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

Dana Foarta
Stanford GSB

Takuo Sugaya
Stanford GSB

August 27, 2019

Abstract

We study the optimal intervention policy to stop projects in a relational contract between a principal and a policymaker. The policymaker is privately informed about his ability and privately chooses how much effort to exert. Before a project is completed, the principal receives a signal about its outcome and can intervene to stop it. Intervention may prevent a bad outcome, but no intervention leads to better learning about the policymaker's ability. In the benchmarks with observable effort or observable ability, optimal intervention follows a threshold rule. With unobservable effort and ability, the optimal policy switches between intervention and no intervention.

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, Federica Izzo, Ken Shotts, Andy Skrzypacz, Victoria Vanasco, Richard van Weelden, as well as seminar partipants at Arizona State, Barcelona GSE, Boston University, Caltech, LSE, Stanford, University of Arizona, University of Chicago, UCSD, 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.

Electronic copy available at: https://ssrn.com/abstract=3243123

1 Introduction

Most institutions include specific mechanisms for interventions to overrule a policymaker and stop an ongoing 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, 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 may be reversed by simply overruling the policymaker rather than removing him from his position. In fact, intervention is usually more likely than outright removal of a policymaker.[1]

Interventions, nevertheless, come with trade-offs. Stopping a potentially faulty project saves resources. Yet, if a project is stopped, its outcome is never seen. Policy is changed, reforms are interrupted, funds are pulled, and the observable events differ from what the policymaker set out to do. This may hamper information acquisition about policy or policymaker effectiveness. Moreover, the expectation of interventions may distort the policymaker's ex-ante choices. Given the ubiquitous use of interventions and the non-obvious implications of these trade-offs, an open question in the design of institutions is how and when to use interventions. 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 duration of which is one period. The project's outcome 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 roadblock that, unless overcome, will result in a bad outcome. After observing the signal, the principal may intervene by paying 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

[1] For instance, interventions by parliaments through legislative reviews are more common than motions of censure that remove a minister or government.

Electronic copy available at: https://ssrn.com/abstract=3243123

revealed. At the end of the period, the principal may continue to the next period with the current agent, who will then start a new project, or she may end the relationship with the agent and draw another agent (or, alternatively, take an outside option).

This simple model is relevant for several applied settings. A prime example is that of transfers from an international funding institution (e.g., the World Bank) towards projects implemented by a local policymaker, e.g., a mayor or head of agency who runs the project locally. The relationship between the lending institution and the local policymaker generally has a repeated nature, as multiple projects are funded over time or in stages. A project may be anything from a building rehabilitation to a broad program aimed at achieving a development objective. The international lender usually audits the project at a pre-specified time during its implementation and, if the audit results are negative, i.e., the signal is bad, then funding for that particular project may be stopped, i.e., an intervention may occur. Finally, the lender can decide to stop funding all future projects ran by that policymaker, by cutting aid to programs ran by that mayor or agency and instead redirecting funds to projects in other jurisdictions.[2]

In the model, 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 a bad signal. Yet, both the effort exerted and the agent's cost of 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. The type captures, for instance, the ability of a policymaker to successfully adapt a project to local conditions. Costly effort 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. 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

[2] Other examples of such institutional relationships include the Structural and Investment Funds run by the European Commission, or transfers from a central government to local governments for running specific programs (for instance, Rwanda's Imihigo program).

Electronic copy available at: https://ssrn.com/abstract=3243123

at the end of the period, the principal updates her belief about the agent's ability.

The model repeats the above sequence of actions each period over an infinite horizon. To deliver results relevant for the motivating examples that involve political actors, we focus on the case in which the principal cannot monetarily compensate the agent.[3] Her only tools for incentives provision are the decision whether or not to intervene after a bad signal and the decision to continue or terminate future funding to the current agent.

We characterize the best Perfect Bayesian Equilibrium for the principal. When taking her actions, the principal faces two unknowns: the agent's type (a selection problem) and his effort (a control problem). To show the role played by each of these two factors in the results, we start with two benchmarks. In the first benchmark, we shut down the control problem by assuming that the high ability agent always exerts the highest effort. We thus obtain a bandit problem for the principal, which delivers a sharp characterization of the optimal intervention policy: it has a cutoff structure. The expected benefit of intervention linearly decreases in the principal's belief about the agent's type. The drawback of intervention is that it reduces the principal's learning about the agent's type. This cost of learning is a single peaked function of the principal's belief about the agent's type. It 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.

In the second benchmark, we shut down the selection problem. We assume that the principal faces an agent who chooses between exerting high effort or no effort. The principal motivates high effort by promising to continue to fund the agent in the future if current outcomes indicate that effort was likely exerted, i.e., if the signal or the project outcome are more likely with high effort. Each additional period of high effort requires a promise of future relationship continuation. Eventually, this leads to the typical result in dynamic

[3] This 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 an effective means of incentive provision.

Electronic copy available at: https://ssrn.com/abstract=3243123

moral hazard problems: it becomes unprofitable to further incentivize the agent to exert effort, and fulfilling the past promises requires continuing the relationship even if the agent stops exerting effort. This dynamic guarantees that there always exists a best equilibrium for the principal which has a cutoff structure: the agent exerts effort and the principal does not intervene until the agent's promised reward reaches a threshold. Afterwards, the agent stops exerting effort, and the principal intervenes after bad signals.

The threshold structures obtained in our two benchmarks are upended in the full model. To motivate the agent to exert effort, the principal must use the promise of higher rewards for 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 an inept 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 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

Electronic copy available at: https://ssrn.com/abstract=3243123

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 threshold rules no longer implement the optimal intervention policy.

Our results lead to several implications for the design of intervention policies. Consider, for instance, the application to international lending agreements. For some projects, control concerns are not first-order, either because effort is observable or the cost of effort is negligible. In such cases, a concern for the international organization may be the policymaker's ability 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 audits, for example, infrastructure projects. When the local policymaker's reputation increases, international lenders become more likely to fund projects which are more difficult to intervene in, for example, broad educational or health programs.[4] As predicted by the model, the latter type of projects are likely to be funded when intervention is not optimal. Another class of projects involve high electoral or personal costs for the policymaker, for instance, projects that require public employee reallocations. Here, the predicted slowdown in implementation effort over time implies an increasing pattern of interventions.

Lastly, many projects raise both selection and control concerns. This category usually includes reforms to the public sector, or the adoption and adaptation of new technologies to local conditions. These projects usually require ongoing auditing and potential intervention, regardless of the policymaker's reputation. Implementing the optimal policy involves costly and complex institutional structures. 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 pursue their preferred projects (or maintain the status quo) leading to a slowdown over time in reforms.[5]

We study dynamic relational contracting under both moral hazard and adverse selection, and our main innovation is to introduce the principal's ability to intervene to stop the agent's

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

[5] This dynamic provides another rationalization of the phenomenon described as "reform fatigue" (Bowen et al., 2016), and it provides an mechanism through which this fatigue may be observed specifically for reforms demanded by supranational agreements.

Electronic copy available at: https://ssrn.com/abstract=3243123

project. A vast literature has intensively analyzed contracting given adverse selection and screening (Rogoff, 1990; Besley and Case, 1995; Coate and Morris, 1995; Besley and Smart, 2007), moral hazard and incentives provision (Barro, 1973; Ferejohn, 1986; Austen-Smith and Banks, 1989; Ray, 2002; Acemoglu et al., 2008; Ales et al., 2014), and both moral hazard and adverse selection (Strulovici, 2011; Halac et al., 2016; Anesi and Buisseret, 2019; Banks and Sundaram, 1993, 1998). However, to the best of our knowledge, this is the first paper to analyze the dynamics of intervention. Levitt and Snyder (1997) have examined a similar type of intervention, but in a static model. Our focus is the dynamic structure of the intervention decisions.

Our focus on interventions and their effect on learning 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) and on dynamic reputation,[6] and specifically the application to the reputation of governments (Herrera et al., 2018). We follow a standard approach in this literature and differentiate between an inept agent type and a strategic agent type.

Finally, the result of switches between intervention and no intervention after bad news on the equilibrium path links our work to the literature on policy cycles (Ales et al., 2014; Dovis et al., 2016). Our result 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 exert effort.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 analyzes two natural benchmarks, a selection problem only and a control problem only. Section 4 analyzes the full model and provides comparative statics. Section 5 concludes, and the Appendix contains the proofs.[7]

[6] Starting with Kreps and Wilson (1982); Mailath and Samuelson (2001, 2006).

[7] The Online Appendix provides a three-period version of the model used to derive analytical results for the comparative statics.

Electronic copy available at: https://ssrn.com/abstract=3243123

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 **G**ood or **B**ad, $y \in \{G, B\}$. 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 $c(0) = 0$, $c'(0) = 0$, $\lim_{e \to 1} c'(e) = \infty \ \forall \theta$,[8] and $c'(e) > 0$, $c''(e) > 0 \ \forall e > 0$.

We assume that the L-type agent 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.[9]

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 probability 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.

[8] We 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.

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

Electronic copy available at: https://ssrn.com/abstract=3243123

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.[10] The monotonicity 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. Finally, the concavity/convexity assumptions, together with the convexity of $c(\cdot)$, guarantee that we can use a first-order condition to characterize the effort provision of the agent.

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. 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, and its outcome $y \in \{G, B\}$ is observed by everyone. If the outcome is B, the principal pays a cost C.

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 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 an H-type 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 is replaced by another minister; if a supranational institutions stops working with a policymaker, it may go on to sponsor projects in a different jurisdiction, under the purview

[10] This 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.

Electronic copy available at: https://ssrn.com/abstract=3243123

of a different policymaker. In some situations, however, an exogenous outside option may be a more natural continuation. For instance, the supranational institution may choose to keep its funds in a reserve fund instead of lending them out. As our analysis will make clear, assuming an exogenous outside option for the principal does not change our qualitative results. We therefore present 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:

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

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.

Given a belief μ that the agent is an H-type, 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:

$$u^P\left(\iota|\mu, e, s\right) = \iota(s) \cdot (-l) + (1 - \iota(s)) \cdot \Pr\left(o = B|\mu, e, s\right) \cdot (-C)\,. \tag{2}$$

By Bayes' Rule,[11]

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

Each period, the H-type agent derives a fixed rent based on whether he is kept ($\rho = 1$) or not ($\rho = 0$), and he a cost for the effort exerted:

$$u(e) = \rho \cdot \left[1 - \gamma \cdot c(e)\right]. \tag{3}$$

[11] 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).

