'(\mu, e, \omega), V^*(\omega))$ is larger for ω with $\mu'(\mu, e, \omega) > \mu$ than for ω with $\mu'(\mu, e, \omega) < \mu$. Since the distribution of $\{\mu'(\mu, e^*, \omega)\}_{\omega}$ given e^* is the mean-preserving spread of the distribution of $\{\mu'(\mu, e, \omega)\}_{\omega}$ given e and we have $\mu'(\mu, e^*, \omega) \geq \mu'(\mu, e, \omega)$ if and only if ω satisfies $\mu'(\mu, e, \omega) \geq \mu$, faster learning increases the continuation payoff.

Together with (24) and (25), this leads to

$$\sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) < \sum_{\omega} \Pr(\omega|\mu, e, \iota^*) J(\mu'(\mu, e^*, \omega), V^*(\omega)).$$

Together with (23), we have proven that P 's payoff increases.

The proof for $V'(\omega) \geq \arg \max_{\hat{V}} J(\mu'(\mu, e, \omega), \hat{V})$ after $\mu'(\mu, e, \omega) \geq \mu$ is completely symmetric, and so it is omitted.

A.7 Proof of Lemma 7

Recall that we refer to intervention as the intervention decision after signal $s = b$, since P never intervenes after $s = g$. Given $s = g$, the principal observes the same information regardless of $\iota(s = b)$. Given $s = b$, she can observe $o \in \{G, B\}$ after $s = b$ without intervention while she can only observe $o = I$ with intervention. Hence, intervention is more informative in the Blackwell sense. Hence, given e , the distribution of the updated beliefs $(\mu'(\mu, e, \omega))_{\omega}$ given no intervention is a mean-preserving spread of that given intervention.

In particular, the belief update is given by

$$\begin{aligned} \mu'(\mu, e, b, I) &= \frac{\mu \Pr(b|e)}{\mu \Pr(b|e) + (1 - \mu) \Pr(b|0)}, \\ \mu'(\mu, e, b, G) &= \frac{\mu \Pr(b, G|e)}{\mu \Pr(b, G|e) + (1 - \mu) \Pr(b, G|0)}, \\ \mu'(\mu, e, b, B) &= \frac{\mu \Pr(b, B|e)}{\mu \Pr(b, B|e) + (1 - \mu) \Pr(b, B|e)}. \end{aligned}$$

Hence, the difference in the variance of $\mu'(\mu, e, \omega)$ is given by

$$\begin{aligned} d(\mu) &: = \sum_{\omega} \Pr(\omega|\mu, e, \iota = 0) (\mu'(\mu, e, \omega) - \mu)^2 - \sum_{\omega} \Pr(\omega|\mu, e, \iota = 1) (\mu'(\mu, e, \omega) - \mu)^2 \\ &= \sum_{y \in \{G, B\}} \frac{\mu^2 (1 - \mu)^2 (\Pr(b, y|e) - \Pr(b, y|0))^2}{\mu \Pr(b, y|e) + (1 - \mu) \Pr(b, y|0)} - \frac{\mu^2 (1 - \mu)^2 (\Pr(b|e) - \Pr(b|0))^2}{\mu \Pr(b|e) + (1 - \mu) \Pr(b|0)}. \end{aligned}$$

Note that this difference is 0 with $\mu = 0$ and $\mu = 1$. Moreover, taking the second derivative of $d(\mu)$ with respect to μ yields

$$\sum_{y \in \{G, B\}} \frac{\Pr(b, y|e)^2 \Pr(b, y|0)^2}{(\mu \Pr(b, y|e) + (1 - \mu) \Pr(b, y|0))^3} - \frac{\Pr(b|e)^2 \Pr(b|0)^2}{(\mu \Pr(b|e) + (1 - \mu) \Pr(b|0))^3}.$$

The function $f(x, y) := \frac{x^2 y^2}{(\mu x + (1-\mu)y)^3}$ is convex since, for each $(a, b) \in \mathbb{R}^2$, we have

$$(a, b) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{2(b^2 x^2 - abxy + a^2 y^2)(\mu^2 x^2 + 4\mu(1-\mu)xy + y^2)}{(\mu x + (1-\mu)y)^5} \geq 0$$

since $b^2 x^2 - abxy + a^2 y^2 = (bx + ay)^2 - abxy = (bx - ay)^2 + abxy$. Given $\Pr(b|e) = \Pr(b, G|e) + \Pr(b, B|e)$, we have $d''(\mu) \leq 0$.

