Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring *

Iván Marinovic† Andrzej Skrzypacz ‡ Felipe Varas §

October 18, 2016

PRELIMINARY AND INCOMPLETE

Abstract

This paper studies the design of monitoring/audit policies in dynamic settings. Firm quality is private information and evolves stochastically via a Markov process with transitions depending on the firm's unobservable effort. The firm has benefits (linearly) from its reputation for quality. A principal designs a monitoring policy that allows him to learn the firm?s quality by conducting costly reviews. Monitoring plays two roles. First, it plays an incentive role, because it affects the firm's reputation. Second, information is directly valuable when the principal?s payoff is convex in reputation, for example because it allows consumers make better choices. Our main result is a characterization of the optimal monitoring policy that induces full effort. The policy is surprisingly simple. It is either deterministic, with a pre-announced date of next monitoring, or random with a constant hazard rate of next inspection. We discuss how the type of optimal monitoring depends on recent history and on the parameters of the problem. We also consider how the optimal monitoring policy is affected by the presence of exogenous news, showing how that the evolution of the hazard of random monitoring depends on whether the absence of news is perceived to be good or bad news about quality.

JEL Classification: C73, D82, D83, D84.
Keywords: Monitoring, Auditing, Dynamic Contracts, Dynamic Games, Reputation.

1 Introduction

Should we test students using random quizzes or pre-announced tests? Should we inspect restaurant hygiene at pre-determined intervals or should we use surprise inspections? How often and how predictably should we test quality of schools, HMOs, health care providers, etc.? How should an industry self-regulate a voluntary licensing program, in particular how and when should it audit its members for compliance? What about timing of internal audits/inspections for the purpose of allocating internal resources within organizations?

*We thank Guy Barles, Douglas Carmichael, and Bob Wilson for helpful feedback and suggestions, and seminar participants at Stanford for helpful comments.
†Stanford University, GSB. email: imvial@stanford.edu
‡Stanford University, GSB. email: skrz@stanford.edu
§Duke University, Fuqua School of Business. email: felipe.varas@duke.edu

These seem to be important questions for provision of complicated products and services that have hard-to-verify quality (or other attributes), especially when efforts to improve or maintain quality are not publicly observable and are subject to moral hazard.

To develop some economic intuition about these related problems, we propose and analyze a stylized model of a principal-agent interaction. The agent/firm provides a service with an unobserved quality. Quality evolves stochastically over time in a partially persistent way that depends on the firms private effort (which is a source of an agency problem). The firm cares about its reputation, which is defined as outsiders belief about current quality. The principal cares about quality and can perform costly monitoring to publicly reveal current state of quality. To identify the effects of the information channel alone, we assume away any explicit fines or bonuses for failing or passing the tests (which is realistic in some of these applications, less so in others) and hence in our model incentives are provided only via reputation. Under a few simplifying assumptions we characterize an optimal monitoring policy.

Our first general insight is that in most of situations that motivate our paper, there are two main reasons for inspections/monitoring. First, since the firm cares about its reputation, providing information to the market about current quality can mitigate the agency problem of under-provision of quality. Second, when the principal's payoff is convex in beliefs, for example because having better information allows the principal allocate resources or consumers better, information has direct value.[1]

Most real-life monitoring policies fall into one of two categories: random inspections/tests, with positive hazard rate of an inspection at most times, or deterministic inspections, taking place at pre-determined dates. At first, neither of these policies seems optimal. A deterministic timing of inspections runs into the risk of window dressing and resting on laurels because a firm knowing that an inspection is coming soon has very strong incentives to put effort into improving its service, while soon after an inspection, the incentives are the weakest. On the other hand, when quality is persistent, a positive hazard rate of inspections right after one has been performed seems unnecessary and wasteful.

Our analysis explains why these two strategies can be part of an optimal monitoring policy. The intuition is that optimal design depends on the role that monitoring plays: when monitoring is mostly for the sake of incentive provision, random monitoring tends to be optimal. When it is mostly about collecting information for better allocation (of consumers or resources) then deterministic monitoring tends to be optimal. Somewhat surprisingly, we show that in our model an optimal policy is not a mixture of random and deterministic monitoring, but instead it is always one of the two extremes (in some cases the optimal policy uses the deterministic testing after bad results and random testing after good results, or vice versa, but the optimal policy never switches between these two modes in between inspections). We characterize how the parameters of the problem affect the optimal choice between these two extreme policies. For example, the less costly is effort, the more likely the optimal policy is deterministic. On the other hand, the more expensive are inspections, the more likely the optimal policy is random.

In our benchmark model the only source of the firms reputation are the results of inspections. In that case we show that when the optimal policy is random monitoring, the optimal intensity is constant over time. We then extend the model to allow exogenous news process as well, for example from third-party reviews or some client experiences that become public. We show how the process of information arrival affects the optimal intensity of monitoring. For example, we provide sufficient conditions for the optimal intensity of testing to be increasing after bad test results.

[1] In some situations inspections play also a third role, that of information sharing. For example, regulators may want to test schools to identify the best performers and then try to transfer knowledge about what makes those schools particularly effective in the hope of improving other schools.

1.1 Related Literature

There is a large empirical literature on the importance of quality monitoring and reporting systems. For example, Epstein et al (2000) argues that public reporting on the quality of health care in the U.S. (via quality report cards) has become the most visible national effort to manage quality of health care. A large literature documents the effect of quality report cards across various industries. Some examples include restaurant hygiene report cards (e.g., Jin and Leslie, 2003), school report cards (e.g. Figlio and Lucas, 2004), and a number of disclosure programs in the health care industry, in particular coronary artery bypass graft (CABG) surgery mortality report cards (e.g., Dranove, Kessler, McClellan, and Satterthwaite, 2003), health plan report cards (e.g., Dafny and Dranove, 2008), hospital rankings (e.g., Pope, 2009), nursing homes report cards (e.g., Lu, 2011), fertility clinics report cards (e.g., Bundorf, Chun, Goda, and Kessler, 2009), and hospital infection rates report cards (e.g., Kim and Black, 2011). Zhang et al. (2011) note that during the past few decades, quality report cards have become increasingly popular, especially in areas such as health care, education, and finance. The underlying rationale for these report cards is that disclosing quality information can help consumers make better choices and encourage sellers to improve product quality. This observation is the basis of our stylized model.

Eccles et al (2007) assert that "in an economy where 70% to 80% of market value comes from hard-to-assess intangible assets such as brand equity, intellectual capital, and goodwill, organizations are especially vulnerable to anything that damages their reputations," suggesting that our focus on the provision of incentives via reputation may be warranted for some markets. Some existing studies provide evidence in support of the effectiveness of report cards, documenting that consumers use them to select better-rated sellers and sellers respond by improving product quality, but others have raised concerns by showing that report cards may induce sellers to game the system in ways that hurt consumers. In the context of schools, some argue that testing school quality can be detrimental for quality provision. For example, Hoffman, Assaf, and Paris (2001) study the results from Texas Assessment of Academic Skills testing and found some evidence that this program has a negative impact on students, especially low achieving and minority students. While our model does not have the richness to address all such issues, it is aimed at contributing to our understanding of properties of good monitoring programs (for example, in our model we assume that testing results in perfect observation of quality, while in reality quality may be sometimes hard to measure and the available noisy measures may be subject to manipulation. Since our focus is on the dynamic effects of monitoring policies, we assume those important problems away).

On the theoretical side, this paper is closely related to Lazear (2006) and Eeckhout et al. (2010) who study the optimal allocation of monitoring resources in static settings and without reputation concerns. Lazear concludes that monitoring should be predictable/deterministic when monitoring is very costly, otherwise it should be random. Both papers are concerned with maximizing the level of compliance given a limited amount of monitoring resources. Optimality requires that the incentive compatibility constraint of complying agents be binding or else some monitoring resources could be redeployed to induce compliance by some non-complying agents. Both papers consider static settings, and ignore the reputation effect of monitoring, which is the focus of our study.

Another related literature is on the deterrence effect of policing and enforcement and the optimal monitoring policy to deter criminal behavior in static settings. See for example Becker (1968), Townsend (1979), Polinsky and Shavell (1984), Reinganum and Wilde (1985), Mookherjee and Png (1989), Bassetto and Phelan (2008), Bond and Hagerty (2010)).

Yet another literature studies the design of reputation systems or rating mechanisms. The literature has explored the design of rating mechanisms. Dellarocas (2006) studies how the frequency of reputation profile updates affects cooperation and efficiency in settings with noisy ratings. Horner and Lambert (2016) study the incentive provision aspect of information systems in a career concern setting similar to Holmstrom (1999). In their setting acquiring information is not costly and does not have value per se. See also Ekmekci (2011) for a study of optimal design of rating systems with commitment types.

We build on the investment and reputation model from Board and Meyer-ter Vehn (2013) where the firm's quality type changes stochastically. Unlike that paper, we analyze the optimal design of monitoring policy, while they take the information process as exogenous (in their model it is a Poisson process of exogenous news). They study equilibrium outcomes of a game, while we solve a design problem (design of a monitoring policy). Moreover, we allow for a principal to have convex preferences in perceived quality, so that information has direct benefits, an assumption that does not have a direct counterpart in their model. Finally, we allow for richer evolution of quality: in Board and Meyer-ter Vehn (2013) it is assumed that if the firm puts full effort, quality never drops from high to low, while in our model even with full effort quality remains stochastic. In the end of the paper we also discuss that some of our results can be extended beyond the Board and Meyer-ter Vehn (2013) model of binary quality levels and we also consider design of optimal monitoring when some information comes exogenously.

Finally, a recent paper by Dilme and Garret (2015) considers the deterrence effect of convictions in a dynamic setting without commitment where a monitor benefits from catching offenders but faces a fixed cost of switching from a passive monitoring state to an active one. The presence of this cost gives rise to interesting dynamics. Our model of monitoring costs are quite different and we also study the commitment case, so the theoretical analysis in these two papers is quite different.

2 Setting

Agents, Technology and Effort: There are two players: a principal and a firm. Time $t \in [0, \infty)$ is continuous.

The firm has a product with quality that changes over time. To model the evolution of quality, we build on Board and Meyer-ter Vehn (2013). At time t, the quality of the product is $\theta_t \in \{L, H\}$, where we normalize $L = 0$ and $H = 1$. Initial quality is exogenous and commonly known. Subsequently quality evolves stochastically in a way that depends on the firm's effort. In each moment, the firm makes a private effort decision $a_t \in [0, \bar{a}]$, $\bar{a} < 1$. Changes in quality are not publicly observed. In most of the paper we assume that when the firm chooses effort a_t then quality switches from low to high with intensity λa_t, while a high quality firm experiences a quality drop with intensity $\lambda(1 - a_t)$. Later in the paper we illustrate how the analysis can be extended to a model with a continuum of levels of quality where effort affects the long-run mean of the quality distribution. Note that unlike Board and Meyer-ter Vehn (2013), we bound a_t below one so that quality is random even if the firm chooses to always exert full effort. The steady-state distribution of quality when the firm puts full effort is $\Pr(\theta = H) = \bar{a}$.

Strategies and Information: At time t, the principal can inspect/monitor the quality of the product, in which case θ_t becomes public information. A monitoring policy M specifies an increasing sequence of dates $(T_n)_{n \geq 1}$ when the principal inspects quality.[2] Accordingly, the information generated by the monitoring

[2]We implicitly assume the principal discloses the quality after the inspection. This is optimal: the principal would never benefit from withholding the quality information because that would weaken the incentive power of monitoring.

policy is given by the filtration $\mathcal{F}_t^M = \sigma(T_n, \theta_{T_n} : T_n \leq t)$. A monitoring policy is represented by a sequence of cumulative density functions, $F_n(t) = \Pr\left(T_n < T_{n-1} + t \mid \mathcal{F}^M_{T_{n-1}}\right)$ over the time of the n-th inspection.

For concreteness we assume that current quality is always privately known by the firm.[3] A strategy for the firm is an effort plan $a = (a_t)_{t\geq 0}$ that is predictable with respect to the filtration generated by $\theta = (\theta_t)_{t\geq 0}$ and the monitoring policy M, that is $\mathcal{F}_t^\theta = \sigma(T_n, \theta_s : T_n \leq t, s \leq t)$.

Reputation and Payoffs: We model the firm's payoffs in a reduced form as being generated by market demand that depends on the firm's reputation. In particular, denote market's beliefs about the firm's effort strategy by $\tilde{a} = (\tilde{a}_t)_{t\geq 0}$. Firm's reputation at time t is given by $x_t \equiv E^{\tilde{a}}(\theta_t | \mathcal{F}_t^M)$ where the expectation is taken with respect to the measure induced by the conjectured effort, $\tilde{a}$. In words, reputation is the market's belief about firm's current quality. It evolves based on the market's conjecture about firm's strategy and in response to the principal's monitoring.

The firm is risk neutral and discounts future payoffs at rate $r > 0$. For tractability we assume that the firm's payoff flow is linear in its reputation.[4] Firm's effort has a marginal cost of k, hence the firm's expected payoff at time t is

$$\Pi_t = E^a \left[\int_t^\infty e^{-r(s-t)} (x_s - k a_s) \mathrm{d}s \Big| \mathcal{F}_t^\theta \right]$$

The principal discounts the future at the same rate r as the firm. The principal's flow payoff when the firm's reputation is x_t is given by a strictly increasing, weakly convex function $u(\cdot)$. Also, monitoring is costly: the lump-sum cost of monitoring at a given point is c. Hence, the principal payoff is

$$U_t = E^{\tilde{a}} \left[\int_t^\infty e^{-r(s-t)} u(x_s) \mathrm{d}s - \sum_{T_n \geq t} e^{-r(T_n - t)} c \Big| \mathcal{F}_t^M \right].$$

Note that we do not include the cost of firm's effort in the principal's payoff. In some applications it may be more natural to assume that the principal internalizes that cost and then we would subtract $-k\tilde{a}_s$ from the welfare flows. However, since in this paper we focus on policies that induce full effort ($a_t = \bar{a}$ for all t), our analysis does not depend on how the principal accounts for the firm's cost of effort in his payoff function (of course the cost still matters indirectly for the incentive constraints).

We consider both the case in which the principal payoff $u(\cdot)$ is linear as well as that in which $u(\cdot)$ is strictly convex. In the convex payoff case, we say that information has social value (in addition to distributive effects and in addition to creating value via incentive provision). Such convexity of the principal's flow payoff represents situations where information about quality affects not only prices but also allocations - for example information may allow relocation of consumers from low-quality to high-quality firms and the principal may internalize consumer surplus. We provide some stylized examples in the next section.

Incentive Compatibility and Optimal Policies An effort policy a is incentive compatible given a monitoring policy M if it is consistent with the firm's optimization given that policy and market beliefs $\tilde{a} = a$ (that are correct on the equilibrium path). In other words, if a is consistent with equilibrium given

[3] As we discuss below, given our assumptions on the evolution of quality, our results immediately extend to the case where the firm does not privately observe quality, which may be a more realistic assumption for some applications.

[4] One interpretation is that the firm sells a unit flow of supply to a competitive market where consumers' willingness to pay is equal to the expected quality, so that in every instance price is equal to the firm's current reputation. We discuss alternative interpretations in the next section.

M:[5]

Definition 1. *Fix a monitoring policy M. An equilibrium is a pair $(\tilde{a}, a)$ such that for every history on the equilibrium path:*

1. *x_t is consistent with Bayes' rule, given M and $\tilde{a}$.*
2. *a maximizes Π.*
3. *$\tilde{a} = a$.*

The goal of this paper is to characterize optimal monitoring policies among those that induce full effort. One interpretation is that we implicitly assume that the parameters of the problem are such that despite agency problems, it is optimal for the principal to induce full effort.

We assume that the principal commits to a monitoring policy from the start and that the firm chooses full effort whenever there exists an equilibrium given M that has full effort. (So that, as usual in contract theory, if there were multiple equilibria given M, we select the one with full effort whenever possible).

Definition 2. *A monitoring policy M is incentive compatible if for that policy there exists an equilibrium with $a_t = \bar{a}$. We call M optimal if it maximizes U over all incentive compatible monitoring polices.*

An optimal policy tries to achieve two goals. First, it tries to minimize the cost of inspections subject to maintaining incentives for effort provision (it is easy to satisfy incentives by very frequent monitoring, but that would be excessively costly). Second, since the principal values information per se (if $u(\cdot)$ is strictly convex) the policy solves a real-option problem of deciding when to spend the cost to learn current quality in order to temporarily benefit from superior information.

2.1 Examples

Before we start the analysis, we discuss a few examples of applications that are captured by our general model. They illustrate how the firm and principal payoffs can be micro-founded.

Example 1: Quality Certification. Consider a classic problem of moral hazard in quality provision, as studied by the reputation literature. In particular, as in Mailath and Samuelson (2001) and Board and Meyer-ter Vehn (2013), consider a monopolist selling a product to a competitive mass of consumers with preferences $x_t - p$, where p is the price of the product. The monopolist sells one unit of output flow per instant. In equilibrium, consumers pay $p_t = x_t$ (by Bertrand competition) and get zero consumer surplus. The firm's profit flow given reputation x_t is then

$$\pi(x_t) = x_t - ka_t,$$

. The principal is a regulator who maximizes total surplus and its payoff flow (not including monitoring costs) is

$$u(x_t) = \alpha\pi(x_t),$$

where α is the weight attached by the regulator to the firm's payoff (since consumers receive no surplus, the regulator's payoff is proportional to the firm's payoff). In this application, principal's preferences are linear

[5] We could define a third player in the model, the market, and then define the equilibrium as a Perfect Bayesian equilibrium of the game induced by the policy M. We hope our simpler definition does not create confusion.

in beliefs, so information has no social value (beyond incentive provision). As a result of this linearity, the optimal policy simply minimizes the cost of monitoring subject to the constraint of inducing full effort.

More generally, there are at least two ways to interpret a model in which the principal's payoff is linear in the firm's payoff. It applies when the principal maximizes total welfare in a market where buyers compete away all consumer surplus. Or, buyers earn some consumer surplus, but the principal represents a self-regulatory organization (SRO) aiming to maximize the overall industry profits (and firms in the industry are local monopolies).

Example 2: School Monitoring. A second application is monitoring of school quality in the presence of horizontal differentiation and school choice. Consider a Hotelling model of school choice with two schools: School A, with a known, constant quality, and school B with changing quality. The two schools, are located at each extreme of the unit line. Evolution of quality of school B depends on the school's hidden effort/investment and is unobservable unless a regulator monitors it. Students are distributed uniformly over the unit line. All schools cost the same amount and students choose them based on distance and perceived quality differences. Assume that the quality of school A is known to be low. If a student is located at location $\ell \in [0,1]$ she derives a utility of attending school A equal to

$$v_A(\ell) = -\ell^2.$$

On the other hand, the utility of attending school B depends on its reputation and is given by

$$v_B(x_t, \ell) = x_t - (1-\ell)^2$$

Given reputation x_t, students above $\ell^*(x_t) = \frac{1-x_t}{2}$ choose school B. Hence the demand for school B is:

$$1 - \ell^*(x_t) = \frac{1+x_t}{2}.$$

Now, assume that for each attending student, the schools receive transfer of \$1 from the government and that marginal costs are normalized to zero. Hence, the profit flows of schools A and B are

$$\begin{aligned}
\pi_A(x_t) &= \ell^*(x_t) = \frac{1-x_t}{2} \\
\pi_B(x_t) &= (1-\ell^*(x_t)) - ka_t = \frac{1+x_t}{2} - ka_t.
\end{aligned}$$

Conditional on school $B's$ reputation x_t, total students' welfare is

$$\begin{aligned}
w(x_t) &= \int_0^{\ell^*(x_t)} v_A(\ell)\mathrm{d}\ell + \int_{\ell^*(x_t)}^1 v_B(x_t,\ell)\mathrm{d}\ell \\
&= \frac{1}{4}x_t^2 + \frac{1}{2}x_t - \frac{1}{12}
\end{aligned}$$

Finally, suppose that the principal's (school regulator's) payoff in each period t is a weighted average of the students' and schools' welfare:

$$u(x_t) = \alpha w(x_t) + (1-\alpha)(\pi_A(x_t) + \pi_B(x_t)),$$

where α is the relative weight attached to students' utility by the principal. Note that the principal's flow utility $u(x_t)$ is an increasing and convex function of reputation, even though the sum of profits of the firms does not depend on it (since the schools just split the subsidy per student, for profits reputation has only distributive effects). The convexity reflects the fact that better information about the quality of school B leads to a more efficient allocation of students and the principal internalizes their welfare.

