Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring *

Felipe Varas † Iván Marinovic‡ Andrzej Skrzypacz §

November 30, 2016

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 benefits from having a 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 when the principal discloses the outcome of inspections to the public, 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 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?

These seem to be important questions for the provision of 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 moral hazard is a concern.

*We thank Guy Barles, Simon Board, Douglas Carmichael, Moritz Meyer-ter-Vehn (discussant) and Bob Wilson for helpful feedback and suggestions, and seminar participants at Stanford, FTG (Princeton), and the Columbia/Duke/MIT/Northwestern IO Theory Conference for helpful comments.

†Duke University, Fuqua School of Business. email: felipe.varas@duke.edu
‡Stanford University, GSB. email: imvial@stanford.edu
§Stanford University, GSB. email: skrz@stanford.edu

To develop economic intuition about these related problems, understand the role of monitoring, and its design, 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 firm's private effort. The firm cares about its reputation, which is defined as the public belief about the firm's current quality, and the principal cares about quality, which the principal can learn by performing a costly inspection, namely by monitoring the firm quality and disclosing the outcome of the inspections to the public. To understand the effects of the information channel alone, we assume away any explicit fines or bonuses levied on the firm 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 the applications that motivate this paper, there are two main reasons driving 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 the principal uses the information to better allocate resources, then information has direct value.[1]

Most real-life monitoring policies fall into one of two categories: random inspections/tests or deterministic inspections taking place at pre-determined dates. At first, neither of these policies seems optimal. A policy of deterministic inspections runs the risk of inducing "window dressing" by the firm to pass the review, while resting on its laurels thereafter: the firm has strong incentives to put effort as it approaches the inspection date, and weak incentives right after the inspection, given the firm knows it will not be inspected in the near future. On the other hand, if quality is persistent, a random policy that assigns a positive hazard rate of inspections right after a review seems unnecessary and wasteful, because after a review there is almost no uncertainty about the firm's quality.

Our analysis explains why these two strategies can be part of the optimal monitoring policy. The intuition is that the optimal policy depends on the role that monitoring plays: random monitoring tends to be optimal when monitoring is mostly driven by incentive provision, while deterministic monitoring tends to be optimal when it is mostly driven by information acquisition considerations, namely the principal's own resource allocation goals. Somewhat surprisingly, we show that 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 deterministic testing after bad results and random testing after good results, or vice versa, but the optimal policy never combines these monitoring modes in between two inspections). We characterize how the parameters of the problem affect the optimal choice between these two extreme policies. For example, deterministic inspections are more likely to be optimal when the cost of effort is low, and random inspections are more likely to be optimal when the cost of inspections is high. Also, the choice between random and deterministic monitoring is determined by the persistence of shocks: highly persistent outcomes are more likely to lead to random inspections, while transitory shocks are more likely to lead to deterministic inspections. These dynamics are driven by the relative importance of incentives versus information acquisition (learning).

In our benchmark model, the only available information are the outcomes of the inspections observed over time. In this setting, we show that, when the optimal policy is random, then the monitoring intensity is constant over time. This result depends, however, on the absence of alternative sources of information.

[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.

Indeed, to understand the scope of this result, we extend the model to allow for exogenous news (as provided, in practice, by consumer reviews or market pundits). Following Board and Meyer-ter Vehn (2010), we consider a setting where conditional on quality good and bad news are exogenous, but may arrive with different intensities. This asymmetry may capture, in reduced form, the fact that bad news are often delayed by firms because of incentive reasons (by contrast, in some settings, media pundits are particularly interested in disclosing negative news, so bad news are observed earlier than good news). In this context, the severity of moral hazard depends on the firm's reputation, but the sign of this relationship depends on the rate of arrival of bad news vis a vis good news. When bad news arrive faster than good news, the moral hazard problem is more acute when the agent's reputation is low, and vice-versa. This asymmetry alters the design of the monitoring system: if bad news arrive faster, the optimal monitoring policy calls for high monitoring intensity when the agent's reputation is low. Hence, after a negative news arrival, the monitoring intensity is increased to mitigate the strong shirking incentives arising after a negative news arrival.

1.1 Related Literature

There is a large empirical literature on the importance of quality monitoring and reporting systems. For example, Epstein (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 (Jin and Leslie, 2009), school report cards (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 (Dranove et al., 2003), health plan report cards (Dafny and Dranove, 2008), hospital rankings (e.g., Pope (2009)), nursing homes report cards (e.g., Feng Lu (2012)), fertility clinics report cards (Bundorf et al., 2009). 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 et al. (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 Holmström (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 Dilmé et al. (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 analysis in is quite different.