A.8 Proof of Proposition 2

The proof consists of the three steps: (1) proving that intervention is optimal in the initial period, (2) intervention is optimal for a sufficiently large T , and (3) in some period $t \geq 2$, no intervention is optimal.

A.8.1 Intervention is Optimal in the Initial Period

Lemma 8 *There exist $\bar{\mu}_H \in (0, 1)$ and $\bar{q} \in (0, 1)$ such that, for each $\mu_H \leq \bar{\mu}_H$ and $\Pr(G|0) \leq \bar{q}$, it is optimal to intervene after $s = b$ in the initial period of an agent's appointment.*

Proof. In period 1 of the agent's appointment, after an $s = b$, the belief is no more than μ_H . Hence, the instantaneous cost of non-intervention is no less than

$$-l - [\mu_H \times \Pr(G|e, b) \times 0 + (1 - \mu_H)(-C)] \geq C - l - \mu_H C.$$

On the other hand, the gain in the continuation payoff of no intervention is at most

$$\delta \left[\mu_H \times 0 + (1 - \mu_H) \max_V J(\mu_H, V) \right] - \delta J(\mu_b, V_b).$$

We now drive an upper bound for $\max_V J(\mu_H, V)$:

$$\max_V J(\mu_H, V) \leq \mu_H \times 0 + (1 - \mu_H) \left(\Pr(y = B|0)(-l) + \delta \max_V J(\mu_H, V) \right).$$

Here, the H -type would deliver the best outcome, the L -type would be replaced immediately after period 1, and we allow the principal to intervene if and only if the outcome is bad, so that we drive an upper bound. Rearranging,

$$\max_V J(\mu_H, V) \leq \frac{-(1 - \mu_H)(1 - \bar{q})l}{1 - (1 - \mu_H)\delta}.$$

In contrast, $J(\mu_b, V_b) \geq \frac{-l}{1-\delta}$ since the principal can always intervene. Hence, the continuation payoff gain is bounded by

$$\delta \left((1 - \mu_H) \frac{-(1 - \mu_H)(1 - \bar{q})l}{1 - (1 - \mu_H)\delta} - \frac{-l}{1 - \delta} \right).$$

Hence, if

$$C - l - \mu_H C > \delta \left((1 - \mu_H) \frac{-(1 - \mu_H)(1 - \bar{q})l}{1 - (1 - \mu_H)\delta} - \frac{-l}{1 - \delta} \right), \quad (26)$$

then intervention is uniquely optimal. At $\mu_H = 0$ and $\bar{q} = 0$, (26) holds since we have $C - l > 0$. Therefore, there exist $\bar{\mu}_H > 0$ and $\bar{q} > 0$ such that, for $\mu_H \leq \bar{\mu}_H$ and $\Pr(G|0) \leq \bar{q}$, we have (26). ■

A.8.2 Intervention at the Limit

Lemma 9 *For any parameter values, after ω with $\Pr_e(\omega|e) < 0$, we have $J_2(\mu'(\mu, e_z, \omega), V'_z(\omega)) = 0$ if we start from $\mu = \mu_H$ and $V = \arg \max_{\bar{V}} J(\mu_H, \bar{V})$.²⁹*

Proof. From Lemma 1, we have $J(\mu'(\mu, e_z, \omega), V'_z(\omega)) = \max_V J(\mu_H, V)$. Hence, Lemma 6 implies the result. ■

We now form the Lagrangian:

$$\begin{aligned} J(\mu, V) &= \int_z (1 - \rho_z) \bar{J} + \rho_z u^P(\iota_z | \mu, e_z) + \delta \sum_{\omega} \Pr(\omega | \mu, e_z, \iota_z) J(\mu'(\mu, e_z, \omega), V'_z(\omega)) dz \\ &+ \lambda \left(V - \int_z \rho_z \left\{ 1 - c(e_z) + \delta \sum_{\omega} \Pr(\omega | e_z, \iota_z) V'_z(\omega) \right\} dz \right) \\ &+ \int_z \rho_z \eta_z \left(\delta \sum_{\omega} \Pr_e(\omega | e_z, \iota_z) V'_z(\omega) - c'(e_z) \right) dz \end{aligned}$$