Example 3: Capital Budgeting and Internal Capital Markets. In the next example we show how the model can be applied to investment problems such as capital budgeting and capital allocation. An extensive literature in finance studies capital budgeting with division managers who have empire building preferences.[6] Following Harris and Raviv (1996), we assume managers enjoy a private benefit from resource allocation in their divisions. In particular, assume the manager's private benefit of allocation ι_t is $b\iota_t$, so the manager's payoff flow at time t is:[7]

$$\pi_t = b\iota_t - ka_t.$$

The division cash-flows follow a compound Poisson process $(Y_t)_{t\geq 0}$ given by

$$Y_t = \sum_{i=1}^{N_t} \theta_{t_i}\iota_{t_i},$$

where N_t is a Poisson process with intensity μ. At each time t, the CFO decides how much resources to allocate to the division where the flow cost of those resources is $\iota^2/2$. Thus, the expected profit flow of the division at time t is $f(\iota_t, \theta_t) = E_t\left[\mu\theta_t\iota_t - \iota_t^2/2\right]$.

The optimal strategy of the CFO is to allocate:

$$\iota_t = \arg\max_{\iota} E[\mu\theta_t\iota - \iota^2/2|\mathcal{F}^M_{t^-}],$$

resources to the division.[8] The solution to this maximization problem is $\iota_t = \mu x_t$ and the principal's expected payoff is

$$u(x_t) = \frac{\mu^2}{2}x_t^2.$$

The manager's reduced-form expected payoff flow given a reputation x_t is $\pi_t = b\mu x_t - ka_t$, as in our general model. In the general model we assume that monitoring of the division is the only source of information about θ. In this application it is natural to assume that when the cash-flows arrive the CFO also learns about the current productivity. After we analyze our model we consider also an extension to such exogenous news (in addition to the endogenous monitoring).

As in the rest of the paper, we ignore in this example the use of monetary transfers in the provision of incentives beyond possibly the transfers proportional to the current allocation and reputation[9]. In some settings, other forms of performance-based compensation are used as an effective tool. But in many cases divisional contracts are simple and earnings proportional to the size of the division, may be the main driver of manager's incentives. Graham, Harvey, and Puri (2015) find evidence that manager's reputation has an

[6] Some examples are found in Hart and Moore (1995), Harris and Raviv (1996), Stein (1997) and Harris and Raviv (1998).

[7] Coefficient b can be also interpreted as incentive pay that is proportional to the size of the allocation to prevent other agency problems, such as cash diversion, not captured explicitly by our model.

[8] Note that the allocation in period t is made before the realization of the cash-flow (the Poisson process), as captured by $\mathcal{F}^M_{t^-}$. Technically, we could write that profits depend on ι_{t^-}, but write simply ι_t since the timing of the game should be well understood.

[9] See Motta (2003) for a capital budgeting model driven by career concerns along these lines.

important role in the internal capital allocation. The use of career concerns as the main incentive device also captures the allocation of resources in bureaucracies as in Dewatripont, Jewitt, and Tirole (1999). The role of financial incentives in government agencies is much more limited than in private firms. Autonomy, control and capital allocation driven by career concerns are more preponderant for worker's motivation.

3 Optimal Monitoring without Agency Problems.

To characterize the optimal monitoring policy in the general model we start with a relaxed problem, assuming away incentive constraints. When information has sufficient direct value ($u(\cdot)$ is sufficiently convex), cost of monitoring is small enough and cost of effort is small enough, the optimal policy in the relaxed problem automatically satisfies the incentive constraints. In that case our solution in this section becomes the optimal solution (it also describes the optimal policy when effort – but not quality – is observable).

Consider the evolution of reputation. Between inspection dates, given that the firm is expected to be putting full effort, $a = \bar{a}$, reputation evolves according to

$$\dot{x}_t = \lambda(\bar{a}_t - x_t). \tag{1}$$

Therefore, reputation at time $T_{n-1} + t < T_n$, given $\theta_{T_{n-1}} = \theta$, is given by

$$x_t^\theta = \theta e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right).$$

In the relaxed problem (ignoring incentive constraints) the principal solves the following stochastic control problem

$$U(x_0) = \sup_{(T_n)_{n\geq 1}} E\left[\int_0^\infty e^{-rt} u(x_t)\mathrm{d}t - \sum e^{-rT_n} c \middle| \mathcal{F}_0^M\right] \tag{2}$$

$$\text{subject to: } \dot{x}_t = \lambda(\bar{a} - x_t), \ \ x_{T_{n-1}} = \theta_{T_{n-1}}$$

The optimal policy is Markovian, with reputation being the state variable. If we let $\mathcal{A}$ be the set of reputations that lead to immediate inspection, then the value function solves the Hamilton-Jacobi-Bellman (HJB) equation

$$rU(x) = u(x) + \lambda(\bar{a} - x)U'(x), \ \ x \notin \mathcal{A} \tag{3a}$$

$$U(x) = xU(1) + (1 - x)U(0) - c, \ \ x \in \mathcal{A}. \tag{3b}$$

We conjecture and verify that the optimal policy is given by an audit set $\mathcal{A} = [\underline{x}, \overline{x}]$, where $\underline{x} \leq \bar{a} \leq \overline{x}$ and the thresholds satisfy boundary conditions:

$$U(\overline{x}) = \overline{x}U(1) + (1 - \overline{x})U(0) - c \tag{4a}$$

$$U(\underline{x}) = \underline{x}U(1) + (1 - \underline{x})U(0) - c \tag{4b}$$

$$U'(\overline{x}) = U(1) - U(0) \tag{4c}$$

$$U'(\underline{x}) = U(1) - U(0) \tag{4d}$$

Given boundary values, we can solve the HJB in closed form to get

$$U(x) = \left(\frac{1-\bar{a}}{x-\bar{a}}\right)^{\frac{r}{\lambda}} U(1) - \int_x^1 \left(\frac{y-\bar{a}}{x-\bar{a}}\right)^{\frac{r}{\lambda}} \frac{u(y)}{\lambda(y-\bar{a})} dy, \;\; x \geq \overline{x}$$
$$U(x) = \left(\frac{\bar{a}}{\bar{a}-x}\right)^{\frac{r}{\lambda}} U(0) - \int_0^x \left(\frac{\bar{a}-y}{\bar{a}-x}\right)^{\frac{r}{\lambda}} \frac{u(y)}{\lambda(\bar{a}-y)} dy, \;\; x \leq \underline{x}$$

Based on this solution we can find the boundary values $U(0)$, $U(1)$ and thresholds $\{\underline{x}, \overline{x}\}$ using the value-matching and smooth-pasting conditions.

Result (Benchmark)**.** *Suppose that U is a function satisfying the HJB equation* (3a)-(3b) *together with the boundary conditions* (4a)-(4d)*. Then U is the value function of the optimization problem* (2) *and the optimal policy is to monitor whenever $x_t \in \mathcal{A} = [\underline{x}, \overline{x}]$.*

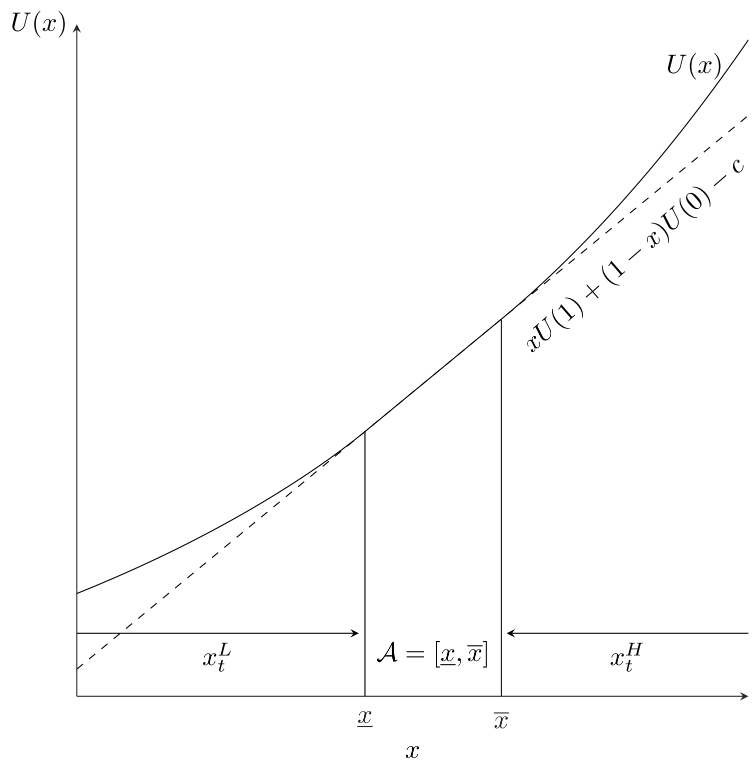


Figure 1: Value Function. The optimal policy requires to monitor whenever $x_t \in \mathcal{A}$.

Figure 1 illustrates the principal's value function (as a function of public beliefs). Observe that after a review is conducted, beliefs reset to either $x = 0$ or $x = 1$ because reviews are fully informative. Then, beliefs begin to drift deterministically toward $\bar{a}$, which lies in the interior of the audit set $\mathcal{A}$. When beliefs hit the boundary of $\mathcal{A}$, the principal monitors the firm for certain. Naturally, the principal acquires information when enough uncertainty has accumulated, namely when the distance between $U(x)$ and the line connecting $U(0)$ and $U(1)$ gets large and when beliefs get close to $\bar{a}$, so that the drift in beliefs becomes small.

The size of the monitoring region $\mathcal{A}$ depends on the convexity of the principal's objective function and the cost of monitoring c, since these parameters capture the value and cost of information, respectively. In the extreme case when $u(\cdot)$ is linear (or c is large relative to the convexity of $u(\cdot)$), the optimal policy is to never monitor the firm and let beliefs converge to $\bar{a}$ (recall we are assuming away incentive constraints – but of course in this case the incentive constraint would be violated since in this model there are no rewards to effort in the absence of information). As $u(\cdot)$ becomes more convex, the monitoring region widens leading to increased monitoring frequency. In some cases this leads to the incentive constraint being always slack, which implies the monitoring policy described above remains optimal in the full problem.

Figure 1 illustrates the optimal policy as a function of beliefs. Notice that between inspection dates beliefs evolve deterministically and monotonically over time, hence there is an equivalent representation of the monitoring policy based on both the time elapsed since last monitoring date, $t - T_{n-1}$, and the outcome observed in the last review, $\theta_{T_{n-1}}$. Specifically, define:

$$\tau_H \equiv \inf\{t : x_t = \overline{x}, x_0 = 1\} = \frac{1}{\lambda} \log\left(\frac{1-\bar{a}}{\overline{x}-\bar{a}}\right)$$
$$\tau_L \equiv \inf\{t : x_t = \underline{x}, x_0 = 0\} = \frac{1}{\lambda} \log\left(\frac{\bar{a}}{\bar{a}-\underline{x}}\right).$$

We can then represent the policy by the $n_{th}-$monitoring time as $T_n = T_{n-1} + \tau_{\theta_{T_{n-1}}}$.[10]

4 Optimal Monitoring with Agency Problems

The monitoring policy we found in the previous section may violate incentive constraints when reviews are too infrequent, leading the firm to shirk in some time intervals. The agency problem is the worst right after an inspection. In this section we solve the constrained problem that explicitly incorporates incentive constraints. As we discussed above, we study the optimal monitoring policy among those that implement full effort.

4.1 Incentive Compatibility

Our first step (towards finding an optimal monitoring policy) is to characterize necessary and sufficient conditions for a monitoring policy to satisfy the incentive compatibility constraints in every instance.

Consider the firm's payoffs given full effort from time T onward, where T is the next review date.

$$\Pi_T = E^{\bar{a}}\left[\int_T^\infty e^{-r(t-T)}(x_t - k\bar{a})\mathrm{d}t \Big| \mathcal{F}_T^\theta\right]$$
$$= \int_T^\infty e^{-r(t-T)}\left(E^{\bar{a}}[x_t|\mathcal{F}_T^\theta] - k\bar{a}\right)\mathrm{d}t.$$

which represents the present value of the firm future revenues net of effort costs. The law of iterated expectations and the Markov property of the evolution of quality imply that $E^{\bar{a}}[x_t|\mathcal{F}_T^\theta] = E^{\bar{a}}[\theta_t|\theta_T]$, where

$$E^{\bar{a}}[\theta_t|\theta_T] = \theta_T e^{-\lambda(t-T)} + \bar{a}\left(1 - e^{-\lambda(t-T)}\right).$$

This means that, in any incentive-compatible monitoring policy, conditional on quality at time T, the firm's

[10]The only exception would be the case when $x_0 \in (0,1)$. In this case $T_1 = \frac{1}{\lambda}\log\left(\frac{x_0-\bar{a}}{\overline{x}-\bar{a}}\right)$ if $x_0 > \overline{x}$; $T_1 = \frac{1}{\lambda}\log\left(\frac{x_0-\bar{a}}{\underline{x}-\bar{a}}\right)$ if $x_0 < \underline{x}$ and $T_1 = 0$ otherwise. After T_1, the policy would be the one described in the text.

continuation value at the review date T is:

$$\Pi(\theta_T) \equiv \frac{\theta_T - \bar{a}}{r+\lambda} + \frac{\bar{a}(1-k)}{r}. \tag{5}$$

Because this conditional continuation value at time T is independent of the history of effort (depends only on effort indirectly via θ_T), we can use the single shot deviation principle to derive the incentive compatibility constraint. For any effort strategy a_t, we can write the process for quality as

$$\theta_t = e^{-\lambda t}\theta_0 + \int_0^t e^{-\lambda(t-s)}(\lambda a_s \mathrm{d}s + \mathrm{d}Z_s),$$

where Z_t is a martingale. The effect of effort on future quality is $\partial\theta_T/\partial a_t = \lambda e^{-\lambda(T-t)}\mathrm{d}t$ so the marginal benefit of effort is $E_t[\lambda e^{-(r+\lambda)(T-t)}](\Pi(H) - \Pi(L))\mathrm{d}t$. This is intuitive: having high quality (rather than low quality) yields gain $\Pi(H) - \Pi(L)$ at the review time. A marginal increase in effort leads to higher quality with probability (flow) λdt. However, to reap the benefits of high quality the firm must wait till the review date, T, facing the risk of an interim drop in quality. Hence, the benefit of having high quality must be discounted according to the interest rate r and the quality depreciation rate λ.

On the other hand, the marginal cost of effort is simply $k\mathrm{d}t$. Combining these observations we can express the incentive compatibility condition for full effort as follows.

Lemma 1. *Let $n = \inf\{n : T_n > t\}$. Full effort is incentive compatible if and only if for all $t \in [T_{n-1}, T_n]$*

$$\frac{1}{r+\lambda} E_t\left[e^{-(r+\lambda)(T_n - t)}\right] \geq \frac{k}{\lambda}.$$

This condition is simple. In essence, it says that for a policy to be incentive compatible the expected next review must be sufficiently close.

Remark 1. *Lemma 1 holds for any quality process satisfying the stochastic differential equation*

$$\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \mathrm{d}Z_t,$$

where Z_t is a martingale. For example, it holds when Z_t is a Brownian motion; in this case, quality follows an Ornstein-Uhlenbeck process.

In order to characterize incentive compatible policies, it is useful to define a state variable

$$q_t = E_t\left[e^{-(r+\lambda)(T_n - t)}\right],$$

where the expectation is taken over the possibly random next monitoring time, T_n. Also let

$$\underline{q} \equiv (r+\lambda)\frac{k}{\lambda}.$$

Then the incentive compatibility constraint in Lemma 1 can be expressed simply as $q_t \geq \underline{q}$.

Variable q_t is the implicit discount rate the firm uses to assess the benefit of having high quality on the review date. It depends upon on the exogenous r and λ and on the endogenous distribution of monitoring times.

We now derive the law of motion of $(q_t)_{t\geq 0}$ to use it as a state variable in the principal's optimization problem. Some notation is in order. Let $F(t) = \Pr(T_n - T_{n-1} < t | \mathcal{F}^M_{T_{n-1}})$ denote the cumulative density of the n-th inspection conditional on the information observed up to time T_{n-1}. Let $\tau = \inf\{t > 0 : F(t) = 1\}$ be the deterministic review time at which the firm is monitored for sure if it has not been monitored before. At any point of continuity, the hazard rate of $F(\cdot)$ is given by $\mathrm{d}F(s)/(1-F(t)) = m_s$ for some function $m : [0,\infty) \to [0,\infty)$. At any point of discontinuity we have that $\mathrm{d}F(s)/(1-F(t)) = \frac{F(s)-F(s^-)}{1-F(t)}$. Accordingly, we can write q_t as

$$q_t = \int_t^\tau e^{-(r+\lambda)(s-t)-\int_t^s m_u \mathrm{d}u} m_s \mathrm{d}s + \sum_{s\in(t,\tau)} e^{-(r+\lambda)(s-t)} \frac{\Delta F(s)}{1-F(t^-)} + e^{-(r+\lambda)(\tau-t)} \frac{1-F(\tau)}{1-F(t)}$$

Whenever $F(t)$ is absolutely continuous we can differentiate the previous equation to get the following differential equation for q_t :

$$\dot{q}_t = (r+\lambda+m_t)q_t - m_t. \tag{6}$$

At any point of discontinuity, we have

$$q_{t^-} = \frac{\Delta F(t)}{1-F(t^-)} + \left(1 - \frac{\Delta F(t)}{1-F(t^-)}\right) q_t.$$

Hence, if we let p_t denote the probability of monitoring at t, conditional on not having monitored prior to t, then, if there is no monitoring at time t, the state q_t experiences a downward jump of

$$\Delta q_t = -\frac{p_t}{1-p_t}(1-q_{t^-}). \tag{7}$$

In order to satisfy incentive compatibility right after t we must have that $p_t \leq (q_{t^-} - \underline{q})/(1-\underline{q})$ (otherwise we would get $q_{t^+} < \underline{q}$).

The cumulative distribution $F(t)$ can be then expressed by the triple (m_t, p_t, τ) as

$$1 - F(t) = \begin{cases} e^{-\int_0^t m_s \mathrm{d}s} \prod_{0<s\leq t}(1-p_s) & \text{if } t < \tau \\ 0 & \text{if } t \geq \tau \end{cases}. \tag{8}$$

Combining equations (6) and (7) we get the following representation for incentive compatible monitoring policies.

Proposition 1 (Incentive Compatibility)**.** *Consider a monitoring policy described by (m_t, p_t, τ). For any interval of time $(T_{n-1}, T_{n-1}+\tau]$, let q_t be the solution to*

$$q_t = q_{T_{n-1}^+} + \int_{T_{n-1}}^t [(r+\lambda+m_s)q_s - m_s]\mathrm{d}s - \sum_{T_{n-1}<s\leq t} \frac{p_s}{1-p_s}(1-q_{s^-}),\ q_{T_{n-1}+\tau} = 1. \tag{9}$$

Full effort $a_t = \bar{a}$ is incentive compatible if and only if $q_t \geq \underline{q}$, all $t \geq 0$, where

$$\underline{q} \equiv (r+\lambda)\frac{k}{\lambda}.$$

Incentive compatibility is formulated as a lower bound on q_t or an upper bound on the expected next review date. Notice that what matters for incentives in any given period is not necessarily the monitoring intensity in that moment but the cumulative discounted likelihood of monitoring in the future: the reason is that effort has persistent effect on quality.

As we shall see, the significance of Proposition 1 is that it allows us to formulate the optimal monitoring policy as a recursive problem, with q_t being the state variable, and use the tools of optimal control theory to characterize the policy.

4.2 Principal's Problem

We now develop a recursive formulation for the principal's optimization problem that we can study using tools from the literature on optimal control theory.

The principal solves the following optimization problem:

$$U(x_0) = \sup_{(T_n)_{n\geq 1}} E\left[\int_0^\infty e^{-rs}u(x_s)\mathrm{d}s - \sum_{T_n} e^{-rT_n}c\Big|\mathcal{F}_0^M\right] \tag{10}$$

$$\text{subject to:} \tag{11}$$

$$\dot{x}_t = \lambda(\bar{a} - x_t), \;\; x_{T_{n-1}} = \theta_{T_{n-1}}$$

$$\underline{q} \leq E_t\left[e^{-(r+\lambda)(T_n - t)}\right], \;\forall t.$$

Note that the only difference between this problem and the problem in Section 3 is the introduction of the incentive constraint.

In Section 3 we have found the optimal policy working on the space of beliefs: that is, we specified the monitoring policy as a function of beliefs. As we discussed above, alternatively, we can specify the policy as a function of the outcome of the last inspection and the time elapsed since then. This alternative formulation turns out to be more convenient in this section.