Electronic copy available at: https://ssrn.com/abstract=3243123

The term $\gamma \geq 0$ parametrizes the size of the control problem faced by the principal. This payoff form corresponds to a purely office-motivated policymaker, who obtains a fixed benefit from being in office (or in the agreement with the international organization).[12] 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 $s = g$,*
$$\frac{\Pr\left((g,B)\,|e\right)}{\Pr\left((g,G)\,|e\right)+\Pr\left((g,B)\,|e\right)} < \frac{l}{C},\ \forall\ e \in [0,1]\,. \tag{4}$$

2. *With $e = 1$,*
$$\frac{\Pr\left((b,B)\,|1\right)}{\Pr\left((b,G)\,|1\right)+\Pr\left((b,B)\,|1\right)} < \frac{l}{C}. \tag{5}$$

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), $\iota(s = b) = 0$ if the principal knows that the agent is an H-type who exerts the highest level of effort.

2.1 Equilibrium Concept

We derive the best Perfect Bayesian Equilibrium for the principal. To construct an equilibrium in this class, we allow P to draw a public randomization device, using the technique also employed in Ales et al. (2014). This randomization device is used to determine the continuation play. 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 chooses whether

[12] We 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.

Electronic copy available at: https://ssrn.com/abstract=3243123

to continue the relationship with the current agent, $\rho_t \in \{0,1\}$, and the intervention rule $\iota_t(s_t)$ for each s_t. The agent chooses effort $e_t \in [0,1]$ conditional on $\rho_t = 1$. After observing the signal s_t, the principal follows the rule $\iota_t(s_t)$ to decide whether to intervene or not.

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 history is the same as the public history h_t. By contrast, the agent's history includes the agent type and effort levels in each past period: $h_t^A = (\theta_0, (z_j, \rho_j, \theta_j, e_j, \iota_j, o_j)_{j=1}^{t-1})$, where θ_0 is the type of the original agent.

The principal's strategy, σ_P, consists of a mapping from (h_t, z_t) to (ρ_t, ι_t).[13] The agent's strategy, σ_A, is a mapping from $(h_t^A, z_t, \rho_t, \iota_t, \theta_t)$ to $\alpha_t \in \Delta([0,1])$, where α_t is the agent's mixed strategy over effort in period t. Given σ_P, let $\Sigma_A^*(\sigma_P)$ be the set of all strategies σ_A such that, after each $(h_t^A, z_t, \rho_t, \iota_t, \theta_t)$, if the current agent is of type H ($\theta_t = H$), then σ_A maximizes

$$\max_{\sigma'_A} \mathbb{E}\left[u(e_t) + \sum_{\tau=t+1}^{\infty} \left(\prod_{j=t+1}^{\tau} \rho_j \right) \cdot \delta^{\tau-t} \cdot u(e_\tau) \middle| \sigma_P, \sigma'_A, h_t^A, z_t, \rho_t, \iota_t, \theta_t \right], \tag{6}$$

and if the current agent is of type L ($\theta_t = L$), then $e_t = 0$ with probability one.

Definition 1 *A Perfect Bayesian Equilibrium (PBE) consists of strategies σ_P for the principal and σ_A for the agent such that (i) the principal maximizes her expected continuation payoff after each (h^t, z_t), (ii) $\sigma_A \in \Sigma_A^*(\sigma_P)$, and (iii) the principal's belief is calculated using Bayes' Rule wherever possible.*

Our goal is to characterize the best PBE for the principal. To this end, we consider a simpler problem such that principal can first commit to σ_P and then the agent best responds to it. Moreover, we define $\hat{\Sigma}_A$ as the set of all pure strategies of the agent that depend only on the public history $(h_t, z_t, \rho_t, \iota_t)$ and the current type θ_t. We refer to these as *pure and semi-public* strategies, and we assume that the agent's chosen strategy is in this

[13] The principal and the agent take actions sequentially, and the principal's actions are observable to the agent, so it is without loss of generality to assume that the principal takes pure strategies.

Electronic copy available at: https://ssrn.com/abstract=3243123

set. Mathematically, we solve

$$\max_{\sigma_P} \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^{t-1} u | \sigma_P, \sigma_A\right] \tag{7}$$

subject to $\sigma_A \in \Sigma_A^*(\sigma_P) \cap \hat{\Sigma}_A$.

The principal's belief μ and her expected utility (7) are well-defined given our model's assumptions. In particular, Assumption 1 implies there is full support over outcomes o. Thus, the belief about h_t^A is well-defined by h_t, σ_A, and Bayes' Rule for each h_t and σ_A, with the following exception: $\iota_t(b) = 0$, the principal believes that either (i) $\theta_t = L$ with probability one or (ii) σ_A assigns probability one to $e_t = 0$, and the principal observes $(s_t, y_t) = (b, G)$. Assumption 2 implies that the principal sets $\iota_t(b) = 1$ if (i) or (ii) is correct. Hence, we can assume that $\Pr(h_t^A | h_t, z_t, \sigma_P, \sigma_A)$ is always well-defined.

The next lemma guarantees that the solution to this commitment problem in fact characterizes the best PBE for the principal.

Lemma 1 *The solution to (7) constitutes the best PBE for the principal.*

The proof consists of the four parts. First, given the principal's commitment, it is without loss of generality to focus on the pure and semi-public strategies of the agent since the past effort levels or agent types do not directly affect the payoff of the principal or that of the agent. Hence, (7) is in fact a relaxed problem. Second, in the solution of (7), the principal's continuation payoff never falls below a minimum value, $\underline{v}_P$. This minimum value is obtained when the agent exerts no effort, the principal always intervenes, and she continues the contract with the agent. Notice that under this play, the principal can give the highest feasible continuation payoff to the agent. Hence, even when the principal wants to reward the agent after good outcomes in order to incentivize high effort, she can guarantee that her payoff is no less than $\underline{v}_P$. Third, we construct a punishment PBE where the principal's payoff after any deviation is no more than $\underline{v}_P$. Intuitively, in this equilibrium, the principal always replaces the agent and the agent does not exert effort. Finally, we define another

Electronic copy available at: https://ssrn.com/abstract=3243123

PBE as follows: on path, the principal follows the solution to (7). As soon as she deviates, we switches to the punishment PBE. This PBE implements the solution to (7) on path.

Notice that, from Assumption 2, no intervention is statically optimal after $s = g$, regardless of σ_A. Moreover, no intervention allows P to have access to more information in the Blackwell sense, as she is able to observe the project's outcome, not just the signal. This implies that no intervention is also dynamically optimal after $s = g$, so $\iota_z(s = g) = 0$. For the rest of the paper, we focus on the intervention decision after $s = b$, and we let $\iota_z \in \{0, 1\}$ denote $\iota_z(s = b) \in \{0, 1\}$ for notational simplicity.

In the next section, we describe two simple benchmark cases. Afterwards, we solve for the best PBE for the principal in the full model.

3 Simplified Settings

3.1 Adverse Selection Only

In the first benchmark, we shut down the agent's effort choice and assume instead that effort is fixed for each agent type: the H-type exerts effort $e = 1$ at no cost ($\gamma = 0$), and the L-type exerts effort $e = 0$. The benchmark is therefore a selection problem for the principal. The difference between agents reduces to the ability to undertake the project.

The game becomes a bandit problem of a single decision-maker, P, since A does not take any meaningful action.[14] Let $\mu \in [0, 1]$ be the probability that A 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 P's current belief. If P removes the current agent, the game restarts with a new agent who is expected to be an H-type with probability μ_H. Hence, P's utility after replacing A, denoted by $\bar{J}$, should satisfy $\bar{J} = J(\mu_H)$.[15]

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

[14] As such, without loss of generality, we can ignore the public randomization device z.

[15] If the outside option were exogenous, it would be represented by a fixed value $\overline{J}$.

Electronic copy available at: https://ssrn.com/abstract=3243123

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

$$J(\mu) = \max_{\substack{\rho \in \{0,1\}, \\ (\iota(s))_s \in \{0,1\}^2}} (1-\rho) \cdot \bar{J} + \rho \cdot \left[\sum_{s \in \{g,b\}} \Pr(s|\mu) \cdot u^P(\iota|\mu, s) + \delta \cdot \sum_{\omega} \Pr(\omega|\mu, \iota) \cdot J(\mu'(\mu, \omega)) \right], \tag{8}$$

where $u^P(\iota|\mu, s)$ is given in (2), but here we omit e since it always equals 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 2 *$J(\mu)$ 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. To see why $J(\mu)$ is convex in μ, suppose that P has an option to observe additional information about A's type. Her expected payoff then weakly increases when this option becomes available (if using this information is not increasing her expected payoff, she can simply ignore it). After observing the new information, the distribution of updated beliefs μ' is a mean-preserving spread of the original belief μ since the belief process is a martingale. Hence, the mean-preserving spread of μ (weakly) increases $J(\mu)$, which means that $J(\mu)$ is convex.

Given the linearity of the value function $J(\mu)$ with respect to ρ, Lemma 2 has an immediate implication for the optimal replacement strategy:

Lemma 3 *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, P continues with A as long as the expected type of a replacement agent is lower than the expected type of the current agent.

We further build our analysis by considering P's optimal strategy in a static game:

Electronic copy available at: https://ssrn.com/abstract=3243123

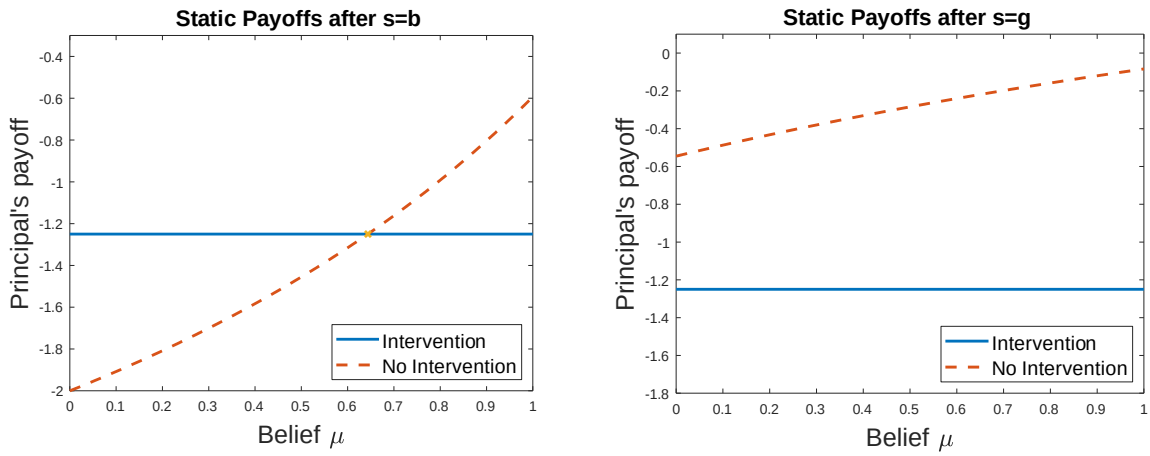

Figure 1: Pricipal's Payoffs when δ=0.