2 Setting

Agents, Technology and Effort: There are two players: a principal and an firm. Time $t \in [0, \infty)$ is continuous. The firm has a product whose quality changes over time. We model the evolution of quality as in 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, and subsequently quality evolves over time and is determined by the firm's effort. At each time t, the firm makes a private effort decision $a_t \in [0, \bar{a}]$, $\bar{a} < 1$, and changes in quality are not publicly observed and only observed by the firm. Throughout most of the paper we assume that when the firm chooses effort a_t, quality switches from low to high with intensity λa_t, and from high to low quality with intensity $\lambda(1 - a_t)$. Later in the paper we illustrate how the analysis can be extended to the case in which quality can take a continuum of levels, and in which 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.[2] 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 specifies an increasing sequence of dates $(T_n)_{n\geq 1}$ at which the principal inspects quality.[3] Letting $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the natural filtration with respect to θ_t, a monitoring policy defines a new filtration $\mathbb{F}^P = (\mathcal{F}_t^P)_{t\geq 0}$ such that $\mathcal{F}_t^P = \mathcal{F}_{T_{N(t)}}$, where $N(t) \equiv \sup\{n : T_n \leq t\}$. A monitoring policy is represented by a sequence of cumulative density functions, $F_n(t) = \Pr\left(T_n \leq t \,|\, \mathcal{F}_t^P\right)$; that is, the distribution of the n-th monitoring (or equivalently, the time elapsed between T_{n-1} and T_n) must be conditional on the information up to the last review date T_{n-1}.

We assume that current quality is always privately known by the firm, but as discussed below, our results extend to the case where the firm does not privately observe quality, which in some applications is more realistic. A strategy for the firm is an effort plan $a = (a_t)_{t\geq 0}$ that is predictable with respect to $\mathbb{F}$.

Reputation and Payoffs: We model the firm's payoffs in a reduced form as driven by the firm's reputation. In particular, denote the market's beliefs about the firm's effort strategy by $\tilde{a} = (\tilde{a}_t)_{t\geq 0}$. Reputation at time t is given by $x_t \equiv E^{\tilde{a}}(\theta_t|\mathcal{F}_t^P)$ 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 the firm's current quality. It evolves based on the market's conjecture about firm's strategy and in response to monitoring outcomes.

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

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

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

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

Note that we do not include the cost of effort in the principal's payoff. In some applications it may be more natural to assume the principal internalizes that cost and then we would subtract $-k\tilde{a}_s$ from the welfare flows. However, since 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 (of course the cost still matters indirectly for the incentive reasons, since it affect the agent's effort incentives).

[2] Board and Meyer-ter Vehn (2010) also considers the case with $\bar{a} < 1$.

[3] 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.

[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.

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 incentive provision value). 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 improve matching of firms and consumers by allowing 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 $(F_n)_{n\geq 1}$ if it is consistent with the firm's optimization given that policy and market beliefs are correct on the equilibrium path, namely $\tilde{a} = a$. In other words, a is incentive compatible if a is consistent with equilibrium beliefs given $(F_n)_{n\geq 1}$:[5]

Definition 1. *Fix a monitoring policy $(F_n)_{n\geq 1}$. 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 $(F_n)_{n\geq 1}$ and $\tilde{a}$.*
2. *a maximizes Π.*
3. *$\tilde{a} = a$.*

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

We assume the principal commits to a policy at the start and the firm chooses full effort whenever there exists an equilibrium given $(F_n)_{n\geq 1}$ that implements full effort. (So that, as usual in contract theory, if there were multiple equilibria given $(F_n)_{n\geq 1}$, we select the one inducing full effort whenever possible).

Definition 2. *A monitoring policy $(F_n)_{n\geq 1}$ is incentive compatible if for that policy there exists an equilibrium with $a_t = \bar{a}$. We call a monitoring policy 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), and second, since the principal values information per se (when $u(\cdot)$ is strictly convex) the policy solves the real-option-information-acquisition problem of deciding when to incur the cost c to learn the firm's current quality and thus benefit from superior information.

2.1 Examples

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

[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 $(F_n)_{n\geq 1}$. We hope our simpler definition does not create confusion.

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 (excluding 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, the principal's preferences are linear in beliefs so information has no social value beyond incentive provision; information only affects the dispersion of prices over time. Hence, the optimal policy minimizes the cost of monitoring subject to 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 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 evolving quality. The two schools are located at opposite extremes of the unit line. Evolution of quality of school B depends on the school's hidden effort/investment and is unobservable to the public 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 location 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 a transfer of \$1 from the government and that marginal costs are normalized to zero. Hence, the profit flows of schools A and B are

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

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 firms' profits does not depend on it (since the two schools just split the subsidy per student, reputation has only distributive effects). The convexity of u reflects here that better information about the quality of 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 increasing investment. In particular, assume the managers' 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's 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 allocate to the division incurring a flow cost of $\iota^2/2$. 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:

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

[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.

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 $\pi_t = b\mu x_t - k a_t$, as in our general model. In the baseline model we assume that monitoring is the only source of information about θ available to the CFO. In this application it is natural to assume that the CFO also learns about the current productivity when the cash-flows arrive. We study the possibility of exogenous news arrivals in Section 5.