In order to specify the problem in this way, let $x_t^\theta \equiv \theta e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right)$ be the principal's beliefs given that the quality observed in the last inspection was θ and the time elapsed since that inspection is t. For any $t < \tau$, we can write the survival function of the time of an inspection as $1 - F(t) = e^{-M_t}$, where

$$M_t = \int_0^t m_s \mathrm{d}s - \sum_{s\leq t}\log(1 - p_s).$$

In addition, let's denote the expected payoff to the principal from inspecting quality at time t by $\mathcal{M}_\theta(U, x_t) \equiv x_t^\theta U_H + (1 - x_t^\theta)U_L - c$, where $U = (U_L, U_H)$ is the continuation payoff under optimal policy upon learning that quality is low or high. Using the recursive formulation of the incentive compatibility constraint in

Proposition 1, we get the following recursive formulation for the principal's problem in (10)

$$\mathscr{G}^{\theta}(U)=\sup_{(m_t,p_t)_{t\in[0,\tau]},\tau,q_0}\int_0^{\tau}e^{-rt-M_{t^-}}\left(u(x_t^{\theta})+m_t\mathcal{M}_{\theta}(U,x_t)\right)\mathrm{d}t+\sum_{s<\tau}e^{-rs-M_{s^-}}p_s\mathcal{M}_{\theta}(U,x_s^{\theta})+e^{-r\tau-M_{\tau}}\mathcal{M}_{\theta}(U,x_{\tau})$$

subject to

$$M_t=\int_0^t m_s\mathrm{d}s-\sum_{s\le t}\log(1-p_s)$$

$$q_t=q_0+\int_0^t[(r+\lambda+m_s)q_s-m_s]\mathrm{d}s-\sum_{s\le t}\frac{p_s}{1-p_s}(1-q_{s^-}),\ q_{\tau}=1$$

$$q_t\in[\underline{q},1],\ \forall t\in[0,\tau]$$

$$0\le m_t$$

$$p_t\in\left[0,\frac{q_{t^-}-\underline{q}}{1-\underline{q}}\right].$$

For a fixed continuation payoff U, this is an optimal control problem with state constraints: the controls are the monitoring intensity m_t, and the probability of monitoring p_t, which captures the possibility of atoms in the monitoring CDF (the introduction of the review date τ is useful in terms of the exposition but formally unnecessary as we are allowing for jumps in the CDF). The magnitude of these jumps is restricted by the constraint that $q_t\ge\underline{q}$ just after the jump.

The solution of the principal's problem is given by the fixed point $(\mathscr{G}^L(U),\mathscr{G}^H(U))=U$. The following technical lemma establishes the existence of such a fixed point.

Lemma 2. *The principal's expected payoff is given by the unique fixed point $(\mathscr{G}^L(U),\mathscr{G}^H(U))=U$.*

We can analyze the previous problem using the theory of optimal control with state constraints. The main idea is to attach a Lagrange multiplier ψ_t to the incentive compatibility constraint $q_t\ge\underline{q}$.[11] It is also convenient to reformulate the problem using the principal's continuation value, U_t, as a state variable (where with some abuse of notation we omit the dependence on θ_0). At any point of continuity, the principal's continuation value satisfies the differential equation

$$\dot{U}_t=(r+m_t)U_t-u(x_t^{\theta})-m_t\mathcal{M}(U,x_t^{\theta}). \tag{12}$$

If there is an atom in the monitoring distribution at time t, then the continuation value satisfies

$$U_{t^-}=p_t\mathcal{M}(U,x_t^{\theta})+(1-p_t)U_{t^+}. \tag{13}$$

If $\tau<\infty$, the following terminal condition must be satisfied $U_{\tau}=\mathcal{M}(U,x_{\tau}^{\theta})$. Finally, when $\tau=\infty$ the following transversality condition must be satisfied

$$\lim_{t\to\infty}e^{-rt-M_t}U_t=0. \tag{14}$$

A control $\tilde{m}_t$ is admissible if and only if condition (14) is satisfied. Using the continuation value of the principal as a state variables, we can formulate the problem as the following optimal control problem in

[11] This method is usually referred as the direct adjoining approach (Hartl et al., 1995).

Mayer form (Cesari, 2012).[12] The benefit of this alternative formulation is that the Hamiltonian is concave, where the Hamiltonian of this problem is

$$\mathcal{H}(q_t,\zeta_t,\nu_t,\psi_t,m_t,t)=\zeta_t\big((r+m_t)U_t-u(x_t^\theta)-m_t\mathcal{M}(U,x_t^\theta)\big)+\nu_t\left((r+\lambda+m_t)q_t-m_t\right)+\psi_t\left(q_t-\underline{q}\right).$$

ζ_t is the (current value) co-state variable associated to U_t, and ν_t is the (current value) co-state variable associated to state variable q_t.[13] According to Pontryagin's maximum principle, the evolution of the co-state variables is given by

$$\dot{\zeta}_t=(r+m_t)\zeta_t-(r+m_t)\zeta_t=0,\ \zeta_0=-1. \tag{15a}$$

$$\dot{\nu}_t=-\lambda\nu_t-\psi_t. \tag{15b}$$

From here, it is immediate that $\zeta_t=-1$. The initial value q_0 can be chosen by the principal, so the initial value of the co-state variable ν_t must satisfy the following condition

$$\nu_0\le 0 \text{ with equality if } q_0>\underline{q} \tag{16}$$

Equation (16) has a simple interpretation: ν_0 reflects the marginal effect that increasing q_0 has on the principal's expected payoff, and this means that if the IC constraint is slack then the the marginal effect must be zero while if the IC constraint is binding the marginal effect must be negative. In addition, if $\tau=\infty$ then ν_t satisfies the transversality condition

$$\lim_{t\to\infty}e^{-rt-M_t}\nu_t=0. \tag{17}$$

The Hamiltonian is linear in m_t with a coefficient given by

$$S(t)=x_t^\theta U_H+(1-x_t^\theta)U_L-c-U_t-(1-q_t)\nu_t. \tag{18}$$

Using the terminology of singular optimal control problems, we call $S(t)$ the switching function. The maximum is finite only if $S(t)\le 0$, in which case the monitoring rate is

$$m_t=\begin{cases}0 & \text{if } S(t)<0\\ [0,\infty] & \text{if } S(t)=0.\end{cases} \tag{19}$$

If the previous condition is not satisfied at some t then the optimal solution requires a jump in the monitoring distribution. Equation (18) shows the main trade-off facing the optimal policy. When the incentive constraint is slack, the principal should monitor and bear the monitoring cost only if information is sufficiently valuable. That is, only if $\mathcal{M}_\theta(U,x_t^\theta)\ge U_t$. The extra term $(1-q_t)\nu_t$ reflects the impact of moral hazard on the optimal policy. The co-state variable ν_t is negative which implies that because of the incentive compatibility constraint it could be optimal to monitor even when information has negative net value. That is, even if $\mathcal{M}_\theta(U,x_t^\theta)<U_t$. This is intuitive: the need to provide incentives forces the principal to increase the level of monitoring above and beyond his own learning needs.

The optimality conditions in (19) require that $S(t)$ vanishes whenever $m_t>0$, and this means the

[12] In Mayer form we maximize U_0 subject to (9), (12), (13), and the state constraint on q_t

[13] If $\tilde{p}_t$ is the co-state for a particular state variable, the current value co-state is defined as $p_t\equiv e^{rt+M_t}\tilde{p}_t$.

switching function must be constant in any interval of time with positive monitoring intensity. In technical terms, the switching function $S(t)$ must be constant across a singular arc, which is a standard condition in the theory of singular optimal control. Accordingly, whenever $m_t > 0$ it must be the case that $\dot{S}(t) = 0$. If we differentiate $S(t)$ and use the optimality condition $m_t S(t) = 0$ the we get

$$
\begin{aligned}
\dot{S}(t) &= \dot{x}_t^\theta (U_H - U_L) - \dot{U}_t + \dot{q}_t \nu_t - (1 - q_t)\dot{\nu}_t \\
&= u(x_t^\theta) + (r + \lambda) q_t \nu_t + (1 - q_t)(\lambda \nu_t + \psi_t) + \dot{x}_t^\theta (U_H - U_L) - rU_t \\
&= 0.
\end{aligned}
$$

This expression is useful to identify the value of the Lagrange multiplier ψ_t. As it is usual in constrained optimization problems, the Lagrange multiplier is positive only if the incentive compatibility constraint is binding so

$$
\psi_t = \begin{cases} 0 & \text{if } q_t > \underline{q} \\ \geq 0 & \text{if } q_t = \underline{q}. \end{cases} \tag{20}
$$

Hence, if the incentive compatibility constraint is binding, then we can use the condition $\dot{S}(t) = 0$ to back out the value of ψ_t

$$
\psi_t = \frac{rU_t - (r\underline{q} + \lambda)\nu_t - u(x_t^\theta) - \dot{x}_t^\theta (U_H - U_L)}{1 - \underline{q}}. \tag{21}
$$

The next step is to characterize the necessary condition that an optimal review date τ must satisfy; in particular, whenever the optimal monitoring policy has a deterministic review date, the following terminal condition must be satisfied at time τ

$$
r\mathcal{M}_\theta(U, x_\tau^\theta) = u(x_\tau^\theta) + (r + \lambda)\nu_\tau + \dot{x}_\tau^\theta (U_H - U_L). \tag{22}
$$

The final step in the characterization of the necessary conditions is to consider the possibility of atoms in the CDF. If there is an atom at time s, then the monitoring probability is given by

$$
p_s \in \arg \max_{p \in [0, (q_{s^-} - \underline{q})/(1 - \underline{q})]} \left\{ \frac{p}{1 - p} \left[\mathcal{M}_\theta(U, x_s^\theta) - U_{s^-} - \nu_{s^+}(1 - q_{s^-}) \right] \right\}, \tag{23}
$$

and so it follows that for any time t without an atom, the following inequality must hold

$$
\mathcal{M}_\theta(U, x_t^\theta) - \nu_t(1 - q_t) \leq U_t. \tag{24}
$$

Equation (24) implies that if $S(t) \leq 0$, for all t, then q_t is continuous and there is at most one atom in the monitoring distribution taking place at time τ. The following Lemma summarizes the necessary conditions of optimality.[14]

Lemma 3. *Given continuation payoffs (U_L, U_H), if the monitoring policy (m_t^*, p_t^*, τ^*) is a solution of the optimal control problem $\mathscr{G}^\theta U$ then there is a trajectory of the co-state variables (ζ_t, ν_t) satisfying* (15a)*,* (15b) *and* (16)*, and a non-negative Lagrange multiplier ψ_t such that:*

1. *The Lagrange multiplier ψ_t is given by* (21) *if $q_t = \underline{q}$ and zero otherwise.*

[14] For necessary and sufficient conditions in optimal control with state constraints see Section 4 in Hartl et al. (1995) and Chapter 5 in Seierstad and Sydsaeter (1986). See Chapter 3 Section 3 in Seierstad and Sydsaeter (1986) for an exposition of the maximum principle in the presence of jumps in the state variable.

2. *The optimal monitoring rate m_t^* is given by* (19).

3. *There is an atom p_t^* at time t only if $S(t) \geq 0$, and $S(t) \leq 0$ holds at any point of continuity.*

4. *If $\tau^* < \infty$ then condition* (22) *is satisfied.*

4.3 Linear Payoffs

We begin by analyzing the case in which the principal flow payoff $u(\cdot)$ is linear (as we discussed above, that could capture a self-regulatory organization). Under linear payoffs, information has no direct social value; the only role of monitoring is incentive provision. As we show, in this case the optimal monitoring policy prescribes a constant monitoring intensity (and no deterministic reviews).

We conjecture (and later verify) that the optimal policy prescribes monitoring if and only if the incentive compatibility constraint binds and that the incentive compatibility constraint always binds (the intuition is that otherwise the principal could save some monitoring cost without affecting the firm's incentives – we return to the intuition below).

For the incentive compatibility constraint to bind all the time we need $\tau = \infty$, so effectively there is no final review date. This claim is true because if $\tau < \infty$ then the incentive compatibility condition would be slack somewhere close to τ.

We can find the optimal monitoring rate at any time t by solving the incentive constraint. Using the law of motion for q_t together with the condition $q_t = \underline{q}$ for all t, we find that $m_t = m^* = (r+\lambda)\underline{q}/(1-\underline{q})$.

The principal's payoff under this policy satisfies:

$$U_H - U_L = \int_t^\infty e^{-r(s-t)} x_s^H \mathrm{d}s - \int_t^\infty e^{-r(s-t)} x_s^L \mathrm{d}s = \frac{1}{r+\lambda}.$$

Let $U_\theta(t)$ be the principal's expected payoff at time t given $\theta_0 = \theta$, so by definition $U_\theta = U_\theta(0)$. In the case that the principal's payoff flow is linear in beliefs and the monitoring rate is constant, the principal's payoff is:

$$U_\theta(t) = \frac{x_t^\theta - \bar{a}}{r+\lambda} + \frac{\bar{a}}{r} - \frac{m^* c}{r}. \tag{25}$$

In order to verify that this policy is optimal, we need to very that the Lagrange multiplier ψ_t in equation (21) is non-negative, which requires that we verify the following inequality:

$$rU_\theta(t) \geq x_t^\theta + (r\underline{q} + \lambda)\nu_t + \dot{x}_t^\theta (U_H - U_L).$$

In the appendix, we verify that this inequality is satisfied for all $t \geq 0$, which yields our result:

Proposition 2. *If $u(x_t)$ is affine, the optimal monitoring policy is given by a Poisson process with arrival rate $m^* = (r+\lambda)\underline{q}/(1-\underline{q})$.*

In sum, the optimal policy is purely random, there are no periodic reviews. Moreover, the optimal policy features a constant monitoring rate that is independent of the history, in particular, of the outcome of the last review. There is no extra scrutiny if the firm quality was low in the last review.[15]

We can gain some intuition comparing the optimal policy to the best policy with deterministic monitoring. When the principal's flow payoffs are linear in beliefs, the optimal policy is the one that minimizes the cost

[15] As we show in Section 5, this independence result can break if we introduce exogenous news about quality.

of monitoring subject to the incentive compatibility constraint. The best deterministic policy is clearly the one with the least amount of monitoring, and this corresponds to the policy that makes the incentive compatibility constraint binding at $t = 0$. This policy requires that the principal monitors at a time τ^{bind} given by

$$\tau^{\text{bind}} = -\frac{\log \underline{q}}{r + \lambda}.$$

The present value of the monitoring cost of this policy is

$$\mathcal{C}^{\text{det}} = \frac{e^{-r\tau^{\text{bind}}}}{1 - e^{-r\tau^{\text{bind}}}} c.$$

On the other hand, the optimal policy is such the monitoring time $\tilde{\tau}^{\text{rand}}$ has an exponential distribution with arrival rate m^*, where m^* is chosen such the IC constraint binds, that is

$$E[e^{-(r+\lambda)\tilde{\tau}^{\text{rand}}}] = \frac{m^*}{r + \lambda + m^*} = \underline{q}.$$

The present value of the monitoring cost in this case is

$$\mathcal{C}^{\text{rand}} = \frac{E[e^{-r\tilde{\tau}^{\text{rand}}}]}{1 - E[e^{-r\tilde{\tau}^{\text{rand}}}]} c,$$

and so the random policy is cheaper if and only if $E[e^{-r\tilde{\tau}^{\text{rand}}}] < e^{-r\tau^{\text{bind}}}$. Replacing τ^{bind}, and computing the term $E[e^{-r\tilde{\tau}^{\text{rand}}}]$ using the fact that $\tilde{\tau}^{\text{rand}}$ has an exponential distribution with parameter m^*, we find that the previous inequality holds if and only if

$$\underline{q}^{\frac{r}{r+\lambda}} > \frac{\underline{q}}{\underline{q} + \frac{r}{r+\lambda}(1 - \underline{q})}.$$

It can be readily verified that the previous inequality is satisfied for any $\underline{q} \in (0, 1)$, and moreover the difference between these two terms is mainly driven by the persistence of the shocks to quality λ. In other words, it is better to use a monitoring policy that smooths incentives over time (so the IC constraints bind at all times) rather than a policy that back-loads the incentives to exert effort.

The intuition is as follows. The policy τ^{bind} implies incentive constraint binds at $t = 0$ and are relaxed in all times $(0, \tau)$. Because quality is transitory, the inspection at τ^{bind} provides a very noisy indicator of $t = 0$ effort. We can improve upon that policy by reducing the probability of monitoring at τ^{bind} and adding some probability of monitoring at time close to $t = 0$, in a way that keeps incentives still slack at $t > 0$ and just satisfied around $t = 0$. This change reduces total costs because the early monitoring is based on a less noisy signal of early effort, so the increase can be much smaller than the reduction of monitoring probability at τ^{bind}. Generalizing this reasoning implies that we can reduce the total expected cost of monitoring by distributing it evenly over time so that the incentive constraints bind in every period.

4.4 Socially Valuable Information

Next, we study the case in which the flow payoff of the principal $u(\cdot)$ is convex in beliefs, so information has social value. First, we establish that the monitoring rate m_t is positive if and only if the incentive compatibility constraint binds.

Lemma 4. *Let (m^*_t, p^*_t, τ^*) be an optimal policy, then $q_t > \underline{q} \Rightarrow m^*_t = 0$.*

Notice this does not mean there will not be monitoring when the incentive constraint is slack. Rather it means that at any point where the incentive constraint is slack, there is either no monitoring at all or there is an atom in the distribution of monitoring; but both cases require m^*_t=0.

Next, we show there cannot be atoms in the distribution of monitoring prior to the final review date τ.

Lemma 5. *Let (m^*_t, p^*_t, τ^*) be an optimal policy, then $p^*_s = 0$ for all $s \in [0, \tau)$.*

Because the optimal policy entails no jumps before τ^*, Lemma 4 implies that q_t is non-decreasing in time. Hence, if $m^*_t = 0$, then $m^*_s = 0$ for all $s \in [t, \tau]$.

This means the optimal policy must be such that: 1) $m^*_t = m^*$ for some (possibly empty) time interval $[0, \hat{\tau}^*]$, and 2) $m^*_t = 0$ on the interval $(\hat{\tau}^*, \tau^*]$. With pure deterministic monitoring $\hat{\tau}^* = 0$ while in the case with pure random monitoring $\tau^* = \infty$. Moreover, whenever $0 < \hat{\tau}^* < \tau^* < \infty$ we have $q_t = e^{-(r+\lambda)(\tau - t)}$ for any $t \in [\hat{\tau}^*, \tau^*]$ and $q_{\hat{\tau}^*} = \underline{q} = e^{-(r+\lambda)(\tau^* - \hat{\tau}^*)}$. Hence, $\hat{\tau}^*$ and τ^* are related by

$$\tau^* = \hat{\tau}^* + \frac{1}{r+\lambda} \log \frac{1}{\underline{q}}$$

The following proposition provides a general characterization of the optimal monitoring policy.

Proposition 3. *Suppose information has social value, or $u(\cdot)$ is strictly convex. Then, there is $(\hat{\tau}^*_\theta, \tau^*_\theta)$, $\theta \in \{L, H\}$ such that the optimal monitoring policy given $\theta_{T_{k-1}} = \theta$ is given by the distribution*

$$F_\theta(t) = \begin{cases} 1 - e^{-\min\{t, \hat{\tau}^*_\theta\} m^*} & \text{if } t < \tau^*_\theta \\ 1 & \text{if } t \geq \tau^*_\theta \end{cases}, \tag{26}$$

where

$$m^* = (r+\lambda) \frac{\underline{q}}{1 - \underline{q}}.$$

Moreover, the random monitoring threshold $\hat{\tau}^*_\theta$ is always an extreme point, that is $\hat{\tau}^*_\theta \in \{0, \infty\}$. In particular,

1. *if $\tau^*_\theta = \hat{\tau}^*_\theta = \infty$, the optimal monitoring policy following $\theta_{T_{k-1}} = \theta$ has constant Poisson monitoring at rate m^*,*
2. *and if $\hat{\tau}^*_\theta = 0$, the optimal monitoring policy following $\theta_{T_{k-1}} = \theta$ has a deterministic review date.*

The optimal policy is remarkably simple. The constant monitoring rate may seem counterintuitive because it means that beliefs about quality don't affect monitoring rates even though beliefs deteriorate over time.[16] The intuition is that given our technology, the incentive constraints are the same for all levels of beliefs (and the same for the high-quality and low-quality firm), and the random monitoring intensity is such that the incentive constraints bind at all time till the inspection.