Lemma 4 *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 $\mu^S < 1$.*

The result follows from Assumption 2. If the signal is bad ($s = b$), then the project is expected to succeed only if there is a sufficiently high probability that A is an H-type. Otherwise, intervention comes at a lower cost than the expected loss from the project's outcome. Figure 1 illustrates the P's expected payoffs in the static problem.

Next, we use the above results to derive the optimal policy 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 4 because P also takes into account the effect of her current intervention decision on her future belief about A's type. Without intervention, the belief update is based on both the signal and the project outcome, while with intervention, it based only on the signal. Thus, not intervening in the current period has a dynamic benefit of better learning A's type. If this benefit is sufficiently high, not intervening is optimal even though it does not maximize instantaneous utility.

Electronic copy available at: https://ssrn.com/abstract=3243123

When $\mu \geq \mu^S$, 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. When $\mu < \mu^S$, however, the static prescription is intervention. This precludes the project's outcome from being revealed. If P chooses no intervention instead, and the project's outcome is highly informative about A'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 $s = b$ 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, an observation of (b, B) updates the belief to $\mu < \mu_H$, 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 very likely an H-type or to removing the agent. This is true for all $\mu \in [\mu^H, \mu^S)$. 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.

To sum up, the optimal intervention policy has a stark characterization: intervention after bad news is optimal up until P is sufficiently confident that A is an H-type.

3.2 Moral Hazard Only

In this second benchmark, we shut down the selection aspect of the model. We assume that the principal faces only the H-type agent, and $\gamma > 0$ in the agent's utility function. The agent may choose effort $e \in \{0, 1\}$, where $c(1) = c > 0$. Notice that now A takes a strategic action by choosing effort each period. The principal must therefore account for A's payoff from any proposed play, and for his incentives given the proposed play. To capture this, we denote by $J(V)$ the principal's value function given the value V promised to the current agent. As before, we denote by $\bar{J} \equiv \max_V J(V)$ the highest expected value for P when the game restarts with a new agent. We obtain this value by selecting the ex-ante best equilibrium for P. This is possible, since at the start of the contract P has the freedom to choose a promised value V which will maximize her expected payoff.

Electronic copy available at: https://ssrn.com/abstract=3243123

Given V and every possible realization of z, the principal's recursive problem resumes to choosing $\rho_z \in [0,1]$, $\iota_z \in \{0,1\}$, the recommended effort $e_z \in \{0,1\}$, and continuation promised values $V_z'(\omega)$ after each publicly observable event $\omega = (s,o)$ to maximize

$$\begin{aligned} J(V) \quad = \quad & \max_{\rho_z,\iota_z,e_z,V_z'(\omega)} \int_z \Bigg\{ (1-\rho_z)\cdot \bar{J} + \rho_z \cdot \Bigg[\sum_s \Pr(s|e_z)\cdot u^P(\iota_z|\mu,e_z,s) \\ & +\delta \cdot \textstyle\sum_\omega \Pr(\omega|e_z,\iota_z)\cdot J(V_z'(\omega))\Bigg]\Bigg\}\, dz, \end{aligned} \tag{9}$$

subject to the constraints

$$V \quad = \quad \int_z \rho_z \cdot \{1 - c\cdot e_z + \delta \cdot \textstyle\sum_\omega \Pr(\omega|e_z,\iota_z)\cdot V_z'(\omega)\}\, dz; \tag{10}$$

$$e_z \quad \in \quad \arg\max_{e\in\{0,1\}} \{1 - c\cdot e + \delta \cdot \textstyle\sum_\omega \Pr(\omega|e,\iota_z)\cdot V_z'(\omega)\}; \tag{11}$$

$$V_z'(\omega) \in \left[0, \frac{1}{1-\delta}\right] \text{ for each } \omega. \tag{12}$$

Constraint (10) is the promised value which P is bound to provide to A in equilibrium. Constraint (11) is the incentive compatibility constraint for A. Finally, constraint (12) places the upper and lower bounds on the promised continuation value. The minimum payoff the agent may receive is 0, and the highest feasible utility for the agent is $(1-\delta)^{-1}$. The latter is implemented by keeping the agent and allowing him to exert effort $e = 0$ forever.

Analogous to the previous section, we first establish the basic properties of $J(V)$.

Lemma 5 *$J(V)$ is concave and decreasing in V.*

The concavity of J with respect to the agent's promised value V follows from the standards arguments, given the use to the randomization device z. The payoff J is decreasing in V because a higher promised value reduces the agent's incentive to exert effort in future periods. This happens because V can only be provided by not removing the agent in future periods.

Lemma 6 *If effort is exerted $(e_z = 1)$, the optimal strategy for the principal is no intervention ($\iota_z = 0$), and the incentive compatibility constraint (11) is binding. If no effort is exerted*

Electronic copy available at: https://ssrn.com/abstract=3243123

$(e_z = 0)$, *the optimal strategy for the principal is intervention* ($\iota_z = 1$), *and* $V_z(\omega) = V_z(\omega')$ *for each* ω, ω'.

By Assumption 2, no intervention is statically optimal after $e = 1$. To see that it is also dynamically optimal, notice that any promised continuation payoffs offered after an intervention can be replicated under no intervention: P can simply ignore the outcome observed after no intervention. Thus, no intervention is also dynamically optimal when the agent exerts effort. Constraint (11) binding implies that P offers the minimal reward that will induce $e = 1$. This happens because P's payoff is decreasing in the agent's reward.

When $e = 0$, there is no gain for P in offering a reward for better outcomes, as these are obtained purely by chance, without the agent's effort contribution. As the agent's continuation payoff does not depend on the outcome, the principal can simply pick the statically optimal action, which is intervention.

Given these insights, we can establish the existence of an equilibrium that has a cutoff structure for the intervention policy.

Proposition 2 *There exists a best PBE for the principal where after any history, if the agent has taken* $e = 0$, *then he will take* $e = 0$ *with probability one. The optimal strategy for the principal is to not intervene as long as* $e = 1$ *and to intervene once* $e = 0$.

Proposition 2 shows that the highest payoff for P can be achieved when the intervention policy follows a threshold rule. The principal does not intervene and the agent exerts effort up until a threshold promised reward. After that threshold, the principal intervenes following bad signals, and the agent does not exert effort. This equilibrium delivers the highest payoff for P because it allows P to frontload all the effort that she can incentivize the agent to exert. In doing so, the principal benefits from higher likelihood of early project successes. These are more valuable than later successes, because of discounting.

The above result makes it clear that the highest payoff for the principal can *always* be achieved through an equilibrium play in which the intervention policy takes a threshold form. As in the first benchmark, this points to a simple institutional implementation of the optimal contract: a regime change happens once the promised value threshold is reached.

Electronic copy available at: https://ssrn.com/abstract=3243123

4 Intervention in the Full Model

We now analyze the optimal intervention policy for the principal in the full model. We present the most general case, in which $e \in [0,1]$, and once an agent is removed, the principal starts a new relationship with randomly drawn new agent.[16]

We solve problem (7) by a recursive method. We denote by $J(\mu, V)$ the principal's value function given the two state variables, the belief μ about the agent's type and the value V promised to the current agent. If the agent is replaced ($\rho = 0$), 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 expected payoff after replacing A satisfies $\bar{J} = \max_V J(\mu_H, V)$.

Given (μ, 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 that agent, given the publicly observable signal-outcome pair $\omega = (s, o)$. 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.

The principal chooses α_z to solve the following dynamic program, for each $\mu \in [0,1]$ and $V \in \left[0, \frac{1}{1-\delta}\right]$:

$$\begin{aligned} J(\mu, V) \;=\; \max_{\alpha_z} \int_z \Bigg\{ (1-\rho_z)\cdot \bar{J} + \rho_z \Bigg[\sum_s \Pr(s|\mu, e_z)\cdot u^P(\iota_z|\mu, e_z, s) \\ + \delta \sum_\omega \Pr(\omega|\mu, e_z, \iota_z)\cdot J\left(\mu'(\mu, e_z, \omega), V_z'(\omega)\right) \Bigg] \Bigg\} dz, \end{aligned} \tag{13}$$

[16] The same qualitative results may be obtained with just binary effort, as in the benchmarks, or if the principal takes an outside option $\bar{J}$ instead of drawing a new agent. Details available upon request.

Electronic copy available at: https://ssrn.com/abstract=3243123

subject to the constraints

$$V = \int_z \rho_z \left\{ 1 - c(e_z) + \delta \sum_\omega \Pr(\omega | e_z, \iota_z) V_z'(\omega) \right\} dz; \tag{14}$$

$$e_z \in \arg\max \left\{ 1 - c(e_z) + \delta \sum_\omega \Pr(\omega | e_z, \iota_z) V_z'(\omega) \right\}; \tag{15}$$

$$V_z'(\omega) \in \left[0, \frac{1}{1-\delta} \right] \text{ for each } \omega. \tag{16}$$

Notice that the principal faces the same constraints as (10)-(12), because only the H-type agent can be incentivized to exert effort. Moreover, in the constraints, we have $\Pr(\omega | e_z)$, because A knows his own type. The principal's objective differs, however, from (9), because P is uncertain about the agent's type. Thus, from P's view point, the distribution of the signal-outcome pair $\Pr(\omega_z | \mu, e_z)$ now also depends on μ.

4.1 Equilibrium Properties

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

Lemma 7 *$J(\mu, V)$ is concave in V, convex in μ, and increasing in μ, with this increase strict if $V \in \left(0, \frac{1}{1-\delta}\right)$.*

These properties were shown, separately, in Lemmas 2 and 5, and the exercise remains the same when selection and incentives provision are combined.

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

Lemma 8 *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 $V(\mu) \in \left[0, \frac{1}{1-\delta}\right]$ 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, $\frac{d}{dV} J(\mu, V)\big|_{V \in [0, V(\mu)]}$, is negative for $\mu < \mu_H$, it is zero for $\mu = \mu_H$, and it is positive for $\mu > \mu_H$.*

Electronic copy available at: https://ssrn.com/abstract=3243123

3. *There exists* $V^*(\mu) \in \arg\max_{V \in \left[0, \frac{1}{1-\delta}\right]} J(\mu, V)$ *such that* $\forall \mu \in (0,1]$, *at* $V = V^*(\mu)$, $V_z'(\omega) \leq (\geq) \arg\max_V J(\mu'(\mu, e_z, \omega), V)$ *after each event* ω *with negative (positive) belief update* $\mu'(\mu, e_z, \omega) \leq (\geq) \mu$.

First, if the value promised to A is 0, only immediate replacement can fulfill this promise. In that case, P starts a new contract with a new agent. 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. For instance, if $V < 1$, the agent is promised less that what he could obtain at the end of a period, without exerting any effort. Thus, for the promise to be kept, he must face some probability of replacement. 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. Promising a smaller V requires setting a higher probability of replacing A. Notice that if the current belief is lower (higher) than μ_H, then setting a smaller V is good (bad) for P. She replaces the agent with higher probability, and a new agent is an H-type with probability μ_H. Hence, the slope of the J function with respect to V, $\frac{d}{dV} J(\mu, V)\big|_{V \in [0, V(\mu)]}$, is negative for $\mu < \mu_H$ and positive for $\mu > \mu_H$. These properties and those described in Lemma 7 are illustrated in Figure 2.