Throughout the paper we ignore the use of monetary transfers beyond transfers that are proportional to the current reputation[9]. In some settings, other forms of performance-based compensation are used to provide incentives, but in many cases divisional contracts are simple and earnings proportional to the size of the division, may be the main driver of the manager's incentives. Graham, Harvey, and Puri (2015) find evidence that manager's reputation has an 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 where autonomy, control and capital allocation driven by career concerns seem more preponderant for worker's motivation.

3 Optimal Monitoring without Agency Problems

As a first step toward characterizing the optimal policy in the general model, we begin by considering a relaxed problem that ignores the agent's incentive constraint. When information has value to the principal (i.e., when $u(\cdot)$ is convex), the cost of monitoring c is small enough, and the cost of effort k is small enough, then the solution of such a relaxed problem satisfies the incentive constraints, being the optimal monitoring policy –the solution of the relaxed problem also characterizes the optimal policy when effort (but not quality) is observable.

Consider the evolution of reputation between inspection dates. Given that the firm puts in 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

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

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

The optimal policy is Markovian with reputation being the state variable.If we let $\mathcal{A}$ be the set of reputa-

[8] Note that the allocation in period t is made before the realization of the cash-flow (the Poisson process), as captured by $\mathcal{F}_{t-}^P$. 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.

tions 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 $\underline{x} \le \bar{a} \le \overline{x}$ 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

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

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 $x_t \in \mathcal{A} = [\underline{x}, \overline{x}]$.*

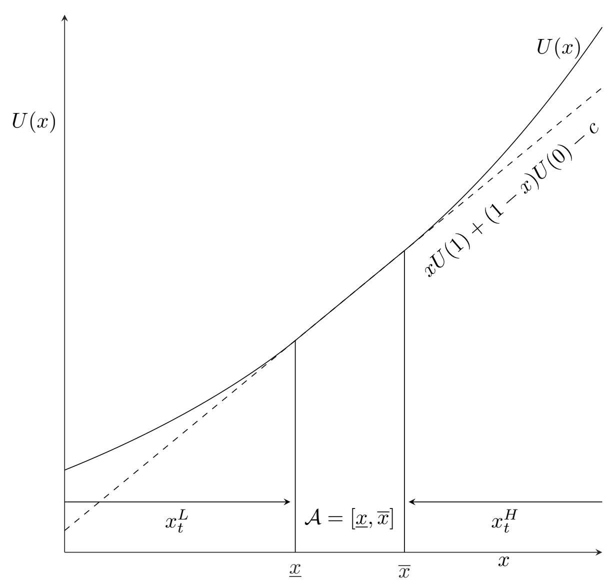


Figure 1: Value Function. The optimal policy requires to monitor whenever $x_t \in \mathcal{A}$.

Figure 1 illustrates the principal's payoff as a function of 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 $\bar{a}$ (but of course in this case the incentive constraint would be violated since 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, as mentioned above, 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 upon the time since last review, $t - T_{n-1}$, 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]

Remark 1. *The monitoring policy characterized in the benchmark corresponds to the case in which both τ_L and τ_H are finite. Depending on the specific parameters of the model, it could be the case that either τ_L or τ_H are infinite, or in other words there is no further monitoring after some outcomes. In terms of the Markov policy, this means that either $\underline{x} = \bar{a}$ or $\overline{x} = \bar{a}$. In this case, the value matching and smooth pasting conditions are only valid at a threshold that is different from $\bar{a}$.*

4 Optimal Monitoring with Agency Problems

The monitoring policy previously discussed may violate incentive constraints when the policy is such that reviews are too infrequent. This may lead the firm to shirk during some time, particularly after an inspection when the agency conflict is most severe. In this section, we solve the constrained problem that explicitly incorporates incentive constraints. As discussed above, we study the optimal monitoring policy among those that implement full effort.

[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.

4.1 Incentive Compatibility

The first step (towards finding an optimal monitoring policy) is to characterize necessary and sufficient conditions for a policy to satisfy the incentive compatibility constraints in every instance. Let's start by considering the firm's continuation payoff under full effort from time T_n onward, where T_n is the next review date.

$$\begin{aligned}\Pi_{T_n} &= E^{\bar{a}}\left[\int_{T_n}^{\infty} e^{-r(t-T_n)}(x_t - k\bar{a})\mathrm{d}t \Big| \mathcal{F}_{T_n}\right] \\ &= \int_{T_n}^{\infty} e^{-r(t-T_n)}\big(E^{\bar{a}}[x_t|\mathcal{F}_{T_n}] - k\bar{a}\big)\mathrm{d}t.\end{aligned}$$

This expression corresponds to the firm continuation value at time T_n assuming the firm exerts full effort, and represents the present value of the firm future revenues net of effort costs. A key insight in the derivation of the incentive compatibility constraint is that the law of iterated expectations along with the Markov nature of the quality process, θ_t, imply that $E^{\bar{a}}[x_t|\mathcal{F}_{T_n}] = E^{\bar{a}}[\theta_t|\theta_{T_n}]$, and it is equal to

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

This means that, in any incentive-compatible monitoring policy, and if the quality at time T_n is public, then the firm's continuation value at time T_n is given by:

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

Because the continuation value at time T_n is independent of the previous history of effort (it depends on effort only indirectly via θ_{T_n}) we can invoke the one-shot deviation principle to derive the agent's incentive compatibility constraint. For any effort strategy a_t, we can write the process of 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, corresponding to the compensated Poisson process of changes in quality.