Proposition 3 allows us to write the principal's problem as a finite-dimensional problem. Let $\mathcal{G}^\theta_{\text{det}}$ be the best incentive compatible deterministic policy given continuation payoffs $U = (U_L, U_H)$, which is given by

$$\mathcal{G}^\theta_{\text{det}}(U) \equiv \max_{\tau \in [0, \tau^{\text{bind}}]} \int_0^\tau e^{-rt} u(x^\theta_t) \mathrm{d}t + e^{-r\tau} \mathcal{M}(U, x^\theta_\tau),$$

[16]As we show in the next section, this result can change if we introduce exogenous information to the model.

where (as before) $\tau^{\text{bind}} \equiv -\log \underline{q}/(r+\lambda)$ is the minimum (deterministic) frequency consistent with incentive compatibility. And let, $\mathcal{G}^{\theta}_{\text{rand}}$ be the payoff of the random policy, which is given by:

$$\mathcal{G}^{\theta}_{\text{rand}}(U) \equiv \int_0^{\infty} e^{-(r+m^*)t} \left(u(x_t^{\theta}) + m^* \mathcal{M}(U, x_t^{\theta})\right) \mathrm{d}t.$$

The solution to the principal's problem is then given by:

$$U_{\theta} = \max\{\mathcal{G}^{\theta}_{\text{det}}(U), \mathcal{G}^{\theta}_{\text{rand}}(U)\}. \tag{27}$$

Figure 2 studies the effect of monitoring cost and effort cost on the optimal policy. The left panel shows that, as the cost of monitoring increases, the optimal policy shifts from relying on scheduled reviews to relying on random reviews. This shift reflects the fact that, as c goes up, the policy becomes more concerned with incentive provision. The right panel shows that as the cost of effort goes up, thus exacerbating the moral hazard issue, the policy relies more on random reviews, which again is the most effective way to mitigate moral hazard issues.

Some of the economic forces behind the optimal policy are similar to those found in static settings such as Lazear (2006) and Eeckhout et al. (2010). While the principal is not concerned with maximizing the fraction of complying agents (which is fixed in our model) our solution features random monitoring. Unlike static settings, here monitoring effort is spread out –not across agents or cross sectionally– but in the time-series. The idea is that, for incentive purposes, having slack incentive constraints is never optimal or else some monitoring resources could be saved or reallocated across periods. These static models may also feature concentration of monitoring efforts: namely deterministic monitoring of a few agents rather than random monitoring of all agents. In our model, the rationale behind deterministic monitoring is driven by learning, whereas in Lazear (2006) the concentration of monitoring across individuals is driven by scarcity: the impossibility of ensuring compliance by all agents calls in their setting for focalized monitoring.

4.5 Brownian Quality Shocks

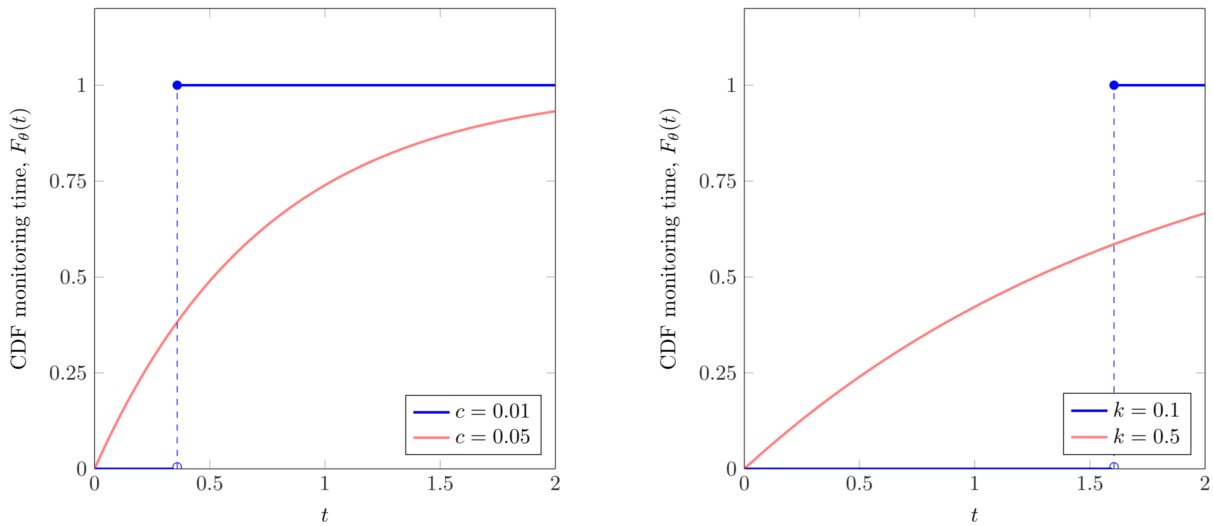
One may be concerned that our results so far are driven by the assumption that quality switches between only two levels. This binary specification makes the analysis more tractable (especially in the presence of exogenous news that we turn to in the next section), but it is our intuition that the dynamics of optimal monitoring are driven by the nature of incentives and not the details of the quality process.

To support this intuition, in this section we show that the qualitative nature of the results continues to hold when we look at another process for quality. As previously mentioned in Remark 1, the incentive compatibility characterization in Proposition 1 holds if quality follows any process such as

$$\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \sigma \mathrm{d}Z_t, \tag{28}$$

where Z_t is any martingale. In particular, it holds when Z_t is a Brownian motion so quality follows an Ornstein-Uhlenbeck process. If the principal's preferences are linear, we get that the principal's expected payoff given a monitoring policy M implementing $a_t = \bar{a}$ is just

$$E\left[\int_0^{\infty} e^{-rs}\theta_s \mathrm{d}s - \sum_{T_n} e^{-rT_n} c \middle| \mathcal{F}_0^M\right] = E\left[\int_0^{\infty} e^{-rs}x_s \mathrm{d}s - \sum_{T_n} e^{-rT_n} c \middle| \mathcal{F}_0^M\right],$$



(a) Comparative statics for c. The optimal monitoring policy is given by $\hat{\tau}^*_\theta = \{0, 0.61, \infty\}$ and $\tau^*_\theta = \{0.35, 1.15, \infty\}$. The first best policy is $\tau^{FB}_\theta = \{0.35, 1.60, 2.41\}$

(b) Comparative statics for k. The optimal monitoring policy is given by $\hat{\tau}^*_\theta = \{0, 0.61, \infty\}$ and $\tau^*_\theta = \{1.60, 1.15, \infty\}$. The first best policy is $\tau^{FB}_\theta = 1.60$

Figure 2: Comparative statics for optimal monitoring distribution. The figure shows the CDF of the monitoring time T_n when $u(x_t) = x_t - 0.5 \times x_t(1 - x_t)$ and $r = 0.1$, $\lambda = 1$, $\bar{a} = 0.5$. When c or k are low, the incentive compatibility constraint is slack under the optimal monitoring policy in the relaxed problem that ignores incentive compatibility constraints. As the monitoring or effort cost increase, deterministic monitoring is replaced by random monitoring. In this example the payoff function and the technology are symmetric so the optimal monitoring policy is independent of θ_0.

where the conditional expectation $x_t = E(\theta_t | \mathcal{F}^M_t)$ satisfies

$$\dot{x}_t = \lambda(\bar{a} - x_t), \;\; x_{T_{n-1}} = \theta_{T_{n-1}}.$$

So, with linear preferences, the principal's problem given quality process (28) is the same as the one with binary quality. Hence, the optimal monitoring policy is still given by Proposition 2. Hence, we have the following result

Corollary 1. *Suppose the principal's preferences are linear in θ_t. Then, the optimal monitoring policy is given by a Poisson process with arrival rate $m^* = (r + \lambda)\underline{q}/(1 - \underline{q})$.*

Whenever the principal's payoff is not linear in quality, we need to specify the principal preferences as a function of beliefs. With non-linear preferences, the optimal policy generally depends on the last outcome of the monitoring process (which in this case has a continuum of outcomes). This fact does not change the core economic forces but makes analysis and computation much more involved. However, the special case in which the principal preferences are linear-quadratic is particularly tractable and we can get a clean characterization of the optimal monitoring policy.[17]

Suppose that the principal has linear quadratic preferences $u(\theta_t, x_t) = \theta_t - \gamma(\theta_t - x_t)^2$. Taking conditional expectations we can write the principal's expected flow payoffs as $u(x_t, \Sigma_t) = x_t - \gamma\Sigma_t$, where

[17] The linear-quadratic case has been extensively used in several applications, for example, the rational inattention literature, because it offers high tractability.

$\Sigma_t \equiv \mathrm{Var}(\theta_t|\mathcal{F}_t^M)$. The distribution of θ_t is Gaussian with moments

$$x_t = \theta_0 e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right)$$
$$\Sigma_t = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right).$$

Moreover, using the law of iterated expectations, we can show that the principal's flow payoffs are linear in expected quality and given by

$$U(\theta) = \frac{\theta - \bar{a}}{r + \lambda} + \frac{\bar{a}}{r} - \mathcal{C},$$

where $\mathcal{C}$ solves

$$\mathcal{C} = \min_{\tau \geq 0, m_t \geq 0, q_0 \geq \underline{q}} \int_0^\tau e^{-rt - M_{t-}}\left(\gamma \Sigma_t + m_t\left(c + \mathcal{C}\right)\right)\mathrm{d}t + e^{-r\tau - M_\tau}\left(c + \mathcal{C}\right)$$

subject to

$$\dot{q}_t = (r + \lambda + m_t)q_t - m_t,\ q_\tau = 1$$
$$q_t \in [\underline{q}, 1],\ \forall t \in [0, \tau].$$

The optimal policy is now formulated recursively as a cost minimization problem, where the costs borne by the firm have two sources: monitoring and uncertainty as captured by the residual variance of quality Σ_t. As before, the principal chooses the monitoring intensity m_t , the probability of a review p_t, and the date of the scheduled review τ. The main state variable is again the expected time till next review, q_t. Using the same arguments as in the binary case, we can show that there is no monitoring if the IC constraint is slack. Hence, the optimal monitoring policy takes the same form as in the binary case. The policy is fully characterized by $(\hat{\tau}, \tau)$ solving

$$\mathcal{C} = \min_{\tau \geq 0} \frac{\int_0^\tau e^{-rt - m^*(t \wedge \hat{\tau})}(\gamma \Sigma_t + \mathbf{1}_{\{t < \hat{\tau}\}} m^* c)\mathrm{d}t + e^{-r\tau - m^* \hat{\tau}} c}{1 - \int_0^{\hat{\tau}} e^{-(r + m^*)t} m^* \mathrm{d}t - e^{-r\tau - m^* \hat{\tau}}}$$
$$\hat{\tau} = \max\left\{\tau - \frac{1}{r + \lambda}\log\frac{1}{\underline{q}}, 0\right\}.$$

Note that given the symmetry in the linear-quadratic case, the optimal policy is independent of the outcome in the last review. Moreover, we have reduced the problem to a simple one dimensional optimization problem. The next proposition shows that in the linear quadratic model with Brownian shocks the optimal policy consist on either periodic reviews or is purely random.

Proposition 4. *If the process of quality, θ_t, is given by the Ornstein-Uhlenbeck process in equation* (28)*, and the principal's expected payoff flows are $u(x_t, \Sigma_t) = x_t - \gamma \Sigma_t$, then the optimal monitoring policy is either deterministic monitoring at intervals of fixed length or purely random with a constant hazard rate (so that $\hat{\tau}^* \in \{0, \infty\}$).*

5 Exogenous News

Exogenous news such as media articles, customer reviews, academic research provide information to the market that may complement or substitute the principal's own monitoring and disclosures. On the surface, it might seem that the presence of news should substitute the principal's monitoring. Anecdotal evidence, however, is that monitoring often seems to intensify following the release of bad news.

To provide some insights about the interaction between optimal monitoring policy and exogenous news, in this section we add to our model an exogenous public news process that may reveal current quality to the market. We characterize how the relation between optimal monitoring and market beliefs depends upon the nature of the news process. When exogenous news convey negative states but are (relatively) silent about positive states, monitoring is more intense in bad times because moral hazard issues exacerbate in bad times (i.e., when beliefs are low). By contrast, when exogenous news convey positive states but are (relatively) silent about positive states, monitoring is more intense in good times (when reputation is higher), since shirking incentives become stronger under optimistic beliefs.

Formally, we assume there are two Poisson news process $(N_t^L)_{t\geq 0}$ and $(N_t^H)_{t\geq 0}$. The process N_t^L is bad news, and arrives with positive probability only if $\theta_t = L$, in which case the mean arrival rate is $\mu_L > 0$. On the other hand, N_t^H is good news and has a mean arrival rate μ_H if $\theta_t = H$ and zero otherwise. When $\mu_L = \mu_H$ we say that news are *symmetric*. In this case the absence of news conveys no information. On the other hand, when $\mu_L \neq \mu_H$ we say that news are *asymmetric*. In this case the absence of news is informative, we say that we are in the *bad news* case when $\mu_L > \mu_H$ and in the *good news* case if $\mu_H > \mu_L$. In the absence of exogenous news and monitoring, beliefs evolve according to

$$\dot{x}_t = \lambda(a_t - x_t) - (\mu_H - \mu_L)x_t(1 - x_t).$$

Thus, the exogenous news introduces a new term in the drift of reputation. That term is positive in the bad news case and negative in the good news case. In general, the market learns from the absence of news since no news are informative when the news processes have asymmetric arrival rates.

5.1 Incentive Compatibility

In the presence of exogenous news, we cannot use a single state variable to characterize incentive compatibility. With persistent state variables we need additional state variables to keep track of the continuation value across states. As in Fernandes and Phelan (2000) we use the continuation value conditional on the firm's private information (i.e., the firm quality). Let Π_t^θ be the firm's continuation value conditional on being type θ_t and define $D_t \equiv \Pi_t^H - \Pi_t^L$. The continuation value must satisfy the Bellman equations

$$r\Pi_t^H = \max_{a\in[0,\bar{a}]} \left\{ x_t - ka_t - \lambda(1-a_t)D_{t^-} + (\mu_H + m_t)(\Pi(H) - \Pi_t^H) + \dot{\Pi}_t^H \right\}$$
$$r\Pi_t^L = \max_{a\in[0,\bar{a}]} \left\{ x_t - ka_t + \lambda a_t D_{t^-} + (\mu_L + m_t)(\Pi(L) - \Pi_t^L) + \dot{\Pi}_t^L \right\},$$

where we use the fact that if $a_t = \bar{a}$ for any $t \geq T_n$ then – whenever $\theta_{T_n} \in \mathcal{F}_{T_n}^M$ – the continuation payoff is $\Pi_0^\theta = \Pi(\theta)$, where $\Pi(\theta)$ was defined in (5). From here it follows that full effort $a_t = \bar{a}$ is incentive compatible if and only if:[18]

$$D_t \geq \frac{k}{\lambda}.$$

Accordingly, given any incentive-compatible policy, the firm's continuation value satisfies the following

[18]This incentive compatibility is analogous to the one in Board and Meyer-ter Vehn (2013) with the difference being that there the only source of information is the exogenous news process and we allow for additional information from costly inspections.

differential equation (in absence of news and monitoring):

$$\dot{\Pi}_t^H = (r + \mu_H + m_t)\Pi_t^H - x_t + k\bar{a} + D_t\lambda(1-\bar{a}) - (\mu_H + m_t)\Pi(H)$$
$$\dot{\Pi}_t^L = (r + \mu_L + m_t)\Pi_t^L - x_t + k\bar{a} - D_t\lambda\bar{a} - (\mu_L + m_t)\Pi(L).$$

Taking the difference between $\dot{\Pi}_t^H$ and $\dot{\Pi}_t^L$ (and keeping $a_t = \bar{a}$) we get that D_t satisfies the following differential equation

$$\dot{D}_t = (r + \lambda + m_t)D_t - \mu_H(\Pi(H) - \Pi_t^H) + \mu_L(\Pi(L) - \Pi_t^L) - m_t\Delta.$$

where we define $\Delta \equiv \Pi(H) - \Pi(L) = 1/(r+\lambda)$. At time T_n, we have $\Pi_{T_n}^{\theta_{T_n}} = \Pi(\theta_{T_n})$ and so $D_{T_n} = \Delta$.

5.2 Principal's Problem

From the principal's viewpoint, it does not matter whether he learns the state due to monitoring or exogenous news. In either case, the problem going forward is the same. Hence, we can write the problem recursively using as state variables the time elapsed since the last time the firm type was observed (either by inspection or news), and the type observed at that time. The optimal control problem (ignoring jumps in the monitoring distribution) becomes

$$\mathscr{G}^\theta(U_L, U_H) = \sup_{\tau, m_t, \Pi_0^{\theta^c}} \int_0^\tau e^{-rt-M_{t-}}\left(u(x_t^\theta) + \mu_H x_t^\theta U_H + \mu_L(1-x_t^\theta)U_L + m_t\mathcal{M}_\theta(U, x_t)\right)\mathrm{d}t + e^{-r\tau - M_\tau}\mathcal{M}_\theta(U, x_\tau)$$

subject to

$$\dot{\Pi}_t^H = (r + \mu_H + m_t)\Pi_t^H - x_t + k\bar{a} + \lambda(1-\bar{a})(\Pi_t^H - \Pi_t^L) - (\mu_H + m_t)\Pi(H),\ \Pi_\tau^H = \Pi(H)$$
$$\dot{\Pi}_t^L = (r + \mu_L + m_t)\Pi_t^L - x_t + k\bar{a} - \lambda\bar{a}(\Pi_t^H - \Pi_t^L) - (\mu_L + m_t)\Pi(L),\ \Pi_\tau^L = \Pi(L)$$
$$\Pi_0^\theta = \Pi(\theta)$$
$$\frac{k}{\lambda} \le \Pi_t^H - \Pi_t^L,\ \forall t \in [0, \tau]$$
$$0 \le m_t.$$

Note that even though $\Pi_0^{\theta_0} = \Pi(\theta_0)$ while if $x_0 = 0$, $\Pi_0^L = \Pi(L)$ it is not true that $\Pi_0^{\neg\theta_0} = \Pi(\neg\theta_0)$ if $x_0^{\theta_0} \neq \theta_0$ because there is a divergence between the market and the firm's beliefs.

5.3 Symmetric News

First, consider the symmetric case in which $\mu_H = \mu_L = \mu$. In this case, the principal observes the current state of the firm at random times. However, because the arrival rate of this news is independent of the current state, the event no news is uninformative. In this case, we can write the evolution of D_t as

$$\dot{D}_t = (r + \mu + \lambda + m_t)D_t - (\mu + m_t)\Delta.$$

From this equation, we can immediately recover the ODE for q_t by looking at D_t/Δ

$$\dot{q}_t = (r + \mu + \lambda + m_t)q_t - \mu - m_t, \tag{29}$$

The only difference between equation (29) and the one in the case without news is that we have added μ to the monitoring rate m_t. This means we can write the principal's problem as

$$\mathscr{G}^\theta(U_L, U_H) = \sup_{\tau, m_t} \int_0^\tau e^{-rt-M_{t-}} \left(u(x_t^\theta) + \mu(x_t^\theta U_H + (1-x_t^\theta)U_L) + m_t \mathcal{M}_\theta(U, x_t)\right) \mathrm{d}t + e^{-r\tau - M_\tau}\mathcal{M}_\theta(U, x_\tau)$$

subject to

$$\dot{q}_t = (r + \mu + \lambda + m_t)q_t - \mu - m_t$$
$$q_t \geq \underline{q}, \ \forall t \in [0, \tau]$$
$$0 \leq m_t.$$

This problem is exactly the same as the one without news, with the exception that now the principal gets some monitoring μ *for free*. When news are symmetric, exogenous news are a perfect substitute of monitoring. As before, if the incentive compatibility constraint is binding on an interval $[t_0, t_1]$, then it must be that $\dot{q}_t = 0$ on this interval. Accordingly, the monitoring rate must be

$$m^* + \mu = (r + \lambda)\frac{\underline{q}}{1 - \underline{q}}.$$

Clearly, the monitoring rate is positive only if μ is low enough. Otherwise, exogenous news suffice for incentive purposes making it unnecessary for the principal to monitor the firm. We can think of this case as arising when the scrutiny performed by customers and market pundits is enough to discipline the firm.

This problem is analogous to that without news except that the principal gets information at rate μ at no cost. Depending on the magnitude of μ the optimal monitoring policy may entail some or no random monitoring. We have the following proposition which is a direct implication of Proposition 3.

Proposition 5. *If $(r + \lambda)\frac{\underline{q}}{1-\underline{q}} \geq \mu$ then the optimal monitoring policy takes the same form as the one in Propositions 2 and 3 with the Poisson monitoring rate now given by*

$$m^* = (r + \lambda)\frac{\underline{q}}{1 - \underline{q}} - \mu.$$

If $(r+\lambda)\frac{\underline{q}}{1-\underline{q}} < \mu$ then there is no monitoring if $u(\cdot)$ is affine, and either no monitoring or only deterministic monitoring if $u(\cdot)$ is strictly convex.