The third property follows from the concavity of J with respect to V. For each $\mu \in [0,1]$, the function $J(\mu, V)$ is maximized at some $V^*(\mu)$. Consider the case when $V = V^*(\mu)$, and P updates her belief to μ' after observing a signal-outcome pair ω. A negative (positive) belief update, i.e., $\mu' < (>) \mu$, implies that ω is the outcome which happens less (more) often with higher effort. Hence, reducing (increasing) $V_z'(\omega)$ incentivizes A to exert more effort. This increased effort has two benefits. First, the instantaneous expected utility is improved. Second, P's belief update is larger, since higher effort makes the H-type more distinguishable from the L-type. If the continuation payoff $V_z'(\omega)$ is higher (lower) than the value that maximizes P's utility, then reducing (increasing) $V_z'(\omega)$ directly improves P's payoff. In total, if ω is the event which happens less (more) often with higher effort, then reducing $V_z'(\omega)$ improves effort provision, learning, and P's continuation payoff.

Electronic copy available at: https://ssrn.com/abstract=3243123

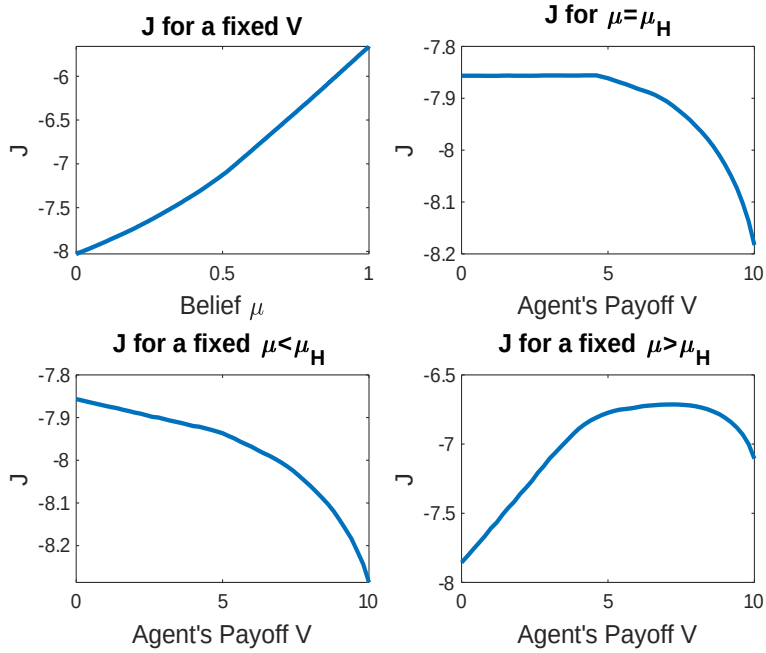

Figure 2: Shape of the Principal's Value Function. For the upper left panel, $V = \frac{0.25}{1-\delta}$. For the upper right panel, $\mu_H = 0.12$, for the lower left panel, $\mu = 0.07$, and the lower right panel, $\mu = 0.6$. We assume quadratic effort cost.

4.2 The Optimal Dynamic Policy

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, not intervening allows the principal to update her belief about the agent more precisely than if she intervened:

Lemma 9 (Learning) *Fixing any targeted effort level e_z, when P 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.*

Electronic copy available at: https://ssrn.com/abstract=3243123

By Lemma 7, J is convex in μ. Hence, the increased variance of the distribution $(\mu'(\mu, e_z, \omega))_\omega$ under no intervention implies that no intervention benefits P by increasing the continuation payoff through better learning. Moreover, this learning effect is stronger if the current belief is intermediate, because in that case there is more room for the belief to move following an update.

Amplification of the learning benefit.

An implication of the above results is that the benefit of learning is higher when higher effort is incentivized. Not intervening allows P to gather more information, and she optimally uses this information to 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 the zero effort exerted by the L-type. Thus, the principal learns the agent's type more accurately.

Switches between intervention and no intervention.

Not intervening after a bad signal offers P dynamic benefits, in terms of better learning and incentives provision; however, this comes with a static cost — the expected loss due to the higher probability of a bad outcome that period. In what follows, we show how this trade-off between the dynamic benefits and the static cost leads to switches between intervention and no intervention on the equilibrium path. We formally show this result in the case in which P observes a sequence of events which positively update her belief and her promised value for A. Afterwards, we discuss other possible sequences of events.

We begin by giving the sufficient conditions for the 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:*

Electronic copy available at: https://ssrn.com/abstract=3243123

(i) $\frac{d}{de}\left[\Pr\left(g,G|\bar{e}\right)\cdot\bar{q}-c\left(\bar{e}\right)\right]\big|_{e=\bar{e}}\geq 0$;

(ii) $\Pr\left(g,G|\bar{e}\right)\geq 1-\bar{q}$;

(iii) $\frac{\Pr(b,B|\bar{e})}{\Pr(b|\bar{e})}<\frac{l}{C}$.

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 P could observe the agent's type, she would prefer not intervene given an H-type. This agent type exerts high effort, which makes a bad outcome sufficiently unlikely.

Given this definition, we establish that a simple threshold strategy as in Propositions 1 and 2 no longer holds in the full model, and in fact we obtain switches between intervention and no intervention on the equilibrium path:

Proposition 3 *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* $\Pr\left(G|0\right)$, *such that if* $\mu_H\leq\bar{\mu}_H$, $\Pr\left(G|0\right)\leq\bar{q}(\mu_H)$, *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 3 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 and increase in V 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

Electronic copy available at: https://ssrn.com/abstract=3243123

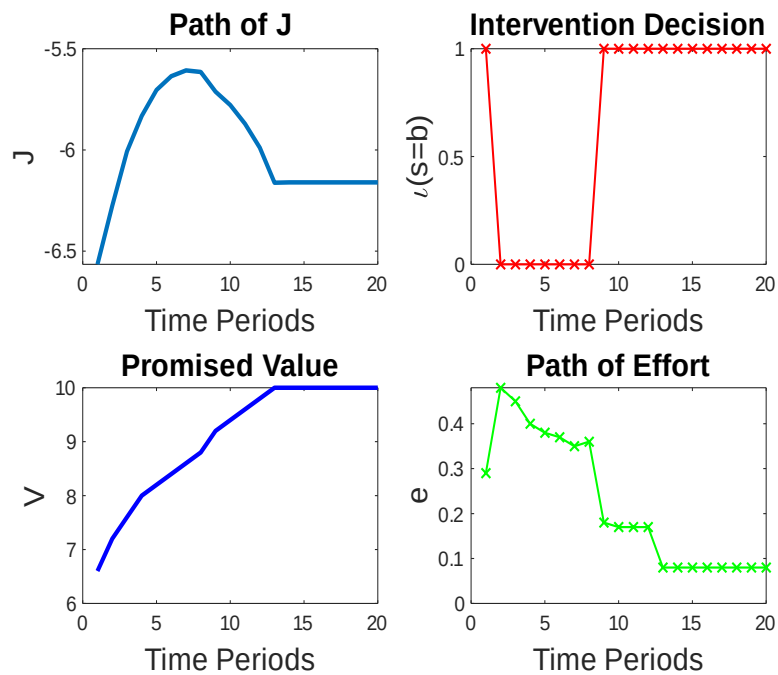

Figure 3: Dynamics after the signal-outcome realization (g,G) over 20 periods.

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 he must therefore intervene because the agent exerts too little effort. Figures 3 illustrates this dynamic in a numerical example of the model.

The mechanism behind our result consists of three elements that form the principal's trade-off when deciding whether to intervene: (i) the static cost of no intervention, (ii) the dynamic benefit of learning, and (iii) the dynamic benefit of incentives provision. Contrasting

Electronic copy available at: https://ssrn.com/abstract=3243123

our result here with our two benchmarks can help illuminate why all three elements together result in switches between intervention and no intervention on the equilibrium path.

Our first benchmark, of adverse selection only, shows that the trade-off between (i) and (ii) alone does not result in switches between intervention and no intervention on the equilibrium path. The static (expected) cost of intervention decreases as the belief μ increases. Even though the dynamic benefit of learning is highest for intermediate values of μ, the overall net effect of these two elements is a decreasing benefit of intervention as the belief increases.

Our second benchmark, of moral hazard only, shows that the trade-off between (i) and (iii) is not sufficient to guarantee that such switches are present in equilibrium. The static (expected) cost of intervention decreases in the agent's current effort, while the dynamic benefit of incentives provision increases in effort. This net effect, together with the principal's gain from frontloading good outcomes (due to discounting), results in an increasing benefit of intervention on the equilibrium path. Note that, even though we assume $e \in [0, 1]$, Proposition 3 is readily extended, if we replace $\bar{e}$ with $e = 1$ in Definition 2.

In the full model, it is the dynamic benefit of learning together with the dynamic benefit of incentives provision that breaks the monotonicity from our benchmark results. A switch from intervention to no intervention is optimal in period $t \geq 2$ in order to capture the benefit of learning, which is highest for intermediate beliefs. A switch from no intervention to intervention is optimal in the long run because the benefit of incentives provision decreases, due to the higher cost of rewarding effort.

Lastly, 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. We pick (g, G) sequence only to make the comparison with the benchmark cases clear. However, the result in Proposition 3 can be extended to obtain switches between intervention and no intervention under a variety of signal-outcome sequences. 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

Electronic copy available at: https://ssrn.com/abstract=3243123

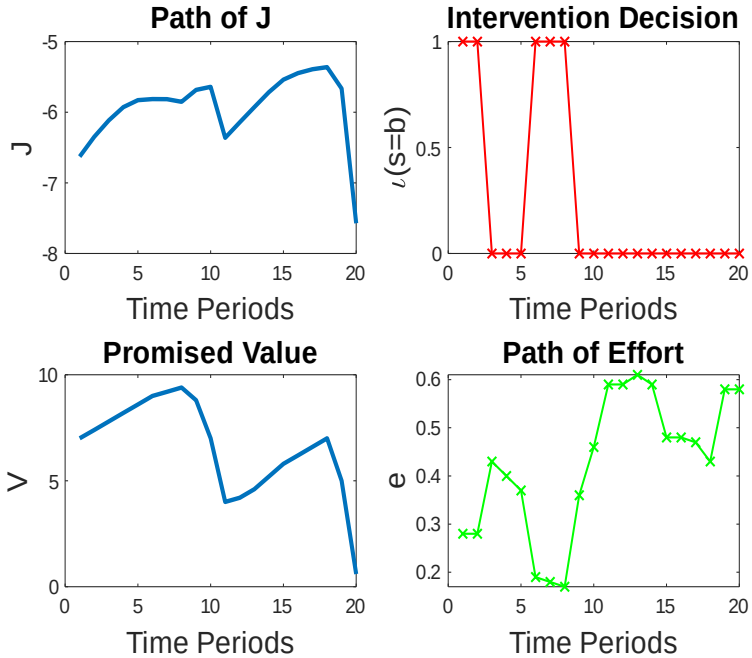

Figure 4: Dynamics after a sequence of alternating 8 periods with the signal-outcome realization (g,G) and 2 periods with the realization (g,B)

is sufficiently low or the agent's promised value is sufficiently high. Figure 4 illustrates this dynamic for such an example of outcome realizations.