with $\eta_z \geq 0$ (higher effort is beneficial). Recall that $\Pr(\omega | e_z, \iota_z) = \Pr(\omega | \mu = 1, e_z, \iota_z)$. By the Envelope theorem, $J_2(\mu, V) = \lambda$. Taking the first order conditions and substituting $J_2(\mu, V) = \lambda$, we obtain

$$\begin{aligned} e_z &: -J_2(\mu, V) c'(e_z) + \eta_z c''(e_z) \\ &= u_e^P(\iota_z | \mu, e_z) + \delta \cdot \sum_{\omega} \Pr_e(\omega | \mu, e_z, \iota_z) \cdot J(\mu'(\mu, e_z, \omega), V'_z(\omega)) \\ &\quad + \delta \cdot \sum_{\omega} \Pr(\omega | \mu, e_z, \iota_z) \cdot J_1(\mu'(\mu, e_z, \omega), V'_z(\omega)) \cdot \mu'_e(\mu, e_z, \omega) \\ &\quad - \delta \cdot \sum_{\omega} \Pr_e(\omega | e_z, \iota_z) \cdot V'_z(\omega) \cdot J_2(\mu, V) \\ &\quad + \delta \cdot \eta_z \cdot \sum_{\omega} \Pr_{ee}(\omega | e_z, \iota_z) \cdot V'_z(\omega), \end{aligned} \quad (27)$$

²⁹Recall that J_n is the derivative of J with respect to its n^{th} argument.

and

$$V'_z(\omega) : J_2(\mu'(\mu, e_z, \omega), V'_z(\omega)) = \frac{\Pr(\omega|e_z, \iota_z)}{\Pr(\omega|\mu, e_z, \iota_z)} J_2(\mu, V) - \eta_z \frac{\Pr_e(\omega|e_z, \iota_z)}{\Pr(\omega|\mu, e_z, \iota_z)}. \quad (28)$$

Using these two first order conditions, we will show that the effort level converges to 0.

Lemma 10 *On the equilibrium path, given a history h such that the belief updates positively, $\mu(h^t) \geq \mu(h^{t-1})$ for each t , effort converges to 0.*

Proof. Fix $(z_t, \omega_t)_{t=1}^\infty$ to satisfy $\mu(h^t) \geq \mu(h^{t-1})$ for each t , and let $(\iota_t, e_t)_{t=1}^\infty$ be the implemented intervention decisions and effort levels along the history. For notational simplicity, we omit $(z_t)_{t=1}^\infty$ since the argument holds conditional on $(z_t)_{t=1}^\infty$.

On such a history, we have $J_2(\mu'(\mu, e_1, \omega_1), V'(\omega_1)) < 0$. To see why, since $\Pr_e(\omega_1|e_1, \iota_1) > 0$ given $\mu(h^t) \geq \mu(h^{t-1})$ and $J_2(\mu_1, V_1) = 0$ in the initial period, given (28), it suffices to show that $\eta > 0$. If $\eta = 0$, since $J_2(\mu, V) = 0$ in the initial period, Lemma 9 and (27) yield

$$\begin{aligned} 0 &= u_e^P(\iota_1|\mu_1, e_1) + \delta \sum_{\tilde{s}_1, \tilde{\delta}_1} \Pr_e(\tilde{\omega}_1|\mu_1, e_1, \iota_1) \cdot J(\mu'(\mu_1, e_1, \tilde{\omega}_1), V'(\tilde{\omega}_1)) \\ &\quad + \delta \sum_{\tilde{\omega}_1: \mu'_e(\mu_1, e_1, \tilde{\omega}_1) > 0} \Pr(\tilde{\omega}_1|\mu_1, e_1, \iota_1) \cdot J_1(\mu'(\mu_1, e_1, \tilde{\omega}_1), V'(\tilde{\omega}_1)) \cdot \mu'_e(\mu_1, e_1, \tilde{\omega}_1). \end{aligned}$$

The first two terms of the right hand side is the benefit of increasing e_1 to the principal's value fixing ι_1 and $V'(\tilde{\omega}_1)$; and the last term is non-negative given $J_1(\mu, V) \geq 0$. Hence, the right hand side is positive.³⁰ This is a contradiction.