In the case of symmetric news, the optimal policy is qualitatively the same as that without news. The monitoring rate is constant and insensitive to exogenous news. The principal's optimal monitoring policy continues to be independent of public beliefs about the firm quality.

5.4 Asymmetric News

Next, consider the asymmetric case, $\mu_H \neq \mu_L$, so that the intensity of news arrival depends on firm's quality. Such asymmetry seems natural. In some industries and under some market conditions good news tend to be revealed faster than bad news, among other things because firms themselves may delay the release of bad news (as some literature on voluntary disclosure seems to document, see for example Kothari et al. (2009)). Sometimes bad news tend to be revealed faster than good news, perhaps because news agencies and TV broadcast face stronger demand for bad news stories (see Trussler and Soroka (2014)).

The main question we address here is how monitoring rates are affected by reputation when exogenous news are asymmetric. For tractability, we restrict attention to the case with linear preferences, $u(x_t) = x_t$. Moreover, given the insights from above, we conjecture that the optimal policy has pure random monitoring (that is, $\tau = \infty$) and the monitoring rate is positive only if the incentive compatibility constraint is binding. In summary, we look into optimal policies that have the following properties 1) $\tau = \infty$, and 2) $m_t > 0$ only if the incentive compatibility constraint is binding, that is if $\Pi_t^H - \Pi_t^L = k/\lambda$. We conjuncture that for linear $u(\cdot)$ this is indeed an optimal monitoring policy, at least for a wide range of parameters.

Given this conjecture, as before, the monitoring rate can be obtained from $(\dot{\Pi}_t^H - \dot{\Pi}_t^L) = 0$ and $\Pi_t^H - \Pi_t^L = k/\lambda$. Using these conditions, the monitoring rate boils down to:

$$m_t = \alpha + \beta \Pi_t^L, \tag{30}$$

where

$$\alpha = \frac{(r+\lambda)k/\lambda + \mu_H(k/\lambda - \Pi(H)) + \mu_L \Pi(L)}{\Delta - k/\lambda}$$
$$\beta = \frac{\mu_H - \mu_L}{\Delta - k/\lambda}.$$

The constant β is positive in the good news case and negative otherwise. Hence, in the bad news case, this expression is positive only if $\Pi_t^L \leq -\alpha/\beta$ while in the good news case is positive only if $\Pi_t^L \geq -\alpha/\beta$. As in Board and Meyer-ter Vehn (2013) this follows from the fact that the effort incentives increase in reputation in the bad news case and decrease in reputation in the good news case.

We follow the same approach as that in Section 4.3: we first compute the conjectured policy and then use the maximum principle to verify it. We relegate the details of the analysis to the appendix.

We focus on the simplest case in which $m_t > 0$ all $t \geq 0$ because it illustrates the effect of introducing exogenous news on the optimal monitoring policy at the lowest cost of technical complications. Such policy is optimal when the rates of exogenous news arrivals are low.[19]

Using the relation $\Pi_t^H = \Pi_t^L + D_t = \Pi_t^L + k/\lambda$ and the monitoring rate (30) we write the evolution of the low quality firm continuation value as

$$\dot{\Pi}_t^L = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi_t^L + \beta(\Pi_t^L)^2 - x_t. \tag{31}$$

If $\theta_0 = L$ then the initial condition is $\Pi_0^L = \Pi(L)$. If $\theta_0 = H$ (and the incentive compatibility is binding) the initial condition is $\Pi_0^L = \Pi(H) - k/\lambda$.[20] Once we have computed the candidate optimal policy, we can use the sufficient conditions of the maximum principle to verify ex-post that our conjectured policy is optimal (see Section D.1 in the appendix).

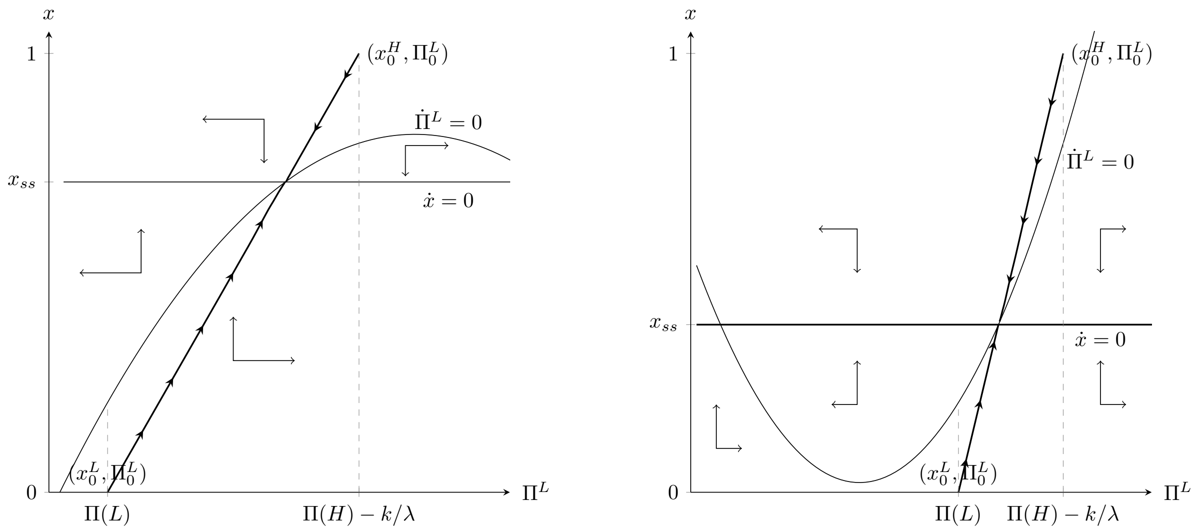
We analyze the evolution of monitoring by studying the phase diagram in the space (x_t, Π_t^L) in Figure 3.

Using the ODE for Π_t^L in equation(31) we get a quadratic equation for the steady state.

$$0 = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi^L + \beta(\Pi^L)^2 - x. \tag{32}$$

[19]When they are large, after some histories the principal will not monitor at all since the exogenous news would be sufficient to provide incentives, as in Board and Meyer-ter Vehn (2013). Our analysis focuses on the cases where some amount of monitoring is needed at all times to solve the agency problem.

[20]If the IC constraint is not binding at time zero then the initial value must be computed indirectly.



(a) Case with bad news

(b) Case with good news

Figure 3: Phase diagram. The (x_t, Π_t^L) system has two steady states. In each case, one of the steady states is a saddle point. Proposition 6 shows that if the optimal solution is such $m_t > 0$ all $t \geq 0$, then the optimal solution corresponds to the trajectory converging to the saddle point. In this case, the analysis of the phase diagram reveals that the trajectory of Π_t^L must be monotone between news arrivals. This immediately implies that the evolution of monitoring between news is monotone as well.

This quadratic equation has two solutions. We show that in the good news case only the largest solution is consistent with a positive monitoring rate, while in the bad news only the smallest one is consistent with a positive monitoring rate. So if the solution has positive monitoring rate at all times, then the solution must correspond to the saddle point trajectory in the phase diagram in Figure 3.

From inspection of the phase diagram it is clear that Π_t^L is monotone: it is decreasing after good news and increasing after bad news. This implies that in the bad news case, monitoring increases after bad news and increases after good news

Proposition 6. *Suppose that in the optimal monitoring policy the incentive compatibility constraint binds at all time, then the monitoring rate m_t is decreasing in beliefs x_t in the bad news case, and increasing in beliefs x_t in the good news case.*

Figure 4 shows the dynamics of monitoring implied by Proposition 6. In the bad news case, monitoring increases after news. The opposite trend is optimal in the good news case. As previously mentioned, this is driven by the dynamics of reputational incentives. In the bad news case, incentives weaken as reputation goes down. As Board and Meyer-ter Vehn (2013) point out, a high reputation firm has more to lose from a collapse in its reputation following a breakdown than a low reputation firm. Hence, monitoring is most useful for incentive purposes when reputation is low. In the good news case, incentives decrease in reputation; a low reputation firm has more to gain from a breakthrough that boosts its reputation than a high reputation firm. In the good news case, monitoring is thus most useful when reputation is high. Accordingly monitoring complements exogenous news, being used when exogenous news are ineffective at providing incentives. The same intuition applies to the good news case, but in the opposite direction. In the good news case, incentives from reputation are weakest for high reputations, hence, monitoring rates are higher when reputations are high.

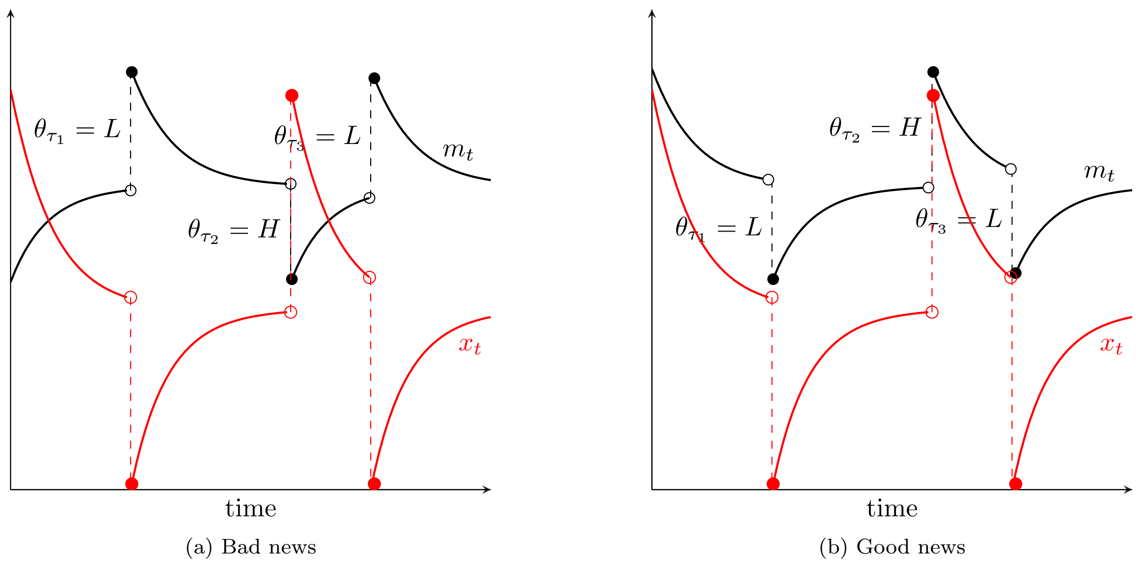


Figure 4: Response of monitoring rates to exogenous news in the bad news and good new cases. In both pictures the starting belief is $x_0 = 1$. The black curves represent optimal monitoring intensity, m_t and the red curves the evolution of reputation, x_t. We picked sample paths for the two cases that have the same realizations of the time of monitoring/news and the same realizations of the revealed quality at those times. In the bad news case (left panel) the rate of monitoring increases after negative news (either from inspection or exogenous news). Moreover, optimal monitoring intensity is decreasing in beliefs. The dynamics of monitoring are the opposite in the good news case.

6 Conclusions

This paper studies optimal monitoring policies in dynamic settings where a firm exerts hidden effort to affect a persistent quality process and has reputation concerns. We used (building on Board and Meyer-ter Vehn (2013)) a capital-theoretic approach to modeling reputation and quality. Our model is flexible and can be applied to situations ranging from quality testing of schools, product/service quality testing, and capital budgeting in multi-divisional firms.

The principal, who can represent a regulator or a self-regulatory organization, commits to a monitoring policy that specifies a probability distribution of costly inspections as a function of past results. We focus on policies that always induce full effort. While we assume that the firm's flow payoffs are linear in its reputation, we allow for arbitrary weakly convex preferences for the principal. The optimal monitoring policy hence has dual roles: learning (due to convexity of the principal's payoff flows in beliefs) and incentive provision. Learning favors deterministic inspections and incentive provision favors random inspections. We show that in our model the optimal policy always takes the form of one of these two extremes.

In practice, monitoring is often triggered by the revelation of some news. We consider the interaction of monitoring and exogenous news and characterize circumstances under which optimal monitoring policies have a higher monitoring intensity when reputations are lower.

We consider policies that implement full effort. Despite full effort is efficient in the first best, given the agency problems in our model, it is possible that for some parameters an optimal policy would call for no inspections and no effort after some histories. The intuition comes from the case $\bar{a} = 1$ so that if the firm achieves high quality, it can prevent any future drops by putting full effort. In that case, an inspection showing low quality would be off-the-equilibrium path and it would optimal to use the worst possible punishment for the firm to relax incentive constraints. That would indeed involve no more effort or monitoring. If $\bar{a}$ is very close to 1, it is possible that such punishments would remain optimal. However, we expect that if cost of inspections is not too large and $\bar{a}$ is sufficiently away from 1, the optimal policy would indeed induce full effort.

Generalizing our analysis to characterize optimal joint design of the monitoring policy and effort choice is a potential avenue for future research. We have considered settings without transfers, where incentives are driven purely by reputation concerns. As Dewatripont et al. (1999) point out many incentives in organizations arise not through explicit formal incentive contracts but rather implicitly through career concerns. That said, extending the analysis to settings where the firm's payoffs are non-linear (e.g., because firms pay part of the monitoring costs or must pay a fine when caught having low quality) is another extension that would allow us to extend the scope of applications of this model. Also, allowing for the possibility of monetary transfers, as in a standard moral hazard or debt contracting setting, would allow us to understand the interaction between compensation and monitoring policies.

Appendix

A Relaxed Problem without Agency Problems

Proof of Result 3

Proof. Differentiating the HJB equation we get that for any $x \notin [\underline{x}, \overline{x}]$ we have

$$(r+\lambda)U'(x) = u'(x) + \lambda(\bar{a} - x)U''(x) \tag{33a}$$

$$(r+2\lambda)U''(x) = u''(x) + \lambda(\bar{a} - x)U'''(x) \tag{33b}$$

Using (33b) we get that for any $x > \bar{a}$ we have $U''(x) = 0 \Rightarrow U'''(x) > 0$. This means that $U''(\overline{x}) \geq 0 \Rightarrow U''(x) > 0$ for all $x > \overline{x}$. Similarly, for any $x < \bar{a}$ we have $U''(x) = 0 \Rightarrow U'''(x) < 0$ which means that $U''(\underline{x}) \geq 0 \Rightarrow U''(x) > 0$ for all $x < \underline{x}$. Evaluating (33a) at $\overline{x}$ and using the smooth pasting condition we find that

$$(r+\lambda)(U(1) - U(0)) = u'(\overline{x}) + \lambda(\bar{a} - \overline{x})U''(\overline{x})$$

Hence, U we have that $U''(\overline{x}) \geq 0$ and $U''(\underline{x}) \geq 0$ if and only if

$$\frac{u'(\underline{x})}{r+\lambda} \leq U(1) - U(0) \leq \frac{u'(\overline{x})}{r+\lambda} \tag{34}$$

The HJB equation together with the boundary conditions imply that

$$r(U(0) + \overline{x}(U(1) - U(0))) = u(\overline{x}) + \lambda(\bar{a} - \overline{x})(U(1) - U(0))$$
$$r(U(0) + \underline{x}(U(1) - U(0))) = u(\underline{x}) + \lambda(\bar{a} - \underline{x})(U(1) - U(0))$$

Taking the difference between these two equations and rearranging terms we find that

$$U(1) - U(0) = \frac{1}{r+\lambda} \frac{u(\overline{x}) - u(\underline{x})}{\overline{x} - \underline{x}}.$$

It follows from the convexity of u that inequality (34) is satisfied. The fact that U is increasing follows directly from the convexity of U and equation (33a).

Next, let's define

$$H(x) \equiv xU(1) + (1-x)U(0) - U(x).$$

The convexity of U implies that H is concave and $H(x) = c$ for $x \in [\underline{x}, \overline{x}]$ and $H(x) < c$ for $x \notin [\underline{x}, \overline{x}]$. Hence, we get that

$$xU(1) + (1-x)U(x) - U(x) \leq c. \tag{35}$$

Similarly, let's define

$$G(x) \equiv u(x) + \lambda(\bar{a} - x)(U(1) - U(0)) - r(xU(1) + (1-x)U(0) - c).$$

Differentiating the previous equation twice we get that $G''(x) = u''(x) > 0$. Because $U(\cdot)$ is continuously differentiable we have that $G(\underline{x}) = G(\overline{x}) = 0$. Hence, we can conclude that $G(x) < 0$ for all $x \in (\underline{x}, \overline{x})$. Accordingly,

$$0 \geq u(x) + \lambda(\bar{a} - x)U'(x) - rU(x), \; x \in [0,1]. \tag{36}$$

The final step is to verify that we can not improve the payoff using an alternative policy. Let $(\tilde{T}_n)_{n\geq 1}$ and let $\tilde{x}_t$

be the belief process induce by this policy. Applying Ito's lemma to the process $e^{-rt}U(\tilde{x}_t)$ we get

$$
\begin{aligned}
e^{-rt}E[U(\tilde{x}_t)] &= U(x_0) + E\left[\int_0^t e^{-rs}(\lambda(\bar{a}-\tilde{x}_t)U'(\tilde{x}_t) - rU(\tilde{x}_t))\mathrm{d}s + \sum_{s\le t} e^{-rs}\big(\tilde{x}_s U(1) + (1-\tilde{x}_s)U(0) - U(\tilde{x}_s)\big)\right] \\
&\le U(x_0) - E\left[\int_0^t e^{-rs}u(\tilde{x}_t)\mathrm{d}s - \sum_{s\le t} e^{-rs}c\right], \qquad (37)
\end{aligned}
$$

where we have used inequalities (35) and (36). Taking the limit when $t\to\infty$ we conclude that

$$
U(x_0) \ge E\left[\int_0^\infty e^{-rs}u(\tilde{x}_t)\mathrm{d}s - \sum_{\tilde{T}_n\ge 0} e^{-r\tilde{T}_n}c\right]
$$

The proof concludes noting that (37) holds with equality for the optimal policy. □

B Case without Exogenous News

B.1 Incentive Compatibility: Proof of Lemma 1

Proof. Let $N_t^{LH} = \sum_{s\le t}\mathbf{1}_{\{\theta_{s^-}=L,\theta_s=H\}}$ and $N_t^{HL} = \sum_{s\le t}\mathbf{1}_{\{\theta_{s^-}=H,\theta_s=L\}}$ be counting processes indicating the number of switches from L to H and from H to L respectively. The processes

$$
\begin{aligned}
Z_t^{LH} &= N_t^{LH} - \int_0^t (1-\theta_s)\lambda a_s \mathrm{d}s \\
Z_t^{HL} &= N_t^{HL} - \int_0^t \theta_s\lambda(1-a_s)\mathrm{d}s,
\end{aligned}
$$

are martingales. Defining $Z_t \equiv Z_t^{LH} - Z_t^{HL}$ and noting that we can write $\mathrm{d}\theta_t = \mathrm{d}N_t^{LH} - \mathrm{d}N_t^{HL}$ we get that θ_t satisfies the stochastic differential equation

$$
\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \mathrm{d}Z_t
$$

We can solve the previous equation and get that

$$
\theta_t = e^{-\lambda t}\theta_0 + \int_0^t e^{-\lambda(t-s)}(\lambda a_s \mathrm{d}s + \mathrm{d}Z_s).
$$

Full effort is incentive compatible if and only if

$$
\begin{aligned}
&E_t^{\bar{a}}\left[\int_t^{T_n} e^{-r(s-t)}(x_s - k\bar{a})\mathrm{d}s + e^{-r(T_n-t)}\big(\theta_{T_n}\Pi(H) + (1-\theta_{T_n})\Pi(L)\big)\right] \ge \\
&E_t^{\tilde{a}}\left[\int_t^{T_n} e^{-r(s-t)}(x_s - k\tilde{a}_s)\mathrm{d}s + e^{-r(T_n-t)}\big(\tilde{\theta}_{T_n}\Pi(H) + (1-\tilde{\theta}_{T_n})\Pi(L)\big)\right]
\end{aligned}
$$

Letting $\Delta \equiv \Pi(H) - \Pi(L)$ and replacing the solution of θ_t, we can write the incentive compatibility condition as

$$
E_t\left[\int_t^{T_n} e^{-r(s-t)}\big(\lambda e^{-(r+\lambda)(T_n-s)}\Delta - k\big)(\bar{a}-\tilde{a}_s)\mathrm{d}s\right] \ge 0.
$$