Implementation of the optimal intervention policy.

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 (different) firm accused of financial fraud. The study finds that,

Electronic copy available at: https://ssrn.com/abstract=3243123

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.

4.3 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 $V_1 \in \left[1, 1 + \delta + \delta^2\right]$. 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 6, 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 analytically in the 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 4 (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.*

Electronic copy available at: https://ssrn.com/abstract=3243123

Higher values of δ increase the value of P's future payoffs, 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 means that the conflict between the principal and the agent is reduced. This has two implications. On the one hand, the lower cost of incentivizing effort in the current period makes more efficient monitoring less important. On the other hand, the lower cost of incentivizing effort in future periods increases continuation payoffs, which increases the value of providing rewards efficiently.

Notice that, while both δ and γ change the cost of providing incentives, they have different effects on the intervention policy. 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 (and will be funded) in future periods. The marginal cost of effort, γ, 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 5.

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

Proposition 5 (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*

Electronic copy available at: https://ssrn.com/abstract=3243123

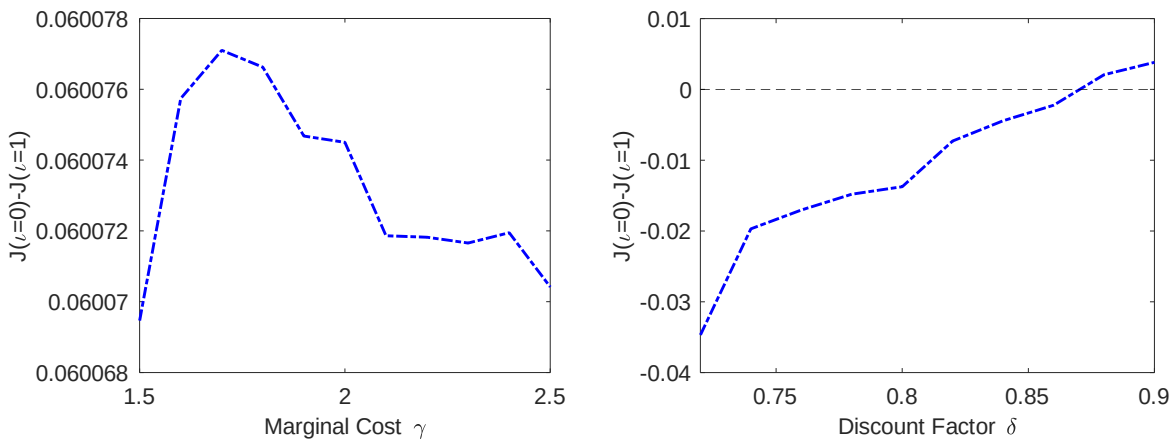

Figure 5: Relative Value of No Intervention as a function of γ (left panel) and as a function of δ (right panel). The grapth is shown for $V = ((0.5)/(1-\delta))$ and $\mu = 0.75$ and quadratic effort cost $\frac{\gamma}{2} \cdot e^2$.

$y \in \{G, B\}$, *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 dynamic cost of fraught audit or oversight systems: they reduce future payoffs by slowing down the learning process about the agent's ability and by 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 6 *The intervention decision is not monotone in the agent's promised payoff V (for $\mu < 1$). 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 3. A similar intuition can be applied when examining the separate roles of V and μ. 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

Electronic copy available at: https://ssrn.com/abstract=3243123

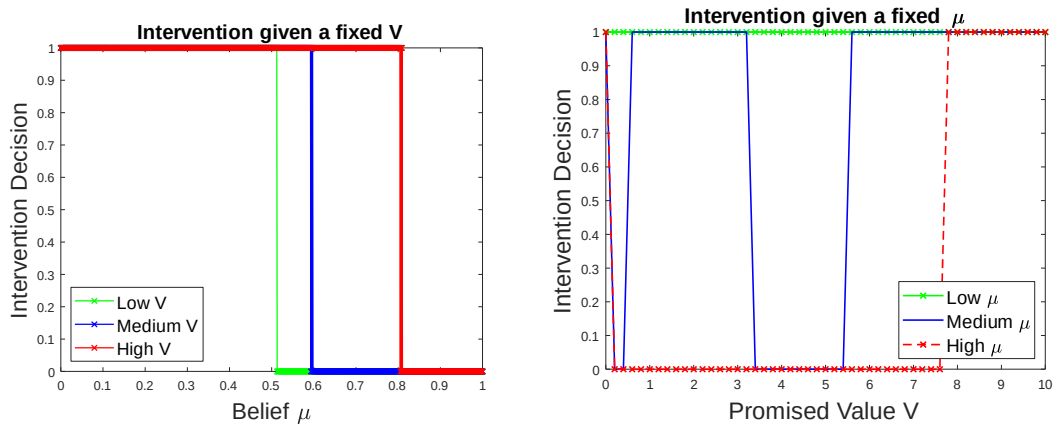

Figure 6: The right panel shows the intervention decision given $V = ((0.2)/(1-\delta)), ((0.3)/(1-\delta)), ((0.4)/(1-\delta))$. The right panel shows the intervention decision given $\mu = 0.3, 0.5, 0.8$

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 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. Notice that this result and its intuition rely on $\mu_H < 1$, thus the difference from our second benchmark.

The effect of increasing μ on the marginal benefit of not intervening may be either positive or negative. On the one hand, no intervention becomes more beneficial since we assume that, with $\mu = 1$, no intervention is optimal in the first period. 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 6 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).

Electronic copy available at: https://ssrn.com/abstract=3243123

5 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 or a control problem only, the optimal intervention policy takes a threshold form. If both selection and control are simultaneous concerns, the optimal policy exhibits switches between intervention and no intervention.

Our setup is relevant for a broad set of applications. In addition to the applications to international institutional contracts discussed in the introduction, we highlight two additional applications. First, another relevant political economy setting is that of two-party governing coalitions in parliamentary democracies. There, a government minister maps the agent in our model, while the parliamentary committee which oversees that ministerial jurisdiction maps to the principal. The parliamentary committee's membership reflects the composition of the ruling coalition and seeks to implement the projects agreed upon inside the coalition. The committee may schedule hearings, gather information — the signal in our model; it can intervene by proposing amendments to the legislation produced by the minister, by delaying or blocking legislation. 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. Martin and Vanberg (2004, 2005) provide empirical evidence on the patterns of parliamentary intervention. As predicted by our model, parliamentary intervention is observed more often on issues that are more "politically divisive" for the coalition. These cases can be mapped to carrying a higher marginal cost of effort for the minister delegated to tackle them.

Electronic copy available at: https://ssrn.com/abstract=3243123

Another relevant application is to contracts between investors and entrepreneurs. These relationships usually have a repeated nature, and they pose selection or control issues. Our results suggest that control concerns along with selection concerns lead to recurring interventions over time. Moreover, as the relationship duration increases, entrepreneurs are less likely to work on projects preferred by the investors.

The model considers the case in which the principal runs only one project each period. In this setting, we could provide a rigorous dynamic analysis of the intervention problem, on which more complexity can be built. A natural such direction is to consider a setting in which the agent can work on multiple projects. In particular, our results highlight the rich structure of the optimal policy that emerges due to the control problem in environments in which the agent cannot be offered a wage schedule. An environment with multiple projects could add an additional instrument for the principal, as she 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.

Anesi, V., Buisseret, P., 2019. Making elections work: Accountability with selection and control. Mimeo, University of Chicago.

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.

Electronic copy available at: https://ssrn.com/abstract=3243123

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.

Fich, E. M., Shivdasani, A., 2007. Financial fraud, director reputation, and shareholder wealth. Journal of Financial Economics 86 (2), 306–336.

Electronic copy available at: https://ssrn.com/abstract=3243123

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., 2018. Political booms, financial crises. Journal of Political Economy (forthcoming).

Kreps, D. M., Wilson, R., 1982. Reputation and imperfect information. Journal of Economic Theory 27 (2), 253–279.

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.

Electronic copy available at: https://ssrn.com/abstract=3243123

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.

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.

Electronic copy available at: https://ssrn.com/abstract=3243123

A Appendix

A.1 Proof of Lemma 1

Our objective is to maximize the objective given in (7) subject to $\sigma_A \in \Sigma_A^*(\sigma_p)$. Notice that, for each σ_P, if A is indifferent between two pure strategies (of the dynamic game), then we can arbitrarily pick the one better for P. In addition, the past effort levels or types do not directly affect P's or A's payoffs. Thus, for each σ_A which depends on the past types or efforts, there exists another σ'_A which (i) does not depend on those variables, (ii) solves (6), and (iii) brings P the same payoff. Hence, without loss of generality, we assume that the agent takes pure strategies that do not depend on past effort levels or types. Now, the relaxed problem becomes (7). Therefore, it remains to show that the solution to (7) constitutes a PBE.

Given the agent's problem (6), the effort level by the H-type agent in period t after (h_t, z_t, ρ_t) depends only on the continuation payoff after each possible realization of o_t conditional on $\theta_t = H$:

$$\left(w(\sigma_P, h_t, z_t, \rho_t, \iota_t, o_t)\right)_{o_t} \equiv \left(\max_{\sigma_A} \mathbb{E}\left[\sum_{\tau=t+1}^{\infty} \prod_{j=t+1}^{\tau} \rho_j \delta^{\tau-t} u(e_\tau) \,|\sigma_P, \sigma_A, h_t, z_t, \rho_t, \iota_t, \{\theta_t = H\}, o_t\right]\right)_{o_t}. \quad (17)$$

By feasibility, all of them are included in $\left[0, \frac{1}{1-\delta}\right]$.