Consider the agent's effort incentives. The effect of effort on future quality is $\partial\theta_{T_n}/\partial a_t = \lambda e^{-\lambda(T_n - t)}\mathrm{d}t$ so the marginal benefit of exerting effort over an interval of size dt is $E_t[\lambda e^{-(r+\lambda)(T_n-t)}](\Pi(H) - \Pi(L))\mathrm{d}t$. 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_n, 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 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 2. *Lemma 1 holds for any quality process satisfying the stochastic differential equation*

$$\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \mathrm{d}Z_t, \tag{6}$$

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. Note that our binary setting is a particular case of (6) *in which Z_t is a compensated Poisson 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. That is, q_t, represents the expected discounted time till the next review, where the effective discount rate incorporates the depreciation rate λ. Letting

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

the incentive compatibility constraint in Lemma 1 becomes $q_t \geq \underline{q}$. The state variable q_t is the implicit discount rate the firm uses to assess the benefit of having high quality at the review date. It depends upon the exogenous r and λ and the endogenous distribution of monitoring times.

The next step is to derive the law of motion of $(q_t)_{t\geq 0}$ to use it as a state variable in the principal's optimization problem. Given the recursive structure of the problem, it is convenient to specify the problem in terms of the time elapsed since the last monitoring time, which we denote by τ (that is, $\tau = t - T_{n-1}$), and represent the monitoring policy by a distribution $F(\tau)$ such that $F(\tau) = \Pr(T_n \leq T_{n-1} + \tau | \mathcal{F}^P_{T_{n-1}+\tau})$. To avoid technical issues related to continuous time, we restrict the principal to use cumulative distribution functions that $F(\tau)$ can be expressed by a triple $(m_\tau, p_\tau, \bar{\tau})$ as[11]

$$1 - F(\tau) = \begin{cases} e^{-\int_0^\tau m_s \mathrm{d}s} \prod_{0<s\leq\tau}(1-p_s) & \text{if } \tau < \bar{\tau} \\ 0 & \text{if } \tau \geq \bar{\tau} \end{cases}. \tag{7}$$

The triple $(m_\tau, p_\tau, \bar{\tau})$ consists of the following: $\bar{\tau} = \inf\{\tau > 0 : F(\tau) = 1\}$ is the deterministic review time at which the firm is monitored for sure if it has not been monitored before. At times where $F(\tau)$ is (absolutely) continuous, the hazard rate of $F(\cdot)$ is given by $\mathrm{d}F(s)/(1-F(\tau)) = m_s$ for some function $m : [0,\infty) \to [0,\infty)$. Finally, at any point of discontinuity at $\tau < \bar{\tau}$, $p_\tau = \mathrm{d}F(\tau)/(1-F(\tau)) = \frac{F(\tau)-F(\tau^-)}{1-F(\tau)}$ is the probability of monitoring at that time conditional on not having monitored earlier.

With this representation of $F(\tau)$ we we can write q_τ as

$$q_\tau = \int_\tau^{\bar{\tau}} e^{-(r+\lambda)(s-\tau)-\int_\tau^s m_u \mathrm{d}u} m_s \mathrm{d}s + \sum_{s\in(\tau,\bar{\tau})} e^{-(r+\lambda)(s-\tau)} \frac{\mathrm{d}F(s)}{1-F(\tau^-)} + e^{-(r+\lambda)(\bar{\tau}-\tau)} \frac{1-F(\bar{\tau})}{1-F(\tau)}$$

Wherever $F(\tau)$ is absolutely continuous, we can differentiate the previous equation to get the following differential equation for q_τ :

$$\dot{q}_\tau = (r + \lambda + m_\tau)q_\tau - m_\tau. \tag{8}$$

[11] We do not allow $F(\tau)$ that are not absolutely continuous almost everywhere, like the Cantor function.

At any point of discontinuity, we have

$$\mathrm{d}q_\tau = -\frac{p_\tau}{1-p_\tau}(1-q_{\tau^-}). \tag{9}$$

Note that in order to satisfy incentive compatibility right after τ we must have that $p_\tau \leq (q_{\tau^-} - \underline{q})/(1-\underline{q})$ (otherwise we would get $q_{\tau^+} < \underline{q}$).

Combining equations (8) and (9) we get the following representation for incentive compatible monitoring policies.