In addition, on such a history, we have $\Pr_e(\omega_t|e_t, \iota_t) \geq 0$ and $\Pr(\omega_t|e_t, \iota_t) \geq \Pr(\omega_t|\mu_t, e_t, \iota_t)$ for each t . Hence, recursively applying to (28), we have that

$$J_2(\mu'(\mu_t, e_t, \omega_t), V_{t+1}(\omega_t)) \leq \frac{\Pr(\omega_t|e_t, \iota_t)}{\Pr(\omega_t|\mu_t, e_t, \iota_t)} J_2(\mu_t, V_t) - \eta_t \frac{\Pr_e(\omega_t|e_t, \iota_t)}{\Pr(\omega_t|\mu_t, e_t, \iota_t)},$$

so it is monotonically decreasing. If e_t does not converge to 0, then μ_t converges to 1 and $\eta_t \geq 0$ converges to 0, since otherwise J_2 diverges to $-\infty$.

³⁰Otherwise, the principal should have implemented $e_1 = 0$ and $V'(\tilde{\omega}_1) = V'(\tilde{\omega}'_1)$ for each $\tilde{\omega}_1, \tilde{\omega}'_1$ given concavity of $J(\mu, V)$. However, (i) the first order condition for e (this is necessary and sufficient given our assumption), (ii) Lemma 9, and (iii) parts A.6.2 and A.6.3 of the proof to Lemma 6, (omitting z for a simple notation) we have

$$c'(e_1) = \delta \sum_{\omega_1: \Pr_e(\omega_1|e_1) > 0} \Pr_e(\omega_1|\iota_1, e_1),$$

which means $e_1 > 0$. This is a contradiction.

Suppose μ_t converge to 1 and η_t converges to 0. At this limit, (27) converges to

$$\begin{aligned} -J_2(1, V) c'(e) &= u_e^P(1, e, \iota) + \delta \sum_{\omega} \Pr_e(\omega|1, e, \iota) J(\mu'(1, e, \omega), V'(\omega)) \\ &\quad + \delta \sum_{\omega} \Pr(\omega|e, \iota) J_1(1, V'(\omega)) \mu'_e(1, e, \omega) \\ &\quad - \delta \sum_{\omega} \Pr_e(\omega|e, \iota) V'(\omega) J_2(1, V). \end{aligned}$$

Since

$$\mu'_e(1, e, \omega) = \lim_{\mu \rightarrow 1} \left(\frac{d}{de} \frac{\mu \Pr(\omega|e)}{\mu \Pr(\omega|e) + (1 - \mu) \Pr(\omega|0)} \right) = 0$$

for each $\Pr(\omega|e)$ with $e > 0$ (recall that we assumed that $e > 0$ for the sake of a contradiction) and $c'(e) = \delta \sum_{\omega} \Pr(\omega|e, \iota) V'(\omega)$ from (9), we have

$$0 = u_e^P(\iota|1, e) + \delta \sum_{\omega} \Pr_e(\omega|e, \iota) J(1, V'(\omega)).$$

This means that the benefit of increasing e to the principal's value fixing $V'(\omega)$, i.e.,

$$\frac{d}{de} [u^P(\iota|1, e) + \delta \sum_{\omega} \Pr(\omega|e, \iota) J(1, V'(\omega))],$$

is 0. This in turn implies that e is equal to 0. Therefore, e_t converges to 0. ■

Given that e converges to 0, intervention is optimal at the limit:

Lemma 11 *There exists $\hat{e} \in (0, 1)$ such that, for any belief $\mu \in [0, 1]$ and promised value V , if the principal implements $e \leq \hat{e}$, then $\iota = 1$ is optimal.*

Proof. With discounting, $e \in [0, 1]$, and $V \in [0, \frac{1}{1-\delta}]$, the principal's payoff is continuous in e . Hence, it suffices to show that it is uniquely optimal for the principal to choose $\iota = 1$ for effort $e = 0$. With $e = 0$, we have $\mu'(\mu, e, \omega) = \mu$. Since $J(\mu, V)$ is concave in V , it is optimal to choose $V'(\omega|\iota) = V'(\omega'|\iota)$ for each ω, ω' . Hence, the continuation payoff is fixed regardless of ι . Since $\iota = 1$ maximizes the instantaneous utility $u^P(\iota|\mu, e, s)$ after $s = b$ given $e = 0$, intervention $\iota(s = b) = 1$ is uniquely optimal. ■

A.8.3 No intervention is Optimal in a Period after the Initial Period

The following lemma ensures that e_1 is bounded below:

Lemma 12 *For sufficiently small $\bar{q} > 0$, if $c'(\bar{e}) \leq \Pr_e(g, G|\bar{e}) \cdot \bar{q}$, then the initial effort level $e(\emptyset)$ is no less than \bar{e} .*

Proof. From (i) the first order condition for e (this is necessary and sufficient given our assumption), (ii) Lemma 9, and (iii) claims A.6.2 and A.6.3 of Lemma 6, (omitting z for a simple notation) we have

$$c'(e_1) = \delta \sum_{\omega_1: \Pr_e(\omega_1|e_1) > 0} \Pr_e(\omega_1|e_1, \iota_1) \geq \delta \Pr_e(g, G|e_1, \iota_1) = \delta \Pr_e(g, G|e_1).$$

Hence, we have $e_1 \geq \bar{e}$. ■

Lemma 13 *For each μ_H and $(\Pr(b|e))_{e \in [0,1]}$, there exists $\bar{q} > 0$ such that, if the effort provision condition holds given \bar{q} and $\Pr(G|0) \leq \bar{q}$, then there exists $t \geq 2$ such that no intervention is optimal in period t .*

Proof. It suffices to show that there exists $t \geq 2$ with $e \geq \bar{e}$, and $\mu'(h^t)$ is sufficiently close to 1 since then no intervention is statically optimal. Note that we first fix $(\Pr(b|e))_{e \in [0,1]}$. Hence, if $\mu'(h^t)$ is sufficiently close to one, $\mu'(h^t, b)$ is also close to one.

On the one hand, if there is no period $t \geq 2$ such that no intervention is optimal along the path of repeated (g, G) . Then, the payoff is bounded by

$$u^P(\iota_1|\mu_H, \bar{e}) + \delta \max \left\{ \max_V J(\mu_H, V), \frac{1}{1-\delta} \Pr(s = b|\bar{e})(-l) \right\}.$$

Here, to obtain an upper bound, we allow the principal to replace the L -type at the end of period 1, and she learns the H -type at the end of period 1 (we then take the maximum of these two continuation payoffs). In the latter event, intervention is optimal after $s = b$ since (i) there is no learning benefit if the principal learns the type and (ii) if the belief were sufficiently high for no intervention to be statically optimal after some history, then it would get sufficiently high along the path of repeated (g, G) .

On the other hand, if the principal implements $e_t = \bar{e}$ without replacement for each $t = 1, \dots, T$ as long as $\omega = (g, G)$, she obtains the payoff at least

$$\begin{aligned} & u^P(\iota_1|\mu_H, \bar{e}) \\ & + \sum_{t=1}^T \delta^{t-1} \left\{ \prod_{\tau=1}^{t-1} \Pr(\omega_\tau = (g, G)) \right\} \Pr(\omega_t \neq (g, G)) \left(-C + \delta \max_V J(\mu_H, V) \right) \\ & + \delta^{T-1} \prod_{\tau=1}^T \Pr(\omega_\tau = (g, G)) \frac{-l}{1-\delta}, \end{aligned}$$

where the probability is determined by the initial belief μ_H and the high type taking \bar{e} . The second line says that, until $\omega_t \neq (g, G)$ is first observed, no cost is incurred, and once $\omega_t \neq (g, G)$ happens, the principal pays $-C$ and replace the agent. The last line says that, if $\omega_t \neq (g, G)$ never happens until period T , then $s = b$ happens all the time and the principal always intervenes for $t = T + 1, \dots$

For each μ_H , for sufficiently large $\Pr(g, G|\bar{e})$ and sufficiently small $\Pr(g, G|0)$, $\mu'(\mu_H, g, G)$ is sufficiently close to 1 and $u^P(\iota_1|\mu_H, 1)$ and $u^P(\iota_1|\mu_H, \bar{e})$ are close to each other. Hence, at $T = \infty$ (namely, $\delta^{T-1} = 0$), the latter is larger.

For each T , for sufficiently small \bar{q} , it is possible to implement $e_t \geq \bar{e}$ for each $t = 1, \dots, T$ by keeping the agent if and only if he generates the outcome (g, G) . Hence, $\lim_{\bar{q} \rightarrow 0} T = \infty$. Therefore, for sufficiently small \bar{q} , there exists $t \geq 2$ with $e \geq \bar{e}$, and $\mu'(h^t)$ sufficiently close to 1. ■