For each $(h_t, z_t, \rho_t, \iota_t)$ and $(w_o)_o$ satisfying $w_o \in \left[0, \frac{1}{1-\delta}\right]$ for each o, there exists σ_P^* such that (i) σ_P^* guarantees that the agent's payoff equals $(w_o)_o$: $w(\sigma_P^*, h_t, z_t, \rho_t, \iota_t, o_t) = w_{o_t}$ after each o_t; and (ii) σ_P^* guarantees that the principal's continuation payoff given each h^{t+1} is no less than $\underline{v}_P$, where

$$\underline{v}_P \equiv \frac{1}{1-\delta}\left(-\Pr(b|0)\, l - \Pr((g,B)\,|0)\, L\right). \quad (18)$$

We construct such a σ_P^* as follows: Given w_o, the principal calculates the probability α_o such that $\alpha_o \frac{1}{1-\delta} = w_o$. Using the public randomization at the beginning of period $t+1$, the principal keeps the agent forever (regardless of the future outcomes) with probability α_o and replaces him with probability $1-\alpha_o$ after $o_t = o$. In the latter case, the principal replaces the future agents after one period regardless of the outcome. The principal always intervenes. Given such σ_P^*, the H-type agent does not provide any effort after each history.

Hence, the principal's continuation payoff in (7) is no less than $\underline{v}_P$ after each history h_t.[17] It remains to show that there exists a punishment equilibrium such that the principal's payoff given h_{t+1} is no more than $\underline{v}_P$.

Consider the following strategy profile: the principal replaces the agent and intervenes

[17] Otherwise, replace the principal's continuation strategy from h_t with σ_P^*; this change improves her continuation payoff from h_t without affecting the agent's incentive before period t.

Electronic copy available at: https://ssrn.com/abstract=3243123

after each history, and the H-type agent does not provide any effort after each history. Clearly this strategy profile is a mutual best response. Moreover, the principal's payoff is no more than $\underline{v}_P$ after each history h^t, (h^t, z_t), or (h^t, z_t, ρ_t); and no more than $\mathbb{E}[u|e=0, \iota_t] + \delta \underline{v}_P \leq \underline{v}_P$ after each history $(h^t, z_t, \rho_t, \iota_t)$, as desired.

A.2 Proof of Lemma 2

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

$$\Pr(s = g|\mu)\, u^P(0|\mu, e, s) + \Pr(s = b|\mu)\, u^P(\iota(s = b)|\mu, e, s), \tag{19}$$

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

Convexity with respect to μ.

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

$$J(\mu) = \mu J(\mu, H) + (1-\mu) J(\mu, L) = J(\mu, L) + \mu [J(\mu, H) - J(\mu, L)]. \tag{20}$$

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

$$J(\mu, L) + \mu_n [J(\mu, H) - J(\mu, L)] \leq J(\mu_n). \tag{21}$$

Hence,

$$\begin{aligned} & \beta J(\mu_1) + (1-\beta) J(\mu_2) \\ \geq\ & \beta J(\mu, L) + \beta\mu_1 [J(\mu, H) - J(\mu, L)] \\ & + (1-\beta) J(\mu, L) + (1-\beta)\mu_2 [J(\mu, H) - J(\mu, L)] \\ =\ & J(\mu, L) + \mu [J(\mu, H) - J(\mu, L)] = J(\mu). \end{aligned} \tag{22}$$

A.3 Proof of Lemma 4

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

$$\frac{l}{C} \leq \Pr(H|\mu, s = b) \cdot \Pr(B|H, b) + \Pr(L|\mu, s = b) \cdot \Pr(B|L, b), \tag{23}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

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

$$\mu \geq \mu_S := \frac{\Pr(b, B|L)\left(1 - \frac{l}{C}\right)}{\left(\Pr(b, B|L) - \Pr(b, B|H)\right)\left(1 - \frac{l}{C}\right) + \frac{l}{C}\Pr(b, G|H)}. \tag{24}$$

A.4 Proof of Proposition 1

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

$$\mu_b = \frac{\mu}{\mu + (1-\mu)\frac{\Pr(b,B|L)+\Pr(b,G|L)}{\Pr(b,B|H)+\Pr(b,G|H)}}. \tag{25}$$

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

$$\begin{aligned} &\mu_b u^H(\iota) + (1-\mu_b)\, u^L(\iota) \\ &+\delta\iota J(\mu_b) + \delta(1-\iota)\left(\alpha\mu_b J(1) + (1-\alpha\mu_b)\, J(\mu')\right), \end{aligned} \tag{26}$$

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

$$\mu' = \frac{\mu \Pr(b, B|H)}{\mu \Pr(b, B|H) + (1-\mu)\Pr(b, B|L)}. \tag{27}$$

Notice that when $\mu \geq \mu^S$, $\iota = 0$ is optimal by Lemma 4. When $\mu = \Delta$, where $\Delta \to 0$, we have $\mu' \to 0$. Thus, $\iota = 1$ is optimal for $\mu \to 0$. 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:

$$\frac{1-\mu^S}{\mu^S} \geq \frac{1-\mu_H}{\mu_H}\frac{\Pr(b, B|H)}{\Pr(b, B|L)}, \tag{28}$$

where μ^S is derived in (24). Thus, substituting for μ^S, the above condition reduces to:

$$\frac{\Pr(b, B|H)}{\Pr(b, G|H)} \leq \frac{\frac{l}{C}}{1 - \frac{l}{C}}\mu_H. \tag{29}$$

Therefore, with the upper bound $\frac{l}{C}\left(1 - \frac{l}{C}\right)^{-1}\mu_H$ 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 (26), the expression

$$\mu_b u^H(\iota) + (1-\mu_b)\, u^L(\iota) + \delta(1-\iota)\left(\alpha\mu_b J(1) + (1-\alpha\mu_b)\, J(\mu')\right) \tag{30}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

is 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.5 Proof of Lemma 5

We show that $J(V)$ is concave in V. 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 (V_1) and (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.

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 $J(V_1)$, and with probability $1-\beta$, she achieves $J(V_2)$, since we fixed μ. Hence she achieves $\beta J(V_1) + (1-\beta) J(V_2)$.

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

A.6 Proof of Proposition 2

Fix an equilibrium. Suppose there exist period $\bar{t}$ (of the current agent's tenure), history $\bar{h}^{\bar{t}}$, and public randomization $\bar{z}_{\bar{t}}$ such that (i) $e(\bar{h}^{\bar{t}}, \bar{z}_{\bar{t}}) = 0$ and (ii) there exists $\tau > \bar{t}$, history h^τ that is a continuation of $\left(h^{\bar{t}}, \bar{z}_{\bar{t}}\right)$, and public randomization z_τ satisfying $e(h^\tau, z_\tau) = 1$.

We show that there exists another equilibrium such that (i) it coincides with the original equilibrium up to period $\bar{t}$, and also after $\left(h^{\bar{t}}, z_{\bar{t}}\right) \neq \left(\bar{h}^{\bar{t}}, \bar{z}_{\bar{t}}\right)$, and (ii) after $\left(\bar{h}^{\bar{t}}, \bar{z}_{\bar{t}}\right)$, P again draws a binary public randomization. After the first realization of the binary draw, the equilibrium is as if we skip period $\bar{t}$, and, after the other realization, the agent retires with probability one (i.e., he exerts $e = 0$ in all future periods). That is, we replace the continuation play after $\left(h^{\bar{t}}, \bar{z}_{\bar{t}}\right)$ with the following two paths: (a) the period of $e(h^\tau, z_\tau) = 1$ is frontloaded by one period; and (b) the agent is allowed to retire. Recursively, we can create another equilibrium in which the agent takes $e = 1$ or he retires.

Let $V(h^t, z_t)$ be the agent's continuation payoff after (h^t, z_t). Since $J(V)$ is concave in V, we have $V(h^t, z_t, \omega) = \frac{1}{\delta}\{V(h^t, z_t) - 1\}$ for each ω. Hence, the principal's payoff equals

$$J\left(V(h^t, z_t)\right) = u(0,1) + \delta J\left(\frac{1}{\delta}\left\{V(h^t, z_t) - 1\right\}\right). \tag{31}$$

Suppose that, at (h^t, z_t), the principal offers the relational contract to bring $\frac{1}{\delta}\{V(h^t, z_t) - 1\}$ with probability δ and the one to bring $\frac{1}{1-\delta}$ (that is, to let the agent retire) with probability

Electronic copy available at: https://ssrn.com/abstract=3243123

$1-\delta$. Then, the agent still obtains the value $V(h^t, z_t)$, and the principal obtains

$$\delta J\left(\frac{1}{\delta}\left\{V(h^t, z_t) - 1\right\}\right) + (1-\delta)\frac{u(0,1)}{1-\delta} = J\left(V(h^t, z_t)\right). \tag{32}$$

Hence, the principal is (weakly) better off.

Note that the best PBE may not be unique since here we start from one equilibrium and create another with frontloaded effort with the same equilibrium payoff for the principal.

A.7 Proof of Lemma 7

Part 1. Concavity with respect to V.

Since μ is fixed, the proof is the same as Lemma 5.

Part 2. Convexity with respect to μ.

Since V is fixed, the proof is the same as Lemma 2.

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 2.

We now show it is strictly increasing for $V \in \left(0, \frac{1}{1-\delta}\right)$. Fix public history h^t with (μ, V) with $V \in \left(0, \frac{1}{1-\delta}\right)$ 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,

$$J(\mu, V) = \frac{V}{\tilde{V}} J(\mu, \tilde{V}) + \left(1 - \frac{V}{\tilde{V}}\right)\bar{J}. \tag{33}$$

Suppose $\tilde{V} = 1$. Then, since $e = 0$, we have $J(\mu, \tilde{V}) = (1-\delta)\underline{v}^P + \delta\bar{J}$, and so $\bar{J} - J(\mu, \tilde{V}) = (1-\delta)\left(\bar{J} - \underline{v}^P\right)$. 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 $\frac{\tilde{V}-1}{\delta} = \frac{\Delta}{\delta}$.