Proposition 1 (Incentive Compatibility)**.** *Consider a monitoring policy described by $(m_\tau, p_\tau, \bar{\tau})$, and for any $\tau \in [0, \bar{\tau}]$, let q_τ be the solution to*

$$q_\tau = q_0 + \int_0^\tau [(r+\lambda+m_s)q_s - m_s]\mathrm{d}s - \sum_{s\leq\tau} \frac{p_s}{1-p_s}(1-q_{s^-}),\ q_{\bar{\tau}} = 1. \tag{10}$$

Full effort is incentive compatible if and only if $q_\tau \geq \underline{q}$, for all $\tau \in [0, \bar{\tau}]$, where

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

Incentive compatibility imposes a lower bound on q_τ or an upper bound on the expected next review date. What matters for incentives at a given point is not necessarily the monitoring intensity at that moment but the cumulative discounted likelihood of monitoring in the near future: the reason is that effort has a persistent effect on quality.

The significance of Proposition 1 is that it allows us to formulate the optimal monitoring policy as a recursive problem, with q_τ being the state variable, and use the tools of optimal control theory to study the optimal policy.

4.2 Principal's Problem

We now develop a recursive formulation for the principal's optimization problem that can be tackled using optimal control theory. The principal solves the following optimization problem:

$$U(x_0) = \sup_{(T_n)_{n\geq 1}} E\left[\int_0^\infty e^{-rs}u(x_s)\mathrm{d}s - \sum_{T_n} e^{-rT_n}c \,\Big|\, \mathcal{F}_0^P\right] \tag{11}$$

$$\text{subject to:} \tag{12}$$

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

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

The only difference between this problem and that stated in Section 3 is the presence of the incentive constraint.

In Section 3 we found the optimal policy in the relaxed problem, by working with the space of beliefs: that is, we specified the monitoring policy as a function of beliefs. 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 is more convenient when dealing with incentive constraints, as we do here.

The analysis differs depending on whether information has value to the principal (social value). Next, we study the case when it does not.

4.3 Linear Payoffs: Information without Social Value

We begin by analyzing the case in which the principal's flow payoff $u(\cdot)$ is linear. As discussed above, this case could capture applications in which the principal is an industry self-regulatory organization that is not directly concerned about consumer surplus but wishes to maximize the industry's overall profits.

Under linear payoffs, information has no social value, hence the only role of monitoring is incentive provision. The principal's problem is equivalent to minimizing the monitoring cost subject to the incentive compatibility constraint. Accordingly, we can reduce the principal's problem to the following cost minimization problem:

$$C_0 = \min_{(T_n)_{n\geq 1}} E\left[\sum_{T_n} e^{-rT_n} c \Big| \mathcal{F}_0^M\right] \tag{13}$$

$$\text{subject to:} \tag{14}$$

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

As we shall show, in this case, the optimal monitoring policy prescribes a constant monitoring intensity (hazard rate) and no deterministic reviews. We conjecture (and later verify) two results: i) the hazard rate is positive if and only if the incentive compatibility constraint binds and ii) the incentive compatibility constraint always binds. The intuition for the latter conjecture is that otherwise the principal could save some monitoring expenses without affecting the firm's incentives.

For the incentive compatibility constraint to bind all the policy should not include a final review date, formally $\bar{\tau} = \infty$. Or else, if $\bar{\tau} < \infty$, then the incentive compatibility condition would be slack somewhere close to $\bar{\tau}$. Hence, we can find the optimal monitoring rate at any time τ using the fact that $\dot{q}_\tau = 0$, along with the condition $q_\tau = \underline{q}$ for all τ. This leads to $m_\tau = m^* = (r+\lambda)\underline{q}/(1-\underline{q})$. We have the following proposition.

Proposition 2. *If $u(x_t) = x_t$, then the optimal monitoring policy is a Poisson process with arrival rate*

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

Proof. Let T be the first monitoring time so the principal cost at time zero satisfies the recursion