Hence, for any deviation we have that

$$E_t\left[\int_t^{T_n} e^{-r(s-t)}\big(\lambda e^{-(r+\lambda)(T_n-s)}\Delta-k\big)(\bar a-\tilde a_s)\mathrm{d}s\right]=$$
$$E_t\left[\int_t^{\infty} \mathbf{1}_{\{T_n>s\}}e^{-r(s-t)}\big(\lambda e^{-(r+\lambda)(T_n-s)}\Delta-k\big)(\bar a-\tilde a_s)\mathrm{d}s\right]=$$
$$E_t\left[\int_t^{\infty} \mathbf{1}_{\{T_n>s\}}e^{-r(s-t)}\big(\lambda E_s[e^{-(r+\lambda)(T_n-s)}|T_n>s]\Delta-k\big)(\bar a-\tilde a_s)\mathrm{d}s\right]=$$
$$E_t\left[\int_t^{T_n} e^{-r(s-t)}\big(\lambda E_s[e^{-(r+\lambda)(T_n-s)}|T_n>s]\Delta-k\big)(\bar a-\tilde a_s)\mathrm{d}s\right]$$

Hence, we can write the incentive compatibility condition as

$$E_t\left[\int_t^{T_n} e^{-r(s-t)}\big(\lambda E_s[e^{-(r+\lambda)(T_n-s)}|T_n>s]\Delta-k\big)(\bar a-\tilde a_s)\mathrm{d}s\right]\ge 0$$

The result in the lemma follows directly after replacing $\Delta=\Pi(H)-\Pi(L)=1/(r+\lambda)$. □

B.2 Principal's Problem

B.2.1 Existence

Proof of Lemma 2

Proof. Let's denote the vector of expected payoffs by $U\equiv(U_L,U_H)$. We have that $U^{\max}=(u(1)-k\bar a)/r<\infty$ is an upper bound for the principal payoff. The monitoring policy $m_t=0$, and τ solving $e^{-(r+\lambda)\tau}=\underline{q}$ provides a lower bound $U_\theta^{\min}>-\infty$. We consider the rectangle $R=[U_L^{\min},U^{\max}]\times[U_H^{\min},U^{\max}]$. Let $\mathscr{G}_\epsilon^\theta$ be the Bellman operator with the extra constraint that $E(e^{-rT})=\int_0^\infty e^{-rt}dF(t)\le e^{-r\epsilon}$. For any bounded functions f,g we have that $|\sup f-\sup g|\le\sup|f-g|$. Clearly, the function $\mathscr{G}_\epsilon=(\mathscr{G}_\epsilon^L,\mathscr{G}_\epsilon^H)$ is bounded in R. Accordingly, we have that

$$\|\mathscr{G}_\epsilon U^0-\mathscr{G}_\epsilon U^1\|\le e^{-r\epsilon}\|U^0-U^1\|,$$

and by the Contraction Mapping Theorem there is a unique fixed-point $\mathscr{G}_\epsilon U_\epsilon=U_\epsilon$. For any sequence $\epsilon_k\downarrow 0$ we get that the sequence U_{ϵ_k} is increasing and bounded above by $U^{\max}$. Accordingly U_{ϵ_k} converges to some limit $\overline{U}$. Moreover, $\mathscr{G}_\epsilon$ is lower semicontinuous as a function of ϵ (Aliprantis and Border, 2006, Lemma 17.29) so

$$\lim_{\epsilon_k\downarrow 0}\mathscr{G}_{\epsilon_k}U_{\epsilon_k}\ge\mathscr{G}\overline{U}.$$

On the other hand, $\mathscr{G}_\epsilon$ is increasing in U, decreasing in ϵ and U_{ϵ_k} is an increasing sequence so

$$\lim_{\epsilon_k\downarrow 0}\mathscr{G}_{\epsilon_k}U_{\epsilon_k}\le\mathscr{G}\overline{U}.$$

Accordingly, $\lim_{\epsilon_k\downarrow 0}\mathscr{G}_{\epsilon_k}U_{\epsilon_k}=\mathscr{G}\overline{U}$ and we conclude that

$$\overline{U}=\lim_{\epsilon_k\downarrow 0}U_{\epsilon_k}=\lim_{\epsilon_k\downarrow 0}\mathscr{G}_{\epsilon_k}U_{\epsilon_k}=\mathscr{G}\overline{U}.$$

□

B.3 Principal's Optimization Problem

The introduction of the deadline τ in the main text is useful for expositional purposes. However, it is redundant because we are allowing for jumps in the cumulative density function. The optimal control problem is Section 4.2 is equivalent to the infinite horizon problem

$$\sup_{m_t,p_t,q_0} \int_0^\infty e^{-rt-M_{t-}} \left(u(x_t^\theta) + m_t \mathcal{M}_\theta(U, x_t)\right) \mathrm{d}t + \sum e^{-rs-M_{s-}} p_s \mathcal{M}_\theta(U, x_s^\theta)$$

subject to

$$M_t = \int_0^t m_s \mathrm{d}s - \sum_{s\le t} \log(1-p_s)$$

$$q_t = q_0 + \int_0^t [(r+\lambda+m_s)q_s - m_s]\mathrm{d}s - \sum_{s\le t} \frac{p_s}{1-p_s}(1-q_{s-}),\ q_\tau = 1$$

$$q_t \in [\underline{q}, 1],\ \forall t \in [0,\tau]$$

$$0 \le m_t$$

$$p_s \le q_{t-}.$$

In this alternative formulation, the deadline is defined as $\tau = \inf\{t > 0 : q_t = 1\}$. Hence, we only need to show that a candidate solution satisfies the necessary and sufficient conditions of the infinite horizon formulation.

B.3.1 Linear Payoffs

Proof of Proposition 2

Proof. We prove first that the proposed policy is optimal among all the policies without atoms. Then we show that this implies that it also dominates any policy with atoms.

We verify that our conjecture solution satisfies necessary and sufficient conditions for optimality given by Theorem 10 in Seierstad and Sydsaeter (1977). The first step is to verify that the non-negativity condition for the Lagrange multiplier is satisfied. Equation (21) for ψ_t reduces to

$$rU_t \ge x_t^\theta + (r\underline{q} + \lambda)\nu_t + \dot{x}_t^\theta (U_H - U_L). \tag{38}$$

We verify that this expression is positive so the state constraint is binding. The condition $S(t) = 0$ implies that

$$U_t = x_t^\theta U_H + (1 - x_t^\theta) U_L - c - (1 - \underline{q})\nu_t$$

Replacing in (38) we get the condition

$$r(x_t^\theta U_H + (1 - x_t^\theta)U_L - c) \ge x_t^\theta + (r+\lambda)\nu_t + \dot{x}_t^\theta (U_H - U_L).$$

Let's define

$$\Gamma(t) = x_t^\theta + \dot{x}_t^\theta (U_H - U_L) - r(x_t^\theta U_H + (1 - x_t^\theta)U_L - c)$$

so the previous condition reduces to $\Gamma(t) + (r+\lambda)\nu_t \le 0$. Replacing U_L and $U_H - U_L$ we get

$$\Gamma(t) \equiv (m^* + r)c$$

This means that $\Gamma(t) + (r+\lambda)\nu_t \le 0$ if and only if

$$(r+\lambda)\nu_t \le -(r+m^*)c \tag{39}$$

If we replace (21) in (15b) we get the differential equation for ν_t

$$\begin{aligned}\dot{\nu}_t &= (r+m^*)\nu_t + \frac{\Gamma(t)}{1-\underline{q}} \\ &= (r+m^*)\nu_t + \frac{(r+m^*)c}{1-\underline{q}} \\ &= \frac{(r+\lambda\underline{q})\nu_t + (r+m^*)c}{1-\underline{q}}\end{aligned}$$

Let ν^{ss} be the steady state of ν_t which is given by

$$\nu^{ss} = -\frac{(r+m^*)c}{r+\lambda\underline{q}}$$

We can immediately verify that inequality (39) is satisfied for ν^{ss}. Hence, it is sufficient to show that (39) is satisfied for ν_0. Condition $S(0) = 0$ implies that

$$\nu_0 = -\frac{c}{1-\underline{q}}.$$

We can verify that (39) is satisfied for ν_0 by replacing ν_0, $\underline{q}$ and m^* in (39). This means that $\psi_t \geq 0$ for all t.

By construction $S(t) = 0$ for all t so $p_t = 0$ for all t satisfies the first order condition (24). Thus, our candidate monitoring policy satisfies all the necessary conditions.

Next, we verify the sufficient conditions. As the co-state variable ν_t is bounded the transversality condition

$$\lim_{t\to\infty} e^{-rt-M_t}\nu_t = 0$$

is automatically satisfied. Moreover, given that for any admissible policy $(\tilde{m}_t, \tilde{p}_t, \tilde{\tau})$ the trajectory of the state variables $\tilde{q}_t$ and $\tilde{U}_t$ is bounded, the transversality condition (17) guarantees that, if q_t and U_t are the trajectories given the optimal policy, and $\tilde{q}_t$ and $\tilde{U}_t$ are trajectories for any other admissible policy, then the following sufficient condition for infinite horizon problems is satisfied (Seierstad and Sydsaeter, 1986, p. 385):

$$\lim_{t\to\infty} e^{-rt-M_t}\big(\nu_t(\tilde{q}_t - q_t) - (\tilde{U}_t - U_t)\big) \geq 0.$$

Finally, the maximized Hamiltonian is linear in the state variables, and so it is automatically concave and so Arrow's sufficient condition is satisfied. This means that the policy is optimal among all the policies with no atoms.

The next step is to show that this policy also dominates any policy with atoms. Let $F(t)$ be the cdf of an arbitrary incentive compatible policy that may have atoms. Because we have allowed the intensity m_t to be unbounded, then for any such $F(t)$ there is some $\tilde{F}(t)$ with no atoms such that $\tilde{F}(t)$ is arbitrarily close to $F(t)$ at any point of continuity t. Hence, we can find a sequence of continuous distributions $\tilde{F}_n(t)$ such that $\tilde{F}_n(t) \to_d F(t)$. Let U_0^* be the principal's profit of the optimal policy above, and let U_0 and $\tilde{U}_0^n$ be the profit's associated to the policies $F(t)$ and $\tilde{F}_n(t)$, respectively. $\tilde{F}_n(t) \to_d F(t)$ implies that $\tilde{U}_0^n \to U_0$. Moreover, $U_0^* \geq \tilde{U}_0^n$, all n, which means that $U_0^* \geq \lim_{t\to n} \tilde{U}_0^n = U_0$. □

B.3.2 Convex Payoffs

Proof of Lemma 4

Proof. In order to prove the lemma we show that any policy that does not satisfy the lemma violates the necessary conditions for optimality. If $q_t > \underline{q}$ then the Lagrange multiplier is $\psi_t = 0$. We show that if $\psi_t = 0$ then it can not be the case that $S(t) = 0$ along a singular arc, which is a necessary condition for $0 < m_t < \infty$.

Looking for a contradiction, suppose that $q_t > \underline{q}$ and $m_t > 0$. Then, it must be the case $S(t)$ is constant and equal to zero along a circular arc. We have that $\dot{S}_t = 0$ and so

$$- rU_t + u(x_t^\theta) + (rq_t + \lambda)\nu_t + \dot{x}_t^\theta(U_H - U_L) = 0 \tag{40}$$

Moreover, because $S(t)$ is constant along the singular arc, it must also be the case that $\ddot{S}(t) = 0$ which means that

$$-r\dot{U}_t + u'(x_t^\theta)\dot{x}_t^\theta + r\dot{q}_t\nu_t + (rq_t + \lambda)\dot{\nu}_t + \ddot{x}_t^\theta(U_H - U_L) = 0$$

Using the first order condition $m_t S(t) = 0$ we can write the evolution of U_t

$$\dot{U}_t = rU_t - u(x_t^\theta) - m_t\nu_t(1 - q_t).$$

Replacing $\dot{U}_t$ and $\dot{\nu}_t$ we get

$$r(-rU_t + u(x_t^\theta) + m_t\nu_t(1 - q_t)) + u'(x_t^\theta)\dot{x}_t^\theta + r\nu_t((r + \lambda + m_t)q_t - m_t) - (rq_t + \lambda)\lambda\nu_t + \ddot{x}_t^\theta(U_H - U_L) = 0,$$

which after some simplification yield

$$- rU_t + u(x_t^\theta) + r^{-1}u'(x_t^\theta)\dot{x}_t^\theta + r^{-1}(r^2 q_t - \lambda^2)\nu_t + r^{-1}\ddot{x}_t^\theta(U_H - U_L) = 0 \tag{41}$$

Combining equations (47) and (48) we get that along the circular arc we have

$$\nu_t = \frac{\dot{x}_t^\theta}{\lambda^2 + r\lambda}\Big(u'(x_t^\theta) - (r + \lambda)(U_H - U_L)\Big) \tag{42}$$

Differentiating (49) we get

$$\begin{aligned}
\dot{\nu}_t &= \frac{\ddot{x}_t^\theta}{\lambda^2 + r\lambda}\Big(u'(x_t^\theta) - (r + \lambda)(U_H - U_L)\Big) + \frac{(\dot{x}_t^\theta)^2}{\lambda^2 + r\lambda}u''(x_t^\theta) \\
&= \frac{\ddot{x}_t^\theta}{\dot{x}_t^\theta}\nu_t + \frac{(\dot{x}_t^\theta)^2}{\lambda^2 + r\lambda}u''(x_t^\theta) \\
&= -\lambda\nu_t + \frac{(\dot{x}_t^\theta)^2}{\lambda(r + \lambda)}u''(x_t^\theta)
\end{aligned}$$

Equation (15b) implies that

$$\dot{\nu}_t = -\lambda\nu_t - \psi_t = -\lambda\nu_t < -\lambda\nu_t + \frac{(\dot{x}_t^\theta)^2}{\lambda(r + \lambda)}u''(x_t^\theta),$$

which yields the contradiction as $u''(x_t^\theta) > 0$. This shows that it cannot be the case that $q_t > \underline{q}$ and $S(t) = 0$ over a singular arc. □

Proof of Lemma 5

Proof. We prove the Lemma showing that the necessary conditions for a jump can only be satisfied at one point and that if an atom exists at t it must be the case that $p_t = 1$. First, we specify the necessary conditions for a jump to be optimal. The maximum principle for optimal control problems with jumps in the state variables in Seierstad and Sydsaeter (1986). The optimality conditions are stated in terms of the non-current-value co-state variables, while ν_t and ζ_t are current value co-states. Hence, when we apply the theorem we need to be careful of using the multipliers $\tilde{\nu}_t = e^{-rt-M_t}\nu_t$ and $\tilde{\zeta}_t = e^{-rt-M_t}\zeta_t$ to account for the effect of jumps in the discounting factor on the co-state variables.

Necessary conditions: Theorem 7, p. 196 in Seierstad and Sydsaeter (1986) requires that at any jump point t the co-state variables satisfy

$$e^{-rt-M_{t^+}}\nu_{t^+} - e^{-rt-M_{t^-}}\nu_{t^-} = -e^{-rt-M_{t^+}}\nu_{t^+}\frac{p_t}{1-p_t}.$$

Hence, we have that

$$\begin{aligned}\nu_{t^+} - e^{M_{t^+}-M_{t^-}}\nu_{t^-} &= \nu_{t^+} - \frac{\nu_{t^-}}{1-p_t}\\ &= -\nu_{t^+}\frac{p_t}{1-p_t}\end{aligned}$$

and so $\nu_{t^+} = \nu_{t^-}$. We also have that $U_{t^-} = p_t\mathcal{M}(U, x_t^\theta) + (1-p)U_{t^+}$. If $p_t > 0$ then the switching function satisfies

$$\begin{aligned}S(t^+) &= x_t^\theta U_H + (1-x_t^\theta)U_L - c - U_{t^+} - (1-q_t)\nu_{t^+}\\ &= \mathcal{M}(U, x_t^\theta) - \frac{U_{t^-}}{1-p_t} + \frac{p_t}{1-p_t}\mathcal{M}(U, x_t^\theta) - (1-q_t)\nu_{t^-}\\ &= S(t^-) + \frac{p_t}{1-p_t}\big(\mathcal{M}(U, x_t^\theta) - U_{t^-}\big) \qquad (43)\end{aligned}$$

Using the continuity of ν_t, the jump size p_t solve

$$p_t \in \arg\max_{p\in[0,(q_{t^-}-\underline{q})/(1-\underline{q})]}\left\{\frac{p}{1-p}\Big[\mathcal{M}_\theta(U, x_t^\theta) - U_{t^-} - \nu_{t^-}(1-q_{t^-})\Big]\right\}.$$

Hence $p_t > 0$ only if $S(t^-) \geq 0$ and $p_t = q_{t^-}$ if $S(t^-) > 0$. Moreover, a necessary condition of optimality (Seierstad and Sydsaeter, 1986, Note 7, p. 197) is that at any jump time $t < \tau$ we have

$$\begin{aligned}0 =& e^{-M_{t^+}}\mathcal{H}(q_{t^+}, \zeta_{t^+}, \nu_{t^+}, \psi_{t^+}, m_{t^+}, t) - e^{-M_{t^-}}\mathcal{H}(q_{t^-}, \zeta_{t^-}, \nu_{t^-}, \psi_{t^-}, m_{t^-}, t)\\ &- e^{-M_{t^+}}\dot{x}_t^\theta(U_H - U_L)\frac{p_t}{1-p_t}.\end{aligned}$$

Factorizing by $e^{-M_{t^+}}$ we find that the previous expression is proportional to

$$\begin{aligned}\mathcal{H}(q_{t^+}, \zeta_{t^+}, \nu_{t^+}, \psi_{t^+}, m_{t^+}, t) - \frac{\mathcal{H}(q_{t^-}, \zeta_{t^-}, \nu_{t^-}, \psi_{t^-}, m_{t^-}, t)}{1-p_t} - \dot{x}_t^\theta(U_H - U_L)\frac{p_t}{1-p_t} =\\ \frac{p_t}{1-p_t}\Big[u(x_t^\theta) - r\mathcal{M}(U, x_t^\theta) + (r+\lambda)\nu_{t^-} - \dot{x}_t^\theta(U_H - U_L) - r(1-q_t)\nu_t\Big] =\\ \frac{p_t}{1-p_t}\Big[u(x_t^\theta) + r\mathcal{M}(U, x_t^\theta) + (r+\lambda)\nu_{t^-} - \dot{x}_t^\theta(U_H - U_L) - r(1-q_t)\nu_t\Big]\end{aligned}$$

This means that it must be the case that

$$r\mathcal{M}(U, x_t^\theta) = u(x_t^\theta) + (r+\lambda)\nu_{t^-} + \dot{x}_t^\theta(U_H - U_L) - r(1-q_t)\nu_t$$

Note that this is the same as condition (22) to determine τ (in which case $q_\tau = 1$). Hence, it is enough to show that there is at most one point satisfying this condition.

Uniqueness of atom at τ: Differentiating the switching function $S(t)$ we get that

$$\dot{S}(t) = rS(t) + G(t),$$

where

$$G(t) \equiv u(x_t^\theta) + (r+\lambda)\nu_{t^-} + \dot{x}_t^\theta(U_H - U_L) - r\mathcal{M}(U, x_t^\theta).$$

Suppose that there is an atom at $\tilde{\tau} < \tau$, then it must be the case that $S(\tilde{\tau}^-) = 0$ and $G(\tilde{\tau}^-) = r(1 - q_{\tilde{\tau}-})\nu_{\tilde{\tau}-} \leq 0$ ($\psi_t \geq 0$ and $\nu_0 \leq 0$ implies that ν_t is non-positive). The next step is to show that the previous conditions imply that it can not be the case that $S(\tau^-) = 0$ and $\dot{S}(\tau^-) = 0$. By the definition of $G(t)$ we have that $G(\tau) = 0$ so it is enough to show that this condition can not be satisfied given that $G(\tilde{\tau}) \leq 0$. We have to consider two cases, $G(\tilde{\tau}) < 0$ and $G(\tilde{\tau}) = 0$.

Case $G(\tilde{\tau}) < 0$: $G(\tau) = 0$ implies that $\dot{S}(\tau) = 0$. At time τ, it must be the case that there is some $\epsilon > 0$ such that for all $t \in (\tau - \epsilon, \tau)$ we have: i) $q_t > \underline{q}$, and i) $S(t) < 0$. Point i) follows from the constraint $p_t \leq (q_{t^-} - \underline{q})/(1 - \underline{q})$ while ii) follows from i) together with Lemma 4. Because $S(t)$ is negative for $t \in (\tau - \epsilon, \tau)$, $S(\tau^-) = 0$ and $\dot{S}(\tau) = G(\tau) = 0$ there must be some interval $(\tau - \epsilon', \tau)$ such that $\dot{S}(t) > 0$ and $\ddot{S}(t) < 0$ for all $t \in (\tau - \epsilon', \tau)$. This means that $\ddot{S}(\tau^-) = \dot{S}(\tau^-) + \dot{G}(\tau^-) = \dot{G}(\tau) \leq 0$.