Electronic copy available at: https://ssrn.com/abstract=3243123

Hence, the principal can achieve the payoff at least

$$(1-\delta)\,\underline{v}^P + \delta\left(\frac{\Delta}{\delta}\left((1-\delta)\,\underline{v}^P + \delta\bar{J}\right) + \left(1-\frac{\Delta}{\delta}\right)\bar{J}\right). \tag{34}$$

Hence,

$$\begin{aligned} & \frac{J(\mu,\tilde{V}+\Delta) - J(\mu,\tilde{V})}{\Delta} \\ \geq\ & \frac{(1-\delta)\,\underline{v}^P + \delta\left(\frac{\Delta}{\delta}\left((1-\delta)\,\underline{v}^P + \delta\bar{J}\right) + \left(1-\frac{\Delta}{\delta}\right)\bar{J}\right) - (1-\delta)\,\underline{v}^P - \delta\bar{J}}{\Delta} \\ \geq\ & -(1-\delta)\left(\bar{J} - \underline{v}^P\right). \end{aligned} \tag{35}$$

In total,

$$\frac{d}{d\tilde{V}}\left[\frac{V}{\tilde{V}}J(\mu,\tilde{V}) + \left(1-\frac{V}{\tilde{V}}\right)\bar{J}\right]\Bigg|_{\tilde{V}=1} \geq 0. \tag{36}$$

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.8 Proof of Lemma 8

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:

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 7, 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

$$\beta \times 0 + (1-\beta) \times \hat{V} = V, \tag{37}$$

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

P maximizes

$$\max_{\beta\in[0,1],\hat{V}\in\left[0,\frac{1}{1-\delta}\right]} \beta J(\mu,0) + (1-\beta)\,J(\mu,\hat{V}) \tag{38}$$

subject to

$$\beta \times 0 + (1-\beta) \times \hat{V} = V \text{ and } \hat{V} \geq 1. \tag{39}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

Substituting the constraint, P's payoff is

$$J(\mu, 0) + \frac{V}{\hat{V}}\left[J(\mu, \hat{V}) - J(\mu, 0)\right]. \tag{40}$$

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

$$V\frac{J(\mu, 0) + \left[J_2(\mu, \hat{V})\hat{V} - J(\mu, \hat{V})\right]}{\hat{V}^2}, \tag{41}$$

where J_n is the derivative of J with respect to its nth 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,

$$\frac{d}{d\hat{V}}\left\{J(\mu, 0) + \left[J_2(\mu, \hat{V})\hat{V} - J(\mu, \hat{V})\right]\right\} = \hat{V}\frac{d^2}{d\hat{V}^2}J(\mu, \hat{V}). \tag{42}$$

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 (40),

$$J(\mu, V) = J(\mu, 0) + V \times [J(\mu, 1) - J(\mu, 0)], \tag{43}$$

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

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 (38), $\hat{V} = 1$ — the smallest continuation payoff without immediate replacement — is the unique optimum. Recall that β is defined as the probability of immediate replacement in (37). 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 $c(0) = \lim_{e \to 0} c'(e) = 0$, 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 $(1 - \delta)\underline{v}^P$. 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 $\hat{V} = 1$,

$$J(\mu, 1) = (1 - \delta)\underline{v}^P + \delta\bar{J}. \tag{44}$$

Recall that $\underline{v}^P$ is defined as the principal's dynamic game payoff when no effort is provided and P intervenes every period.

Electronic copy available at: https://ssrn.com/abstract=3243123

It will be useful to verify that the payoff at the arrival of a new agent is higher than $\underline{v}^P$. To see why, the principal can improve upon $\underline{v}^P$ as follows: 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 $\underline{v}^P$ in the first period, and then obtain the continuation payoff of $\delta\underline{v}^P$. In total, we have $\bar{J} > \underline{v}^P$. Given $\bar{J} > \underline{v}^P$, for each (μ, V) with $V \in \left(0, \frac{1}{1-\delta}\right)$, by concavity of $J(\mu, \cdot)$,

$$J(\mu, V) \geq \frac{\frac{1}{1-\delta} - V}{\frac{1}{1-\delta}} \bar{J} + \frac{V}{\frac{1}{1-\delta}} J\left(\mu, \frac{1}{1-\delta}\right) > \underline{v}^P. \tag{45}$$

For $\mu = \mu_H$, (44) together with (45) 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

$$J(\mu_H, V) = \bar{J} \text{ for } V \in [0, V(\mu_H)]. \tag{46}$$

For $\mu > \mu_H$, by Lemma 7, we have $J(\mu, 1) > J(\mu_H, 1) \geq \bar{J}$, which contradicts (44). Hence, $V(\mu) > 1$ as well.

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, (46) implies the slope of the linear part is negative for $\mu < \mu_H$ and positive for $\mu > \mu_H$.

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

Define

$$u^P(\iota_z|\mu, e_z) \equiv \sum_s \Pr(s|\mu, e_z) \cdot u^P(\iota_z|\mu, e_z, s). \tag{47}$$

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:

$$J(\mu) = \max_{(e, \iota, V')} \{u^P(\iota|\mu, e) + \delta \sum_\omega \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega))\}, \tag{48}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

subject to incentive compatibility constraint:

$$e \in \arg\max 1 - c(e) + \delta \sum_{\omega} \Pr(\omega|e,\iota) V'(\omega). \tag{49}$$

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

$$\mu'(\mu,e,\bar{\omega}) = \frac{\mu \Pr(\bar{\omega}|e,\iota)}{\mu \Pr(\bar{\omega}|e,\iota) + (1-\mu)\Pr(\bar{\omega}|0,\iota)} \leq \mu, \tag{50}$$

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

$$\begin{aligned} 0 &= \frac{d}{dV'(\bar{\omega})}\{u^P(\iota|\mu,e) + \delta \sum_{\omega} \Pr(\omega|\mu,e,\iota) J(\mu'(\mu,e,\omega), V'(\omega))\} \\ &= \{u_e^P(\iota|\mu,e) \\ &\quad +\delta \sum_{\omega} \Pr_e(\omega|\mu,e,\iota) J(\mu'(\mu,e,\omega), V'(\omega)) \\ &\quad +\delta \sum_{\omega} \Pr(\omega|\mu,e,\iota) J_1(\mu'(\mu,e,\omega), V'(\omega)) \mu_e'(\mu,e,\omega)\} \frac{de}{dV'(\bar{\omega})} \\ &\quad +\delta \Pr(\bar{\omega}|\mu,e,\iota) J_2(\mu'(\mu,e,\bar{\omega}), V'(\bar{\omega})), \end{aligned} \tag{51}$$

where J_n is the derivative of J with respect to its nth 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,

$$\begin{aligned} &\{u_e^P(\iota|\mu,e) + \delta \sum_{\omega} \Pr_e(\omega|\mu,e,\iota) J(\mu'(\mu,e,\omega), V'(\omega)) \\ &+\delta \sum_{\omega} \Pr(\omega|\mu,e,\iota) J_1(\mu'(\mu,e,\omega), V'(\omega)) \mu_e'(\mu,e,\omega)\} < 0. \end{aligned} \tag{52}$$

Similarly, if there exists $\hat{\omega}$ such that $\Pr(\hat{\omega}|e,\iota)$ is decreasing in e but

Electronic copy available at: https://ssrn.com/abstract=3243123

$V'(\hat{\omega}) \leq \arg\max_{\hat{V}} J(\mu'(\mu, e, \hat{\omega}), \hat{V})$, then the symmetric argument implies that

$$\{u_e^P(\iota|\mu, e) + \delta \sum_{\omega} \Pr_e(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega))$$
$$+\delta \sum_{\omega} \Pr(\omega|\mu, e, \iota) J_1(\mu'(\mu, e, \omega), V'(\omega)) \mu_e'(\mu, e, \omega)\} \geq 0, \tag{53}$$

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,

$$u^P(\iota|\mu, e^*) > u^P(\iota|\mu, e). \tag{54}$$

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

$$\sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V'(\omega)) < \sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)). \tag{55}$$

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

$$J(\mu'(\mu, e, \omega), V^*(\omega)) < J(\mu'(\mu, e, \hat{\omega}), V^*(\hat{\omega})) \tag{56}$$

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

$$\sum_{\omega} \Pr(\omega|\mu, e, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)) < \sum_{\omega} \Pr(\omega|\mu, e^*, \iota) J(\mu'(\mu, e, \omega), V^*(\omega)). \tag{57}$$

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. J is convex in μ, so

$$J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \leq J(\mu_2, V^*(\mu_1)). \tag{58}$$

$V^*(\mu_2)$ maximizes $J(\mu_2, V)$ at μ_2, so

$$J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \leq J(\mu_2, V^*(\mu_2)). \tag{59}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

Also,

$$J\left(\mu_1, V^*\left(\mu_1\right)\right) \geq J\left(\mu_2, V^*\left(\mu_2\right)\right) - J_1\left(\mu_2, V^*\left(\mu_2\right)\right)\left[\mu_2 - \mu_1\right], \tag{60}$$

since J is convex in μ. From the first inequality of the proof,

$$\begin{aligned} J\left(\mu_1, V^*\left(\mu_1\right)\right) \quad \geq \quad & J\left(\mu_1, V^*\left(\mu_1\right)\right) + J_1\left(\mu_1, V^*\left(\mu_1\right)\right)\left[\mu_2 - \mu_1\right] \\ & -J_1\left(\mu_2, V^*\left(\mu_2\right)\right)\left[\mu_2 - \mu_1\right]. \end{aligned} \tag{61}$$

Hence,

$$0 \geq \left[J_1\left(\mu_1, V^*\left(\mu_1\right)\right) - J_1\left(\mu_2, V^*\left(\mu_2\right)\right)\right]\left(\mu_2 - \mu_1\right). \tag{62}$$

■

Given this claim, $J_1\left(\mu'\left(\mu, e, \omega\right), V^*\left(\omega\right)\right)$ is larger for ω with $\mu'\left(\mu, e, \omega\right) > \mu$ than for ω with $\mu'\left(\mu, e, \omega\right) < \mu$. Since the distribution of $\left\{\mu'\left(\mu, e^*, \omega\right)\right\}_\omega$ given e^* is the mean-preserving spread of the distribution of $\left\{\mu'\left(\mu, e, \omega\right)\right\}_\omega$ given e and we have $\mu'\left(\mu, e^*, \omega\right) \geq \mu'\left(\mu, e, \omega\right)$ if and only if ω satisfies $\mu'\left(\mu, e, \omega\right) \geq \mu$, faster learning increases the continuation payoff. Together with (55) and (57), this leads to

$$\sum_\omega \Pr\left(\omega|\mu, e, \iota\right) J\left(\mu'\left(\mu, e, \omega\right), V'\left(\omega\right)\right) < \sum_\omega \Pr\left(\omega|\mu, e, \iota^*\right) J\left(\mu'\left(\mu, e^*, \omega\right), V^*\left(\omega\right)\right). \tag{63}$$

Together with (54), we have proven that P's payoff increases.

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

A.9 Proof of Lemma 9

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, and, given e, the distribution of the updated beliefs $\left(\mu'\left(\mu, e, \omega\right)\right)_\omega$ after no intervention is a mean-preserving spread of that after intervention.