$$C_0 = E_0[e^{-rT}](c + C_0)$$

and the incentive compatibility constraint at time zero is

$$E_0[e^{-(r+\lambda)T}] \geq \underline{q}$$

We show that if there is any time τ such that the incentive compatibility constraint is slack, then we can find a new policy that satisfies the incentive compatibility constraint and yields a lower cost to the principal. In fact, it is enough to show if that the IC constraint is slack at some time $\tilde{\tau}$ then we can find an alternative policy that leaves $E_0[e^{-(r+\lambda)T}]$ unchanged at time zero, remains IC at $\tau > 0$ and reduces $E_0[e^{-rT}]$. We only consider the case in which there is a positive density just before $\tilde{\tau}$ as the argument for the case in which there is an atom at $\tilde{\tau}$ and zero probability just before $\tilde{\tau}$ is analogous. Suppose the IC constraint is slack at time $\tilde{\tau}$ and let $\tau^\dagger = \sup\{\tau < \tilde{\tau} : \text{IC constraint binds}\}$: such a date must exist as otherwise we could increase the date of monitoring and still satisfy the IC constraint. Moreover, we can assume without loss of generality that $\tau^\dagger = 0$. Suppose the monitoring distribution $F(\tau)$ is such that $f(\tau) > 0$ for some interval $(\tilde{\tau} - \epsilon, \tilde{\tau})$, then we can find small ϵ_0 and η and construct an alternative monitoring distribution $\hat{F}(\tau)$ that coincides with $F(\tau)$ outside the intervals $(0, \epsilon_0)$ and $(\tilde{\tau} - \epsilon_0, \tilde{\tau} + \epsilon_0)$. For any $\tau \in (\tilde{\tau} - \epsilon_0, \tilde{\tau})$ the density of the alternative policy is

$$\hat{f}(\tau) = f(\tau) - \eta,$$

while for $\tau \in (0, \epsilon_0)$ is

$$\hat{f}(\tau) = f(\tau) + \alpha\eta$$

and for $\tau \in (\tilde{\tau}, \tilde{\tau} + \epsilon_0)$ is

$$\hat{f}(\tau) = f(\tau) + (1 - \alpha)\eta$$

We can pick $\alpha \in (0, 1)$ such that IC constraint is not affected at $\tau = 0$, that is $\alpha \in (0, 1)$ satisfies

$$\alpha \int_0^{\epsilon_0} e^{-(r+\lambda)\tau} \mathrm{d}\tau + (1 - \alpha) \int_{\tilde{\tau}}^{\tilde{\tau}+\epsilon_0} e^{-(r+\lambda)\tau} \mathrm{d}\tau - \int_{\tilde{\tau}-\epsilon_0}^{\tilde{\tau}} e^{-(r+\lambda)\tau} \mathrm{d}\tau = 0,$$

and we can pick ϵ_0 and η small enough so that the IC constraint still holds for all $\tau > 0$. Because the IC constraint is not affected at $\tau = 0$ we have that

$$\int_0^\infty e^{-(r+\lambda)\tau} \mathrm{d}F(\tau) = \int_0^\infty e^{-(r+\lambda)\tau} \mathrm{d}\hat{F}(\tau),$$

and defining the random variable $z \equiv e^{-(r+\lambda)\tau}$, and letting G and $\hat{G}$ be the respective CDFs of z, we get that

$$\int_0^1 z \mathrm{d}G(z) = \int_0^1 z \mathrm{d}\hat{G}(z).$$

By construction $G(z)$ and $\hat{G}(z)$ have same mean and cross only once which means that $\hat{G}(z)$ is a mean preserving spread of $G(z)$. Noting that

$$\int_0^\infty e^{-r\tau} \mathrm{d}F(\tau) = \int_0^1 z^{\frac{r}{r+\lambda}} \mathrm{d}G(z),$$

where $z^{r/(r+\lambda)}$ is a strictly concave function, and using the fact that $\hat{G}(z)$ is a mean preserving spread of $G(z)$, we immediately conclude that

$$\int_0^1 z^{\frac{r}{r+\lambda}} \mathrm{d}\hat{G}(z) < \int_0^1 z^{\frac{r}{r+\lambda}} \mathrm{d}G(z),$$

and so the monitoring distribution $\hat{F}(\tau)$ yields a lower cost of monitoring: This contradicts the optimality of $F(\tau)$ and implies that the optimal policy must be such the IC constraint binds at all time, and so it is given by a constant monitoring rate m^*. □

Remark 3. *It follows directly from the proof of Proposition 2 that the result extends to the case in which the principal and the Agent have different discount rates as long as the principal is patient enough. If* *the principal has a discount rate ρ, then Proposition 2 still holds as long as $\rho < r + \lambda$. If the principal is sufficiently impatient, that is if $\rho > r + \lambda$, then the optimal policy in the linear case involves purely deterministic monitoring.*

4.4 Convex Payoffs: Information with Social Value

In many applications the principal may derive benefits from the information for reasons other than incentive provision. We represent this case by assuming the principal's utility $u(x)$ is convex in beliefs. In this context the principal designs the monitoring system facing a trade off between cost minimization and information acquisition so we cannot use a simple argument as in the linear case; we need to analyze the full optimization problem.

We can reformulate the problem as an optimal control problem with state constraints. To solve it we adopt the following approach: we use the necessary conditions from the maximum principle to show that the optimal policy is recursive and can be characterized by a final review date $\bar{\tau}$. We thus reduce the problem to a one dimensional optimization problem and we then maximize over $\bar{\tau}$ directly.