Given the hypothesis that $G(\tilde{\tau}) < 0$ and given that $G(\tau) = 0$, there must be some interval $(\tau - \epsilon, \tau)$ such that $\dot{G}(t) > 0$ for all $t \in (\tau - \epsilon, \tau)$. The derivative $\dot{G}(t)$ satisfies the differential equation

$$\ddot{G}(t) = -\lambda \dot{G}(t) + (\dot{x}_t^\theta)^2 u''(x_t^\theta).$$

Solving for $\dot{G}(t)$ we get

$$\dot{G}(t) = -\int_t^\tau e^{\lambda(s-t)} (\dot{x}_s^\theta)^2 u''(x_s^\theta) \mathrm{d}s + e^{\lambda(\tau - t)} \dot{G}(\tau).$$

Because $\dot{G}(\tau) \leq 0$ and $(\dot{x}_s^\theta)^2 u''(x_s^\theta) \geq 0$ we get that $\dot{G}(t) \leq 0$ for all $t \in (\tilde{\tau}, \tau)$ which yields a contradiction to the fact that $\dot{G}(t) > 0$ for all $t \in (\tau - \epsilon, \tau)$.

Case $G(\tilde{\tau}) = 0$: If $G(\tilde{\tau}) = 0$ then $\dot{S}(\tilde{\tau}) = 0$ and, by a similar argument that the one we use for τ in the previous case applied to $\tilde{\tau}$, $\ddot{S}(t) \leq 0$ in an interval $(\tilde{\tau} - \epsilon, \tilde{\tau})$. But then it must be the case that $\dot{G}(\tilde{\tau}) \leq 0$. If $\dot{G}(\tilde{\tau}) = 0$ then our previous equation implies that $\dot{G}(t) > 0$ for some positive measure set $\mathcal{T} \subset (\tilde{\tau}, \tau)$. This immediately implies that $G(\tau) > 0$ giving a contradiction. On the other hand, if $\dot{G}(\tilde{\tau}) < 0$, then there is $\tilde{t} > \tilde{\tau}$ such that $G(\tilde{t}) < 0$. We can then replicate the same argument as in the case with $G(\tilde{\tau}) < 0$ starting from time $\tilde{t}$.

Hence, if the necessary conditions for an atom are satisfied at τ then these conditions can not be satisfied at any other time $\tilde{\tau} < \tau$.

Atom at τ implies $p_\tau = 1$: The only remaining step is to argue that if there is an atom at time τ then $p_\tau = 1$. Suppose this is not the case, and in particular suppose that at time τ the atom is $p < 1$. If $p < 1$ then we have that $S(\tau^+) < 0$ (equation (43)). Accordingly, it is the case that $m_t = 0$ and so $q_{t+\epsilon} > \underline{q}$ for small $\epsilon > 0$. Lemma 4 implies then that $m_{t+\epsilon} = 0$ and so $m_t = 0$ and $\dot{q}_t > 0$ for all $t > \tau$. But then, there is some $\bar{\tau} > \tau$ such that $q_{\bar{\tau}} = 1$ but this can only be the case if there is an atom in the CDF at time $\bar{\tau}$. This contradicts the previous result that $F(t)$ has at most one atom.

Proof that $\hat{\tau}_\theta^*$ is an extreme point

We consider the optimization problem

$$U_\theta = \max_{\tau \geq 0} \int_0^\tau e^{-rt - m^* t \wedge \hat{\tau}} \left(u(x_t^\theta) + m^* \mathcal{M}(U, x_t^\theta) \mathbf{1}_{t < \hat{\tau}} \right) \mathrm{d}t + e^{-r\tau - m^* \tau \wedge \hat{\tau}} \mathcal{M}(U, x_\tau^\theta)$$

$$\hat{\tau} = \max \left(\tau + \frac{\log \underline{q}}{r + \lambda}, 0 \right).$$

The objective function is

$$\mathcal{U}(\tau) = \int_0^{\hat{\tau}} e^{-(r+m^*)t} \left(u(x_t^\theta) + m^* \mathcal{M}(U, x_t^\theta) \right) \mathrm{d}t + e^{-m^* \hat{\tau}} \int_{\hat{\tau}}^\tau e^{-rt} u(x_t^\theta) \mathrm{d}t + e^{-r\tau - m^* \hat{\tau}} \mathcal{M}(U, x_\tau^\theta)$$

Let's consider τ such that $\hat{\tau} > 0$ (that is, interior candidates). The first derivative is

$$\begin{aligned}\mathcal{U}'(\tau) &= e^{-(r+m^*)\hat{\tau}}\left(u(x^\theta_{\hat{\tau}}) + m^*\mathcal{M}(U, x^\theta_{\hat{\tau}})\right) - m^* e^{-m^*\hat{\tau}} \int_{\hat{\tau}}^{\tau} e^{-rt} u(x^\theta_t)\mathrm{d}t \\ &+ e^{-r\tau - m^*\hat{\tau}} u(x^\theta_\tau) - e^{-(r+m^*)\hat{\tau}} u(x^\theta_{\hat{\tau}}) - (r+m^*) e^{-r\tau - m^*\hat{\tau}} \mathcal{M}(U, x^\theta_\tau) \\ &+ e^{-r\tau - m^*\hat{\tau}} \dot{x}^\theta_\tau (U_H - U_L) \\ &= e^{-m^*\hat{\tau}} \Big[m^* e^{-r\hat{\tau}} \mathcal{M}(U, x^\theta_{\hat{\tau}}) - m^* \int_{\hat{\tau}}^{\tau} e^{-rt} u(x^\theta_t)\mathrm{d}t \\ &+ e^{-r\tau} u(x^\theta_\tau) - (r+m^*) e^{-r\tau} \mathcal{M}(U, x^\theta_\tau) + e^{-r\tau} \dot{x}^\theta_\tau (U_H - U_L) \Big]\end{aligned}$$

Let's define the function

$$\begin{aligned}H(\tau) &= m^* e^{-r\hat{\tau}} \mathcal{M}(U, x^\theta_{\hat{\tau}}) - m^* \int_{\hat{\tau}}^{\tau} e^{-rt} u(x^\theta_t)\mathrm{d}t + e^{-r\tau} u(x^\theta_\tau) - (r+m^*) e^{-r\tau} \mathcal{M}(U, x^\theta_\tau) + e^{-r\tau} \dot{x}^\theta_\tau (U_H - U_L) \\ &= m^* \underline{q}^{-\frac{r}{r+\lambda}} e^{-r\tau} \mathcal{M}(U, x^\theta_{\hat{\tau}}) - m^* \int_{\hat{\tau}}^{\tau} e^{-rt} u(x^\theta_t)\mathrm{d}t + e^{-r\tau} u(x^\theta_\tau) - (r+m^*) e^{-r\tau} \mathcal{M}(U, x^\theta_\tau) + e^{-r\tau}(U_H - U_L)\dot{x}_\tau\end{aligned}$$

so we can write the first derivative as

$$\mathcal{U}'(\tau) = \underline{q}^{-\frac{m^*}{r+\lambda}} e^{-m^*\tau} H(\tau)$$

and the second derivative as

$$\mathcal{U}''(\tau) = -m^* \mathcal{U}'(\tau) + e^{-m^*\hat{\tau}} H'(\tau)$$

where

$$\begin{aligned}H'(\tau) &= e^{-r\tau} \Big[-rm^* \underline{q}^{-\frac{r}{r+\lambda}} \mathcal{M}(U, x^\theta_{\hat{\tau}}) + m^* \underline{q}^{-\frac{r}{r+\lambda}} (U_H - U_L) \dot{x}^\theta_{\hat{\tau}} + m^* q^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - (r+m^*) u(x^\theta_\tau) + u'(x^\theta_\tau) \dot{x}^\theta_\tau \\ &+ r(r+m) \mathcal{M}(U, x^\theta_\tau) - (r+m^*)(U_H - U_L) \dot{x}^\theta_\tau - (r+\lambda)(U_H - U_L) \dot{x}^\theta_\tau \Big]\end{aligned}$$

We show that any candidate interior solution τ that satisfies the first order conditions, necessarily violates the second order conditions (so it's a local minimizer rather than maximizer). That is, we show that $\mathcal{U}'(\tau) = 0 \Rightarrow \mathcal{U}''(\tau) \geq 0$, which is equivalent to show that $H(\tau) = 0 \Rightarrow H'(\tau) \geq 0$. If $H(\tau) = 0$, then we can write

$$\begin{aligned}H'(\tau) = e^{-r\tau} \Big[&\left(m^* \underline{q}^{-\frac{r}{r+\lambda}} \dot{x}^\theta_{\hat{\tau}} - (r+m^*) \dot{x}^\theta_\tau - \lambda \dot{x}^\theta_\tau \right) (U_H - U_L) \\ &+ u'(x^\theta_\tau) \dot{x}^\theta_\tau + m^* \left(q^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - u(x^\theta_\tau) - e^{r\tau} \int_{\hat{\tau}}^{\tau} re^{-rt} u(x^\theta_t)\mathrm{d}t \right) \Big]\end{aligned}$$

After replacing $\dot{x}^\theta_t = \lambda(\bar{a} - \theta)e^{-\lambda t}$ in the first term, we find that

$$\begin{aligned}m^* \underline{q}^{-\frac{r}{r+\lambda}} \dot{x}^\theta_{\hat{\tau}} - (r+m^*) \dot{x}^\theta_\tau - \lambda \dot{x}^\theta_\tau &= \lambda(\bar{a} - \theta)\big(m^* \underline{q}^{-\frac{r}{r+\lambda}} e^{-\lambda\hat{\tau}} - (r+m^*) e^{-\lambda\tau} - \lambda e^{-\lambda\tau} \big) \\ &= \lambda(\bar{a} - \theta)\big(m^* \underline{q}^{-1} - r - m^* - \lambda \big) \\ &= 0.\end{aligned}$$

Hence, we get that

$$H'(\tau) = e^{-r\tau} \left[u'(x^\theta_\tau) \dot{x}^\theta_\tau + m^* \left(\underline{q}^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - u(x^\theta_\tau) + e^{r\tau} \int_{\hat{\tau}}^{\tau} (-re^{-rt}) u(x^\theta_t)\mathrm{d}t \right) \right]$$

Using integration by parts

$$e^{r\tau}\int_{\hat{\tau}}^{\tau}(-re^{-rt})u(x_t^{\theta})\mathrm{d}t = u(x_\tau^{\theta}) - e^{r(\tau-\hat{\tau})}u(x_{\hat{\tau}}^{\theta}) - e^{r\tau}\int_{\hat{\tau}}^{\tau} e^{-rt}u'(x_t^{\theta})\dot{x}_t^{\theta}\mathrm{d}t$$
$$= u(x_\tau^{\theta}) - \underline{q}^{-\frac{r}{r+\lambda}}u(x_{\hat{\tau}}^{\theta}) - e^{r\tau}\int_{\hat{\tau}}^{\tau} e^{-rt}u'(x_t^{\theta})\dot{x}_t^{\theta}\mathrm{d}t$$

So we get

$$H'(\tau) = e^{-r\tau}u'(x_\tau^{\theta})\dot{x}_\tau^{\theta} - (r+\lambda)\frac{\underline{q}}{1-\underline{q}}\int_{\hat{\tau}}^{\tau} e^{-rt}u'(x_t^{\theta})\dot{x}_t^{\theta}\mathrm{d}t$$
$$= \lambda(\bar{a}-\theta)\left[e^{-(r+\lambda)\tau}u'(x_\tau^{\theta}) - \frac{\underline{q}}{1-\underline{q}}\int_{\hat{\tau}}^{\tau}(r+\lambda)e^{-(r+\lambda)t}u'(x_t^{\theta})\mathrm{d}t,\right]$$

where the second line comes from replacing $\dot{x}_t^{\theta} = \lambda(\bar{a}-\theta)e^{-\lambda t}$. If $\theta = 0$, then we have that $u'(x_\tau^{\theta}) > u'(x_t^{\theta})$ for all $t \leq \tau$ so

$$H'(\tau) > \lambda\bar{a}u'(x_\tau^{\theta})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}(r+\lambda)\int_{\hat{\tau}}^{\tau} e^{-(r+\lambda)t}\mathrm{d}t\right]$$
$$= \lambda\bar{a}u'(x_\tau^{\theta})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}\left(e^{-(r+\lambda)\hat{\tau}} - e^{-(r+\lambda)\tau}\right)\right]$$
$$= \lambda\bar{a}u'(x_\tau^{\theta})e^{-(r+\lambda)\tau}\left[1 - \frac{1}{1-\underline{q}}\left(1-\underline{q}\right)\right]$$
$$= 0.$$

On the other hand, if $\theta = 1$, then we have that $u'(x_\tau^{\theta}) < u'(x_t^{\theta})$ for all $t < \tau$ so

$$H'(\tau) > \lambda(1-\bar{a})u'(x_\tau^{\theta})\left[-e^{-(r+\lambda)\tau} + \frac{\underline{q}}{1-\underline{q}}(r+\lambda)\int_{\hat{\tau}}^{\tau} e^{-(r+\lambda)t}\mathrm{d}t\right]$$
$$= \lambda(1-\bar{a})u'(x_\tau^{\theta})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}\left(e^{-(r+\lambda)\hat{\tau}} - e^{-(r+\lambda)\tau}\right)\right]$$
$$= \lambda(1-\bar{a})u'(x_\tau^{\theta})e^{-(r+\lambda)\tau}\left[1 - \frac{1}{1-\underline{q}}\left(1-\underline{q}\right)\right]$$
$$= 0.$$

Hence, $H'(\tau) \geq 0$ so $\mathcal{U}''(\tau) \geq 0$ whenever $\mathcal{U}'(\tau) = 0$ which means that the maximum is never an interior point. □

C Proof Brownian Linear-Quadratic Model

We first proved that the monitoring rate is positive only if the IC constraint binds, the proof that there are no jumps before τ is similar to the binary case and so omitted. Let $U_t = -C_t$. The ODE for U_t is

$$\dot{U}_t = (r+m_t)U_t + \gamma\Sigma_t + m_t(c+\mathcal{C}). \tag{44}$$

The Hamiltonian is

$$\mathcal{H}(q_t,\zeta_t,\nu_t,\psi_t,m_t,t) = \zeta_t\big((r+m_t)U_t + \gamma\Sigma_t + m_t(c+\mathcal{C})\big) + \nu_t\left((r+\lambda+m_t)q_t - m_t\right) + \psi_t\left(q_t - \underline{q}\right).$$

The switching function is

$$S(t) = (c + \mathcal{C}) - U_t - (1 - q_t)\nu_t. \tag{45}$$

The co-state ν_t evolves according to

$$\dot{\nu}_t = -\lambda\nu_t - \psi_t. \tag{46}$$

We show that any policy that does not satisfy the condition $q_t > \underline{q} \Rightarrow m_t = 0$ violates the necessary conditions for optimality. If $q_t > \underline{q}$ then the Lagrange multiplier is $\psi_t = 0$. We show that if $\psi_t = 0$ then it can not be the case that $S(t) = 0$ along a singular arc, which is a necessary condition for $0 < m_t < \infty$.

Looking for a contradiction, suppose that $q_t > \underline{q}$ and $m_t > 0$. Then, it must be the case $S(t)$ is constant and equal to zero along a circular arc.

$$\dot{S}(t) = -\dot{U}_t + \dot{q}_t\nu_t - (1 - q_t)\dot{\nu}_t.$$

We have that $\dot{S}_t = 0$ and so

$$-rU_t - \gamma\Sigma_t + (rq_t + \lambda)\nu_t = 0 \tag{47}$$

Moreover, because $S(t)$ is constant along the singular arc, it must also be the case that $\ddot{S}(t) = 0$ which means that

$$-r\dot{U}_t - \gamma\dot{\Sigma}_t + r\dot{q}_t\nu_t + (rq_t + \lambda)\dot{\nu}_t = 0$$

Using the first order condition $m_t S(t) = 0$ we can write the evolution of C_t

$$\dot{U}_t = rU_t + \gamma\Sigma_t - m_t\nu_t(1 - q_t).$$

Replacing $\dot{U}_t$ and $\dot{\nu}_t$ we get

$$r(-rU_t - \gamma\Sigma_t + m_t\nu_t(1 - q_t)) - \gamma\dot{\Sigma}_t + r\nu_t((r + \lambda + m_t)q_t - m_t) - (rq_t + \lambda)\lambda\nu_t = 0,$$

which after some simplification yield

$$-rU_t - \gamma\Sigma_t - r^{-1}\gamma\dot{\Sigma}_t + r^{-1}(r^2 q_t - \lambda^2)\nu_t = 0 \tag{48}$$

Combining equations (47) and (48) we get that along the circular arc we have

$$\nu_t = \frac{\gamma\dot{\Sigma}_t}{\lambda(r + \lambda)}, \tag{49}$$

where

$$\begin{aligned}\dot{\Sigma}_t &= \sigma^2 - 2\lambda\Sigma_t \\ \ddot{\Sigma}_t &- 2\lambda\Sigma_t\end{aligned}$$

Differentiating (49) we get

$$\begin{aligned}\dot{\nu}_t &= \frac{\gamma\ddot{\Sigma}_t}{\lambda(r + \lambda)} \\ &= -2\lambda\frac{\gamma\dot{\Sigma}_t}{\lambda(r + \lambda)} \\ &= -2\lambda\nu_t\end{aligned}$$

Equation (15b) implies that

$$\dot{\nu}_t = -\lambda\nu_t - \psi_t = -\lambda\nu_t < -2\lambda_t\nu_t,$$

as $\nu_t < 0$ if the IC constraint is binding. We have a contradiction; it cannot be the case that $q_t > \underline{q}$ and $S(t) = 0$

over a singular arc.

Proof of Proposition 4

We prove that the solution is bang-bang by showing that any interior policy violates the second order conditions, and so it is a local maximum. Given the principal's payoff $\mathcal{C}$, the optimal policy solves the following minimization problem

$$\min_{\tau\geq 0} \int_0^{\hat{\tau}} e^{-(r+m)t}\gamma\Sigma_t \mathrm{d}t + e^{-m\hat{\tau}} \int_{\hat{\tau}}^{\tau} e^{-rt}\gamma\Sigma_t \mathrm{d}t + \frac{m(c+\mathcal{C})}{r+m}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} e^{-(r+m)\tau}\right) + \underline{q}^{-\frac{m}{r+\lambda}} e^{-(r+m)\tau}(c+\mathcal{C}),$$

where

$$\int_0^{\hat{\tau}} e^{-(r+m)t}\gamma\Sigma_t \mathrm{d}t = \frac{\gamma\sigma^2}{2\lambda(r+m)}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} e^{-(r+m)\tau}\right) - \frac{\gamma\sigma^2}{2\lambda(r+m+2\lambda)}\left(1-\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}} e^{-(r+m+2\lambda)\tau}\right)$$

$$\int_{\hat{\tau}}^{\tau} e^{-rt}\gamma\Sigma_t \mathrm{d}t = \frac{\gamma\sigma^2}{2\lambda r}\left(\underline{q}^{-\frac{r}{r+\lambda}} e^{-r\tau} - e^{-r\tau}\right) - \frac{\gamma\sigma^2}{2\lambda(r+2\lambda)}\left(\underline{q}^{-\frac{r+2\lambda}{r+\lambda}} e^{-(r+2\lambda)\tau} - e^{-(r+2\lambda)\tau}\right)$$

Let's define $G(\tau)$ as

$$\begin{aligned}
G(\tau) &\equiv \int_0^{\hat{\tau}} e^{-(r+m)t}\gamma\Sigma_t \mathrm{d}t + e^{-m\hat{\tau}} \int_{\hat{\tau}}^{\tau} e^{-rt}\gamma\Sigma_t \mathrm{d}t + \frac{m(c+\mathcal{C})}{r+m}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} e^{-(r+m)\tau}\right) + \underline{q}^{-\frac{m}{r+\lambda}} e^{-(r+m)\tau}(c+\mathcal{C}) \\
&= \frac{\gamma\sigma^2}{2\lambda(r+m)}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} e^{-(r+m)\tau}\right) - \frac{\gamma\sigma^2}{2\lambda(r+m+2\lambda)}\left(1-\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}} e^{-(r+m+2\lambda)\tau}\right) \\
&+ \frac{\gamma\sigma^2}{2\lambda r}\left(\underline{q}^{-\frac{r+m}{r+\lambda}} - \underline{q}^{-\frac{m}{r+\lambda}}\right) e^{-(r+m)\tau} - \frac{\gamma\sigma^2}{2\lambda(r+2\lambda)}\left(\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}} - \underline{q}^{-\frac{m}{r+\lambda}}\right) e^{-(r+m+2\lambda)\tau} \\
&+ \frac{m(c+\mathcal{C})}{r+m}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} e^{-(r+m)\tau}\right) + \underline{q}^{-\frac{m}{r+\lambda}} e^{-(r+m)\tau}(c+\mathcal{C})
\end{aligned}$$