In particular, the belief update is given by

$$\begin{aligned} \mu'(\mu, e, b, I) &= \frac{\mu \Pr\left(b|e\right)}{\mu \Pr\left(b|e\right) + \left(1 - \mu\right) \Pr\left(b|0\right)}, \\ \mu'(\mu, e, b, G) &= \frac{\mu \Pr\left(b, G|e\right)}{\mu \Pr\left(b, G|e\right) + \left(1 - \mu\right) \Pr\left(b, G|0\right)}, \\ \mu'(\mu, e, b, B) &= \frac{\mu \Pr\left(b, B|e\right)}{\mu \Pr\left(b, B|e\right) + \left(1 - \mu\right) \Pr\left(b, B|e\right)}. \end{aligned} \tag{64}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

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

$$\begin{aligned} d(\mu) \quad : &= \sum_{\omega} \Pr(\omega|\mu, e, \iota = 0) \left(\mu'(\mu, e, \omega) - \mu\right)^2 - \sum_{\omega} \Pr(\omega|\mu, e, \iota = 1) \left(\mu'(\mu, e, \omega) - \mu\right)^2 \\ &= \sum_{y \in \{G,B\}} \frac{\mu^2 (1-\mu)^2 \left(\Pr(b, y|e) - \Pr(b, y|0)\right)^2}{\mu \Pr(b, y|e) + (1-\mu) \Pr(b, y|0)} \\ &\quad - \frac{\mu^2 (1-\mu)^2 \left(\Pr(b|e) - \Pr(b|0)\right)^2}{\mu \Pr(b|e) + (1-\mu) \Pr(b|0)}. \end{aligned} \tag{65}$$

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

$$\sum_{y \in \{G,B\}} \frac{\Pr(b, y|e)^2 \Pr(b, y|0)^2}{\left(\mu \Pr(b, y|e) + (1-\mu) \Pr(b, y|0)\right)^3} - \frac{\Pr(b|e)^2 \Pr(b|0)^2}{\left(\mu \Pr(b|e) + (1-\mu) \Pr(b|0)\right)^3}. \tag{66}$$

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

$$\begin{aligned} (a, b) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} &= \frac{2\left(b^2 x^2 - abxy + a^2 y^2\right)\left(\mu^2 x^2 + 4\mu(1-\mu) xy + y^2\right)}{\left(\mu x + (1-\mu) y\right)^5} \\ &\geq 0, \end{aligned} \tag{67}$$

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

A.10 Proof of Proposition 3

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.

Intervention is Optimal in the Initial Period

Lemma 10 *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

$$-l - \left[\mu_H \times \Pr(G|e, b) \times 0 + (1 - \mu_H)(-C)\right] \geq C - l - \mu_H C. \tag{68}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

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

$$\delta\left[\mu_H \times 0 + (1-\mu_H)\max_V J(\mu_H, V)\right] - \delta J(\mu_b, V_b). \tag{69}$$

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

$$\max_V J(\mu_H, V) \leq \mu_H \times 0 + (1-\mu_H)\left(\Pr(y=B|0)(-l) + \delta \max_V J(\mu_H, V)\right). \tag{70}$$

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

$$\max_V J(\mu_H, V) \leq \frac{-(1-\mu_H)(1-\bar{q})\,l}{1-(1-\mu_H)\,\delta}. \tag{71}$$

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

$$\delta\left((1-\mu_H)\frac{-(1-\mu_H)(1-\bar{q})l}{1-(1-\mu_H)\,\delta} - \frac{-l}{1-\delta}\right). \tag{72}$$

Hence, if

$$C - l - \mu_H C > \delta\left((1-\mu_H)\frac{-(1-\mu_H)(1-\bar{q})l}{1-(1-\mu_H)\,\delta} - \frac{-l}{1-\delta}\right), \tag{73}$$

then intervention is uniquely optimal. At $\mu_H = 0$ and $\bar{q} = 0$, (73) holds since $C - l > 0$. 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 (73). ■

Intervention at the Limit

Lemma 11 *For any parameter values, if we start from $\mu = \mu_H$ and $V = \arg\max_{\tilde{V}} J(\mu_H, \tilde{V})$, then after ω with $\Pr_e(\omega|e) < 0$, we have $J_2(\mu'(\mu, e_z, \omega), V_z'(\omega)) = 0$.*[18]

Proof. From Lemma 2, we have $J(\mu'(\mu, e_z, \omega), V_z'(\omega)) = \max_V J(\mu_H, V)$. Hence, Lemma 8 implies the result. ■

[18] J_n is the derivative of J with respect to its nth argument.

Electronic copy available at: https://ssrn.com/abstract=3243123

We form the Lagrangian

$$\begin{aligned} J(\mu, V) &= \int_z (1-\rho_z)\bar{J} + \rho_z u^P(\iota_z|\mu, e_z) + \delta \sum_\omega \Pr(\omega|\mu, e_z, \iota_z) J(\mu'(\mu, e_z, \omega), V_z'(\omega))\, dz \\ &\quad + \lambda \left(V - \int_z \rho_z \left\{ 1 - c(e_z) + \delta \sum_\omega \Pr(\omega|e_z, \iota_z) V_z'(\omega) \right\} dz \right) \\ &\quad + \int_z \rho_z \eta_z \left(\delta \sum_\omega \Pr_e(\omega|e_z, \iota_z) V_z'(\omega) - c'(e_z) \right) dz \end{aligned} \tag{74}$$

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

$$\begin{aligned} e_z \quad &: \quad -J_2(\mu, V) c'(e_z) + \eta_z c''(e_z) \\ &= \quad u_e^P(\iota_z|\mu, e_z) + \delta \cdot \sum_\omega \Pr_e(\omega|\mu, e_z, \iota_z) \cdot J(\mu'(\mu, e_z, \omega), V_z'(\omega)) \\ &\quad + \delta \cdot \sum_\omega \Pr(\omega|\mu, e_z, \iota_z) \cdot J_1(\mu'(\mu, e_z, \omega), V_z'(\omega)) \cdot \mu_e'(\mu, e_z, \omega) \\ &\quad - \delta \cdot \sum_\omega \Pr_e(\omega|e_z, \iota_z) \cdot V_z'(\omega) \cdot J_2(\mu, V) \\ &\quad + \delta \cdot \eta_z \cdot \sum_\omega \Pr_{ee}(\omega|e_z, \iota_z) \cdot V_z'(\omega), \end{aligned} \tag{75}$$

and

$$V_z'(\omega) : J_2(\mu'(\mu, e_z, \omega), V_z'(\omega)) = \frac{\Pr(\omega|e_z, \iota_z)}{\Pr(\omega|\mu, e_z, \iota_z)} J_2(\mu, V) - \eta_z \frac{\Pr_e(\omega|e_z, \iota_z)}{\Pr(\omega|\mu, e_z, \iota_z)}. \tag{76}$$

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

Lemma 12 *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 (76), it suffices to show that $\eta > 0$. If $\eta = 0$, since $J_2(\mu, V) = 0$ in the initial period, Lemma 11 and (75) yield

$$\begin{aligned} 0 &= u_e^P(\iota_1|\mu_1, e_1) + \delta \sum_{\tilde{s}_1, \tilde{o}_1} \Pr_e(\tilde{\omega}_1|\mu_1, e_1, \iota_1) \cdot J(\mu'(\mu_1, e_1, \tilde{\omega}_1), V'(\tilde{\omega}_1)) \\ &\quad + \delta \sum_{\tilde{\omega}_1 : \mu_e'(\mu_1, e_1, \tilde{\omega}_1) > 0} \Pr(\tilde{\omega}_1|\mu_1, e_1, \iota_1) \cdot J_1(\mu'(\mu_1, e_1 \tilde{\omega}_1), V'(\tilde{\omega}_1)) \cdot \mu_e'(\mu_1, e_1, \tilde{\omega}_1). \end{aligned} \tag{77}$$

Electronic copy available at: https://ssrn.com/abstract=3243123

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.[19] 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 (76),

$$J_2(\mu'(\mu_t, e_t, \omega_t), V_{t+1}(\omega_t)) \leq \frac{\Pr(\omega_t|e_t, \iota_t)}{\Pr(\omega_t|\mu_t, e_t, \iota_t)} J_2(\mu_t, V_t) - \eta_t \frac{\Pr_e(\omega_t|e_t, \iota_t)}{\Pr(\omega_t|\mu_t, e_t, \iota_t)}, \tag{78}$$

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 $-\infty$.

Suppose μ_t converge to 1 and η_t converges to 0. At this limit, (75) converges to

$$\begin{aligned} -J_2(1, V)\, c'(e) &= u_e^P(1, e, \iota) + \delta \sum_\omega \Pr_e(\omega|1, e, \iota)\, J(\mu'(1, e, \omega), V'(\omega)) \\ &\quad + \delta \sum_\omega \Pr(\omega|e, \iota)\, J_1(1, V'(\omega))\, \mu_e'(1, e, \omega) \\ &\quad - \delta \sum_\omega \Pr_e(\omega|e, \iota)\, V'(\omega)\, J_2(1, V). \end{aligned} \tag{79}$$

Since

$$\mu_e'(1, e, \omega) = \lim_{\mu \to 1} \left(\frac{d}{de} \frac{\mu \Pr(\omega|e)}{\mu \Pr(\omega|e) + (1-\mu) \Pr(\omega|0)} \right) = 0 \tag{80}$$

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 (15),

$$0 = u_e^P(\iota|1, e) + \delta \sum_\omega \Pr_e(\omega|e, \iota)\, J(1, V'(\omega)). \tag{81}$$

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

$$\frac{d}{de}[u^P(\iota|1, e) + \delta \sum_\omega \Pr(\omega|e, \iota)\, J(1, V'(\omega))], \tag{82}$$

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:

[19] Otherwise, 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 11, and (iii) parts 2 and 3 of the proof to Lemma 8(omitting z for notational simplicity) imply

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

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

Electronic copy available at: https://ssrn.com/abstract=3243123

Lemma 13 *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 $V \in \left[0, \frac{1}{1-\delta}\right]$, 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. ■

No intervention is Optimal in a Period after the Initial Period

The following lemma ensures that e_1 is bounded below:

Lemma 14 *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 11, and (iii) parts 2 and 3 of Lemma 8 (omitting z for notational simplicity), we have

$$c'(e_1) = \delta \sum_{\omega_1 : \Pr_e(\omega_1|e_1) > 0} \Pr_e(\omega_1|e_1, \iota_1) \geq \delta \Pr_e(g, G|e_1, \iota_1) = \delta \Pr_e(g, G|e_1). \tag{83}$$

Hence, $e_1 \geq \bar{e}$. ■

Lemma 15 *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

$$u^P(\iota_1|\mu_H, \bar{e}) + \delta \max\left\{\max_V J(\mu_H, V), \frac{1}{1-\delta} \Pr(s = b|\bar{e})(-l)\right\}. \tag{84}$$

Here, to obtain an upper bound, we allow P to replace the L-type at the end of period 1, and she learns that the agent is an 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 P learned 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).

Electronic copy available at: https://ssrn.com/abstract=3243123

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

$$\begin{aligned} & u^P(\iota_1|\mu_H, \bar{e}) \\ & + \sum_{t=1}^{T} \delta^{t-1} \left\{ \prod_{\tau=1}^{t-1} \Pr(\omega_\tau = (g, G)) \right\} \Pr(\omega_t \neq (g, G)) \left(-C + \delta \max_V J(\mu_H, V) \right) \\ & + \delta^{T-1} \prod_{\tau=1}^{T} \Pr(\omega_\tau = (g, G)) \frac{-l}{1-\delta}, \end{aligned} \tag{85}$$

where the probability is determined by the initial belief μ_H and the H-type agent 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 replaces 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 $t = T + 1, ...$

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

Electronic copy available at: https://ssrn.com/abstract=3243123