As we just mentioned, the first step to formulate the problem as an optimal control problem. We start by introducing some notation. Let $x_\tau^\theta \equiv \theta e^{-\lambda\tau} + \bar{a}\left(1 - e^{-\lambda\tau}\right)$ be the principal's beliefs given that last inspection quality was θ and the time elapsed since that inspection is τ. For any $\tau < \bar{\tau}$, we can write the survival function of the time of an inspection as $1 - F(\tau) = e^{-M_\tau}$, where

$$M_\tau = \int_0^\tau m_s \mathrm{d}s - \sum_{s \le \tau} \log(1 - p_s).$$

In addition, let's denote the expected payoff to the principal from inspecting quality at time τ by $\mathcal{M}_\theta(U, x_\tau) \equiv x_\tau^\theta U_H + (1 - x_\tau^\theta) U_L - c$, where $U = (U_L, U_H)$ is the continuation payoff under the 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 formulation for the principal's problem in (13)

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

subject to

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

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

$$q_\tau \in [\underline{q}, 1], \ \forall \tau \in [0, \bar{\tau}]$$

$$0 \le m_\tau$$

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

For a fixed continuation payoff U, this is an optimal control problem with state constraints where the controls are the monitoring intensity m_τ, and the probability of monitoring p_τ. The latter variable captures the possibility of atoms in the monitoring CDF (Given this, the review date $\bar{\tau}$ is useful in terms of the exposition but somewhat redundant, as we are allowing for jumps in the CDF). The magnitude of these jumps is restricted by the constraint that $q_\tau \geq \underline{q}$ 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* $(\mathscr{G}^L(U), \mathscr{G}^H(U)) = U$.

To solve this problem using the theory of optimal control with state constraints requires that we attach a Lagrange multiplier ψ_τ to the incentive compatibility constraint $q_\tau \geq \underline{q}$.[12] It is also convenient to reformulate the problem using the principal's continuation value, U_τ, as a state variable (where with some abuse of notation we omit the dependence on θ_0). At any continuity point, the principal's continuation value satisfies the differential equation

$$\dot{U}_\tau = (r + m_\tau)U_\tau - u(x^\theta_\tau) - m_\tau \mathcal{M}(U, x^\theta_\tau). \tag{15}$$

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

$$U_{\tau^-} = p_\tau \mathcal{M}(U, x^\theta_\tau) + (1 - p_\tau)U_{\tau^+}. \tag{16}$$

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

$$\lim_{\tau \to \infty} e^{-rt - M_\tau} U_\tau = 0. \tag{17}$$

A control $\tilde{m}_\tau$ is admissible if and only if condition (17) is satisfied. Using the continuation value of the principal as a state variables, we can now formulate the problem in Mayer form (Cesari, 2012).[13] The Hamiltonian of this problem is

$$\mathcal{H}(q_\tau, \zeta_\tau, \nu_\tau, \psi_\tau, m_\tau, \tau) = \zeta_\tau\big((r + m_\tau)U_\tau - u(x^\theta_\tau) - m_\tau \mathcal{M}(U, x^\theta_\tau)\big) + \nu_\tau\left((r + \lambda + m_\tau)q_\tau - m_\tau\right) + \psi_\tau\left(q_\tau - \underline{q}\right).$$

ζ_τ is the (current value) co-state variable associated to U_τ, and ν_τ is the (current value) co-state variable associated to q_τ.[14] According to Pontryagin's maximum principle, the evolution of the co-state variables is given by

$$\dot{\zeta}_\tau = (r + m_\tau)\zeta_\tau - (r + m_\tau)\zeta_\tau = 0, \ \zeta_0 = -1. \tag{18a}$$

$$\dot{\nu}_\tau = -\lambda\nu_\tau - \psi_\tau. \tag{18b}$$

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

$$\nu_0 \leq 0 \text{ with equality if } q_0 > \underline{q} \tag{19}$$

Equation (19) 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

[12]This method is usually referred as the direct adjoining approach (Hartl et al., 1995).
[13]In Mayer form we maximize U_0 subject to (10), (15), (16), and the state constraint on q_τ
[14]If $\tilde{p}_\tau$ is the co-state for a particular state variable, the current value co-state is defined as $p_\tau \equiv e^{rt+M_\tau}\tilde{p}_\tau$.

must be zero while if the IC constraint is binding the marginal effect must be negative.

The Hamiltonian is linear in m_τ with a coefficient given by

$$S(\tau) = x^\theta_\tau U_H + (1 - x^\theta_\tau) U_L - c - U_\tau - (1 - q_\tau)\nu_\tau. \tag{20}$$

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

$$m_\tau = \begin{cases} 0 & \text{if } S(\tau) < 0 \\ [0, \infty] & \text{if } S(\tau) = 0. \end{cases} \tag{21}$$

If the previous condition is not satisfied at some τ then the optimal solution requires a jump in the monitoring distribution. Equation (20) 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 $\mathcal{M}_\theta(U, x^\theta_\tau) \geq U_\tau$. The extra term $(1 - q_\tau)\nu_\tau$ reflects the impact of moral hazard on the optimal policy. The co-state variable ν_τ is negative, which implies that because of incentive reasons it could be optimal to monitor even when information has negative net value, namely $\mathcal{M}_\theta(U, x^\theta_\tau) < U_\tau$. This is intuitive: the need to provide incentives forces the principal to increase the level of monitoring above and beyond his informational needs.