It is convenient to work with the change of variable $z \equiv e^{-(r+m+2\lambda)\tau}$ and define the constant $\phi = \frac{r+m}{r+m+2\lambda} \in (0,1)$ so we get

$$\begin{aligned}
G(z) &= \frac{\gamma\sigma^2}{2\lambda(r+m)}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} z^{\phi}\right) - \frac{\gamma\sigma^2}{2\lambda(r+m+2\lambda)}\left(1-\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}} z\right) \\
&+ \frac{\gamma\sigma^2}{2\lambda r}\left(\underline{q}^{-\frac{r+m}{r+\lambda}} - \underline{q}^{-\frac{m}{r+\lambda}}\right) z^{\phi} - \frac{\gamma\sigma^2}{2\lambda(r+2\lambda)}\left(\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}} - \underline{q}^{-\frac{m}{r+\lambda}}\right) z \\
&+ \frac{m(c+C)}{r+m}\left(1-\underline{q}^{-\frac{r+m}{r+\lambda}} z^{\phi}\right) + \underline{q}^{-\frac{m}{r+\lambda}} z^{\phi}(c+C)
\end{aligned}$$

To simplify the previous expressions, we define the constants

$$\begin{aligned}
A &= \frac{\gamma\sigma^2}{2\lambda}\left[\frac{1}{r+m} - \frac{1}{r+m+2\lambda}\right] + \frac{m}{r+m}(c+\mathcal{C}) \\
B &= \frac{\gamma\sigma^2}{2\lambda(r+2\lambda)}\underline{q}^{-\frac{r+m+2\lambda}{r+\lambda}}\left[\underline{q}^{\frac{r+2\lambda}{r+\lambda}} - \frac{(r+\lambda)\underline{q}}{r+\lambda+\lambda(1-\underline{q})}\right] \\
D &= \underline{q}^{-\frac{r+m}{r+\lambda}}\left[\frac{m}{r+m} - \underline{q}^{\frac{r}{r+\lambda}}\right]\left(\frac{\gamma\sigma^2}{2\lambda r} - c - \mathcal{C}\right)
\end{aligned}$$

so we can write

$$G(z) = A + Bz + Dz^{\phi},$$

with derivatives

$$G_z(z, C) = B + \phi D z^{\phi-1}$$
$$G_{zz}(z, C) = \phi(\phi - 1) D z^{\phi-2}.$$

Suppose that z^* satisfies the first order condition $G_z(z^*) = 0$, then we have that

$$(z^*)^{\phi-1} = -\frac{B}{\phi D},$$

and so the second derivative is

$$G_{zz}(z^*) = (1 - \phi)\frac{B}{z^*}$$

To prove the proposition is suffice to show that $B < 0$, which means that it suffice to show that the function

$$f(z) = z^{\alpha+1} - \frac{z}{1 + \alpha(1 - z)},$$

where $\alpha \equiv \lambda/(r + \lambda)$, is negative for all $z \in (0, 1)$. The value of the function at zero and one is $f(0) = f(1) = 0$, and the derivative is

$$f'(z) = (\alpha + 1)\left[z^{\alpha} - \frac{1}{(1 + \alpha(1 - z))^2}\right],$$

which means that $f'(0) < 0$ and $f'(1) = 0$. Moreover, the second derivative is

$$f''(z) = (\alpha + 1)\left[\alpha z^{\alpha-1} - \frac{2\alpha}{(1 + \alpha(1 - z))^3}\right],$$

so at $z = 1$ we have $f''(1) < 0$. We can conclude from here that there is $\epsilon > 0$ such that $f(z) < 0$ for all $z \in (1 - \epsilon, 1)$. If there is $z^{\dagger}$ such that $f(z^{\dagger}) > 0$, the by continuity there is $\tilde{z}$ with $f(\tilde{z}) = 0$. Such $\tilde{z}$ satisfies

$$\tilde{z}^{\alpha} = \frac{1}{1 + \alpha(1 - \tilde{z})},$$

which means that

$$f'(\tilde{z}) = (\alpha + 1)\left[\frac{1}{1 + \alpha(1 - \tilde{z})} - \frac{1}{(1 + \alpha(1 - \tilde{z}))^2}\right] > 0.$$

Accordingly, we have that $f(z) \geq 0$ for all $z \in [\tilde{z}, 1]$, but this contradicts the fact that $f(z) < 0$ for all $z \in (1 - \epsilon, 1)$, and so it must be the case that $f(z) < 0$ for all $z \in (0, 1)$.

As it was previously mentioned, this implies that $B < 0$, which means that z^* is a local maximum. Hence, we conclude that the optimal z^* must belong to $\{0, 1\}$, and from the definition of z we get that $\tau^* \in \{-\log(\underline{q})/(r+\lambda), \infty\}$.

D Exogenous News

Proof of Proposition 5

Proof. Defining $\tilde{m}_t \equiv m_t + \mu$ and $\tilde{M}_t = \int_0^t \tilde{m}_s \mathrm{d}s$ we can write the optimization problem as

$$\mathscr{G}^\theta U = \sup_{\tau, \tilde{m}_t} \int_0^\tau e^{-rt-\tilde{M}_{t-}} \Big(u(x_t^\theta) - \mu c + \tilde{m}_t \mathcal{M}_\theta(U, x_t) \Big) \,\mathrm{d}t + e^{-r\tau - \tilde{M}_\tau} \mathcal{M}_\theta(U, x_\tau)$$

$$\text{subject to}$$

$$\dot{\tilde{M}}_t = \tilde{m}_t$$

$$\dot{q}_t = (r + \lambda + \tilde{m}_t) q_t - \tilde{m}_t$$

$$q_t \geq \underline{q}, \ \forall t \in [0, \tau]$$

$$\mu \leq m_t.$$

This problem is essentially the same as in the one in the absence of exogenous news with the exception that the non-negativity constraint $m_t \geq 0$ is replaced by the lower bound $\tilde{m}_t \geq \mu$. We need to distinguish two cases $(r+\lambda)\frac{\underline{q}}{1-\underline{q}} \geq \mu$ and $(r+\lambda)\frac{\underline{q}}{1-\underline{q}} < \mu$.

Case $(r+\lambda)\frac{\underline{q}}{1-\underline{q}} \geq \mu$. In this case we have that $\dot{q}_t = 0\big|_{q_t = \underline{q}}$ requires that $\tilde{m}_t > \mu$. This means that the solution for $\tilde{m}_t = m_t + \mu$ is given by Propositions 2 and 3.

Case $(r+\lambda)\frac{\underline{q}}{1-\underline{q}} < \mu$. In this case exogenous news alone are sufficient to provide incentives so random monitoring is never part of the optimal monitoring policy. In the linear case, if we set $\tau = \infty$ and $m_t = 0$ we get $q_t = \mu/(r+\lambda+\mu) > \underline{q}$. Clearly, this policy is optimal because the cost of monitoring is zero and there is no direct benefit of monitoring. When $u(\cdot)$ is convex, we argue that it must be the case that $q_t > \underline{q}$, all $t \geq 0$, and so by Lemma 4 $m_t = 0$. Suppose that this is not the case and there is a time $t^\dagger$ such that $q_{t^\dagger} = \underline{q}$, then we have that

$$\dot{q}_{t^\dagger} = (r+\lambda)\underline{q} - \tilde{m}_{t^\dagger}(1-\underline{q}) < (r+\lambda)\underline{q} - \mu(1-\underline{q}) < 0.$$

This means that for some small $\epsilon > 0$ we have $q_{t^\dagger+\epsilon} < \underline{q}$ so the incentive compatibility constraint is violated. Accordingly, it must be the case that $q_t > \underline{q}$ for all $t \geq 0$. □

D.1 Sufficient Conditions with Asymmetric News

The Hamiltonian for the optimal control problem is

$$\begin{aligned}\mathcal{H}(\Pi_t^L, \Pi_t^H, \zeta_t, \nu_t^L, \nu_t^H, \psi_t, m_t, t) = {} & \zeta_t\big((r+m_t)U_t - x_t^\theta - \mu_H x_t^\theta U_H - \mu_L(1-x_t^\theta)U_L - m_t\mathcal{M}_\theta(U, x_t)\\ & + \psi_t(\Pi_t^H - \Pi_t^L - k/\lambda) + \nu_t^H\big((r+\mu_H+m_t)\Pi_t^H - x_t + k\bar{a} + \lambda(1-\bar{a})(\Pi_t^H - \Pi_t^L)\\ & - (\mu_H+m_t)\Pi(H)\big) + \nu_t^L\big((r+\mu_L+m_t)\Pi_t^L - x_t + k\bar{a} - \lambda\bar{a}(\Pi_t^H - \Pi_t^L) - (\mu_L+m_t)\Pi(L)\big)\end{aligned}$$

As before, we have that $\zeta_t- = 1$ and the evolution of the remaining co-state variables is The evolution of the co-state variables is given by

$$\begin{aligned}\dot{\nu}_t^H &= -(\mu_H + \lambda(1-\bar{a}))\nu_t^H - \psi_t + \lambda\bar{a}\nu_t^L\\ \dot{\nu}_t^L &= -(\mu_L + \lambda\bar{a})\nu_t^L + \psi_t + \lambda(1-\bar{a})\nu_t^H.\end{aligned}$$

The switching function $S(t)$ is given by

$$S(t) = \mathcal{M}_\theta(U, x_t) + \nu_t^H(\Pi_t^H - \Pi(H)) + \nu_t^L(\Pi_t^L - \Pi(L)) - U_t.$$

We pin-down the boundary condition for the co-state variables ν_t^θ by looking at the switching function. The rate of monitoring is positive (and finite) at time zero only if $S(0) = 0$ which implies that

$$0 = \mathcal{M}_\theta(U, \theta) - U_\theta + \nu_0^H(\Pi_0^H - \Pi(H)) + \nu_0^L(\Pi_0^L - \Pi(L)).$$

If the incentive compatibility constraint is binding at time zero, so $\Pi_0^H - \Pi_0^L = k/\lambda$, then when $\theta_0 = L$ and $m_0 > 0$ the initial value of the co-state variable ν_0^H is

$$c = -\nu_0^H \left(\frac{1}{r+\lambda} - \frac{k}{\lambda} \right).$$

The initial value of the co-state variable ν_0^L is determined by the transversality condition $\lim_{t\to\infty} \nu_t^L = \nu_{ss}^L$. If the incentive compatibility constraint at time zero were slack (that is $m_0 = 0$) then the initial value would be $\nu_0^H = 0$. The determination of ν_0^L is more complicated in this latter case as ν_t^L can jump at the junction time τ^m in which the IC constraint becomes binding. Similarly, if $\theta = H$ then we have that ν_0^L is given by

$$c = \nu_0^L \left(\frac{1}{r+\lambda} - \frac{k}{\lambda} \right)$$

while ν_0^H is determined by the transversality condition $\lim_{t\to\infty} \nu_t^H = \nu_{ss}^H$. As for $\theta_0 = L$, the same qualification for the case in which the IC constraint is slack at time zero applies. In the same way as we did in the case without news, we can use the condition that the switching function is constant on a singular arc, $\dot{S}_t = 0$, to back out the value of the Lagrange multiplier ψ_t

$$\begin{aligned}\psi_t\big((\Pi_t^H - \Pi_t^L) - (\Pi(H) - \Pi(L))\big) = \dot{x}_t^\theta(U_H - U_L) - \dot{U}_t + (-(\mu_H + \lambda(1-\bar{a}))\nu_t^H + \lambda\bar{a}\nu_t^L)(\Pi_t^H - \Pi(H)) + \nu_t^H\dot{\Pi}_t^H \\ + (-(\mu_L + \lambda\bar{a})\nu_t^L + \lambda(1-\bar{a})\nu_t^H)(\Pi_t^L - \Pi(L)) + \nu_t^L\dot{\Pi}_t^L\end{aligned}$$

If the incentive compatibility constraint is binding, $\Pi_t^H - \Pi_t^L = k/\lambda$, then we can write the Lagrange multiplier as

$$\begin{aligned}\psi_t = \frac{1}{k/\lambda - \Delta}\Big[\dot{x}_t^\theta(U_H - U_L) - \dot{U}_t - \big(\mu_H\nu_t^H + \mu_L\nu_t^L\big)(\Pi_t^L - \Pi(L)) + \big((\mu_H + \lambda(1-\bar{a}))\nu_t^H - \lambda\bar{a}\nu_t^L\big)\left(\frac{1}{r+\lambda} - \frac{k}{\lambda}\right) \\ +(\nu_t^L + \nu_t^H)\dot{\Pi}_t^L\Big].\end{aligned}$$

Given that the maximized Hamiltonian is linear in the sate variables, a sufficient condition for our conjectured monitoring policy m_t to be optimal is that the Lagrange multiplier ψ_t is non-negative whenever the incentive compatibility constraint is binding. The monitoring policy m_t is positive if and only if this constraint is binding; hence, the sufficiency condition reduces to verify that $\psi_t m_t \geq 0$. Given the higher dimensionality of the state space, we can no longer check this condition analytically. However, this condition can be easily verified numerically after solving for the system of ODEs.

Proof of Proposition 6

Proof. Looking at the phase diagram in Figure 3 we see that if the optimal solution is given by the saddle path then the trajectory towards the steady state is monotonic which implies that m_t is decreasing in x_t. We show next that the trajectory to the stable steady state violates the non-negativity condition of the monitoring rate. The roots of the equation for the steady state are

$$\frac{-(r + \mu_L + \alpha - \beta\Pi(L)) \pm \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta}}{2\beta}.$$

Let's denote by Π_-^L and Π_+^L the smaller and larger solution to the quadratic equation (32), respectively. We show next that only one of this roots is consistent with $m_t \geq 0$.

Claim 1 (Bad News). *If* $\mu_L > \mu_H$ *then*

$$\alpha + \beta \Pi_+^L < 0.$$

Given that we are in the bad news cas+e, $m_t > 0$ only if $\Pi_t < -\alpha/\beta$. When $\mu_L > \mu_H$, the larger root Π_+^L is

$$\begin{aligned}\Pi_+^L &= \frac{r + \mu_L + \alpha - \beta\Pi(L) + \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 - 4((\mu_L + \alpha)\Pi(L) + x_{ss})(-\beta)}}{-2\beta} \\ &> \frac{2(r + \mu_L + \alpha - \beta\Pi(L)) + 2\sqrt{((\mu_L + \alpha)\Pi(L) + x_{ss})(-\beta)}}{-2\beta} \\ &= -\frac{\alpha}{\beta} + \frac{r + \mu_L - \beta\Pi(L)}{-\beta} + \frac{\sqrt{((\mu_L + \alpha)\Pi(L) + x_{ss})(-\beta)}}{-\beta} \\ &> -\frac{\alpha}{\beta}.\end{aligned}$$

Hence, in the bad news case only the trajectory towards the saddle point is consistent with $m_t > 0$.

Claim 2 (Good News). *If* $\mu_L < \mu_H$ *then*

$$\alpha + \beta \Pi_-^L < 0.$$

In the good news case, $m_t > 0$ only if $\Pi_t > -\alpha/\beta$. The smaller root is

$$\Pi_-^L = \frac{-(r + \mu_L + \alpha - \beta\Pi(L)) - \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta}}{2\beta}$$

If $\Pi_-^L \leq 0$ then there is nothing to prove as the payoff of the firm cannot be negative. Accordingly, let's restrict attention to parameters such that $\Pi_-^L > 0$. We have that $\Pi_-^L > 0$ if and only if

$$(r + \mu_L - \beta\Pi(L)) + \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta} < -\alpha$$

Monitoring is positive at iff $\Pi_-^L > -\alpha/\beta$ which requires

$$(r + \mu_L - \alpha + \beta\Pi(L)) + \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta} < 0$$

We consider two separate cases:

Case $\alpha \leq 0$ Using the condition for $\Pi_-^L > 0$ we get the inequality

$$\begin{aligned}r + \mu_L - \alpha + \beta\Pi(L) + \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta} > \\ 2(r + \mu_L + \beta\Pi(L)) - \alpha + 2\sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta} > 0\end{aligned}$$

which contradicts the condition for positive monitoring $\Pi_-^L > -\alpha/\beta$.

Case $\alpha > 0$ If $(r + \mu_L + \alpha - \beta\Pi(L)) > 0$ then we get an immediate contradiction with the hypothesis that $\Pi_-^L > 0$. Hence, assume that $(r + \mu_L + \alpha - \beta\Pi(L)) < 0$. For any $b > 0$ and $a < 0$ we have the following inequality

$$\sqrt{a^2 + b} > |a| \Rightarrow -a - \sqrt{a^2 + b} < -a - |a| = 0.$$

If $\alpha > 0$ then we have $4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta > 0$. Setting $a = (r + \mu_L + \alpha - \beta\Pi(L)) < 0$ and $b = 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta > 0$ in the previous inequality we get

$$\Pi_-^L = \frac{-(r + \mu_L + \alpha - \beta\Pi(L)) - \sqrt{(r + \mu_L + \alpha - \beta\Pi(L))^2 + 4((\mu_L + \alpha)\Pi(L) + x_{ss})\beta}}{2\beta} < 0,$$

which yields a contradiction to $\Pi_-^L > 0$. □

References

Aliprantis, C. D. and K. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Bassetto, M. and C. Phelan (2008). Tax riots. *The Review of Economic Studies 75*(3), 649–669.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The Economic Dimensions of Crime*, pp. 13–68. Springer.

Board, S. and M. Meyer-ter Vehn (2013). Reputation for quality. *Econometrica 81*(6), 2381–2462.

Bond, P. and K. Hagerty (2010). Preventing crime waves. *American Economic Journal: Microeconomics 2*(3), 138–159.

Cesari, L. (2012). *Optimization–Theory and Applications: Problems with Ordinary Differential Equations*, Volume 17. Springer.

Dewatripont, M., I. Jewitt, and J. Tirole (1999). The economics of career concerns, part II: Application to missions and accountability of government agencies. *The Review of Economic Studies 66*(1), 199–217.

Eeckhout, J., N. Persico, and P. Todd (2010). A theory of optimal random crackdowns. *The American Economic Review 100*(3), 1104–1135.

Ekmekci, M. (2011). Sustainable reputations with rating systems. *Journal of Economic Theory 146*(2), 479–503.

Fernandes, A. and C. Phelan (2000). A recursive formulation for repeated agency with history dependence. *Journal of Economic Theory 91*(2), 223–247.

Graham, J. R., C. R. Harvey, and M. Puri (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics 115*, 449–470.

Harris, M. and A. Raviv (1996). The capital budgeting process: Incentives and information. *The Journal of Finance 51*(4), 1139–1174.

Harris, M. and A. Raviv (1998). Capital budgeting and delegation. *Journal of Financial Economics 50*(3), 259 – 289.

Hart, O. and J. Moore (1995). Debt and seniority: An analysis of the role of hard claims in constraining management. *The American Economic Review 85*(3), 567–585.

Hartl, R. F., S. P. Sethi, and R. G. Vickson (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM review 37*(2), 181–218.

Horner, J. and N. S. Lambert (2016). Motivational ratings.

Kothari, S. P., S. Shu, and P. D. Wysocki (2009). Do managers withhold bad news? *Journal of Accounting Research 47*(1), 241–276.

Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, 1029–1061.

Mailath, G. J. and L. Samuelson (2001). Who wants a good reputation? *The Review of Economic Studies 68*(2), 415–441.

Mookherjee, D. and I. Png (1989). Optimal auditing, insurance, and redistribution. *The Quarterly Journal of Economics 104*(2), 399–415.

Motta, A. (2003). Managerial incentives and internal capital markets. *The Journal of Finance 58*(3), 1193–1220.

Polinsky, A. M. and S. Shavell (1984). The optimal use of fines and imprisonment. *Journal of Public Economics 24*(1), 89–99.

Reinganum, J. F. and L. L. Wilde (1985). Income tax compliance in a principal-agent framework. *Journal of Public Economics 26*(1), 1–18.

Seierstad, A. and K. Sydsaeter (1977). Sufficient conditions in optimal control theory. *International Economic Review*, 367–391.

Seierstad, A. and K. Sydsaeter (1986). *Optimal Control Theory with Economic Applications*. Elsevier North-Holland.

Stein, J. C. (1997). Internal capital markets and the competition for corporate resources. *The Journal of Finance 52*(1), 111–133.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic theory 21*(2), 265–293.

Trussler, M. and S. Soroka (2014). Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics 19*(3), 360–379.

Zhang, Y., S. Farmer, J. Gong, M. Satterthwaite, T. Sarver, T.-h. Sng, et al. (2011). Are two report cards better than one? the case of cabg surgery and patient sorting. *The Case of CABG Surgery and Patient Sorting (September 28, 2011)*.