The optimality conditions in (21) require that $S(\tau)$ is zero whenever $m_\tau > 0$, and this means the switching function must be constant over any time interval where the monitoring intensity is positive. In technical terms, the switching function $S(\tau)$ must be constant across a singular arc, which is a standard condition in the theory of singular optimal control. Accordingly, whenever $m_\tau > 0$ it must be the case that $\dot{S}(t) = 0$. If we differentiate $S(\tau)$ and use the optimality condition $m_\tau S(\tau) = 0$ then we get

$$\begin{aligned} \dot{S}(\tau) &= \dot{x}^\theta_\tau (U_H - U_L) - \dot{U}_\tau + \dot{q}_\tau \nu_\tau - (1 - q_\tau)\dot{\nu}_\tau \\ &= u(x^\theta_\tau) + (r + \lambda) q_\tau \nu_\tau + (1 - q_\tau)(\lambda \nu_\tau + \psi_\tau) + \dot{x}^\theta_\tau (U_H - U_L) - r U_\tau \\ &= 0. \end{aligned}$$

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

$$\psi_\tau = \begin{cases} 0 & \text{if } q_\tau > \underline{q} \\ \geq 0 & \text{if } q_\tau = \underline{q}. \end{cases} \tag{22}$$

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

$$\psi_\tau = \frac{r U_\tau - (r\underline{q} + \lambda)\nu_\tau - u(x^\theta_\tau) - \dot{x}^\theta_\tau (U_H - U_L)}{1 - \underline{q}}. \tag{23}$$

The next step is to characterize the necessary condition that an optimal review date $\bar{\tau}$ must satisfy; in particular, whenever the optimal monitoring policy entails a deterministic review date, the following terminal condition must be satisfied at time $\bar{\tau}$:

$$r\mathcal{M}_\theta(U, x^\theta_{\bar{\tau}}) = u(x^\theta_{\bar{\tau}}) + (r + \lambda)\nu_{\bar{\tau}} + \dot{x}^\theta_{\bar{\tau}} (U_H - U_L). \tag{24}$$

Finally we must consider the possibility of atoms in the CDF. If there is an atom at time s, then the optimal monitoring probability is given by

$$p_s \in \arg \max_{p \in [0,(q_{s^-}-\underline{q})/(1-\underline{q})]} \left\{ \frac{p}{1-p} \left[\mathcal{M}_\theta(U, x_s^\theta) - U_{s^-} - \nu_{s^+}(1-q_{s^-}) \right] \right\}, \tag{25}$$

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

$$\mathcal{M}_\theta(U, x_\tau^\theta) - \nu_\tau(1-q_\tau) \leq U_\tau. \tag{26}$$

Equation (26) implies that if $S(\tau) \leq 0$ for all τ, then q_τ is continuous and there is at most one atom in the monitoring distribution taking place at time $\bar{\tau}$. Lemma 3 summarizes the necessary conditions of optimality.[15]

Lemma 3. *Given continuation payoffs (U_L, U_H), if the monitoring policy $(m_\tau^*, p_\tau^*, \bar{\tau}^*)$ is a solution of the optimal control problem $\mathscr{G}^\theta U$ then there is a trajectory of the co-state variables (ζ_τ, ν_τ) satisfying* (18a), (18b) *and* (19)*, and a non-negative Lagrange multiplier ψ_τ such that:*

1. *The Lagrange multiplier ψ_τ is given by* (23) *if $q_\tau = \underline{q}$ and zero otherwise.*
2. *The optimal monitoring rate m_τ^* is given by* (21)*.*
3. *There is an atom p_τ^* at time τ only if $S(\tau) \geq 0$, and $S(\tau) \leq 0$ holds at any point of continuity.*
4. *If $\bar{\tau}^* < \infty$ then condition* (24) *is satisfied.*

4.5 Optimal Policy

Armed with the necessary conditions we can characterize the optimal policy. The general characterization of the optimal policy follows from two results: First, we show that the monitoring intensity is interior only if the incentive compatibility constraint is binding. Second, we show that the optimal policy has at most one atom which corresponds to the final review date $\bar{\tau}$.

We first prove that the monitoring rate m_τ is positive only if the incentive compatibility constraint binds. Formally, we have the following result:

Lemma 4. *Let $(m_\tau^*, p_\tau^*, \bar{\tau}^*)$ be an optimal policy, then $q_\tau > \underline{q} \Rightarrow m_\tau^* = 0$.*

It is important to note that the previous Lemma does not predict zero monitoring at times when the incentive constraint is slack. Instead, it establishes that at any point when the incentive compatibility constraint is slack, there is either no monitoring or an atom in the distribution of monitoring, but both cases require m_τ^*=0.

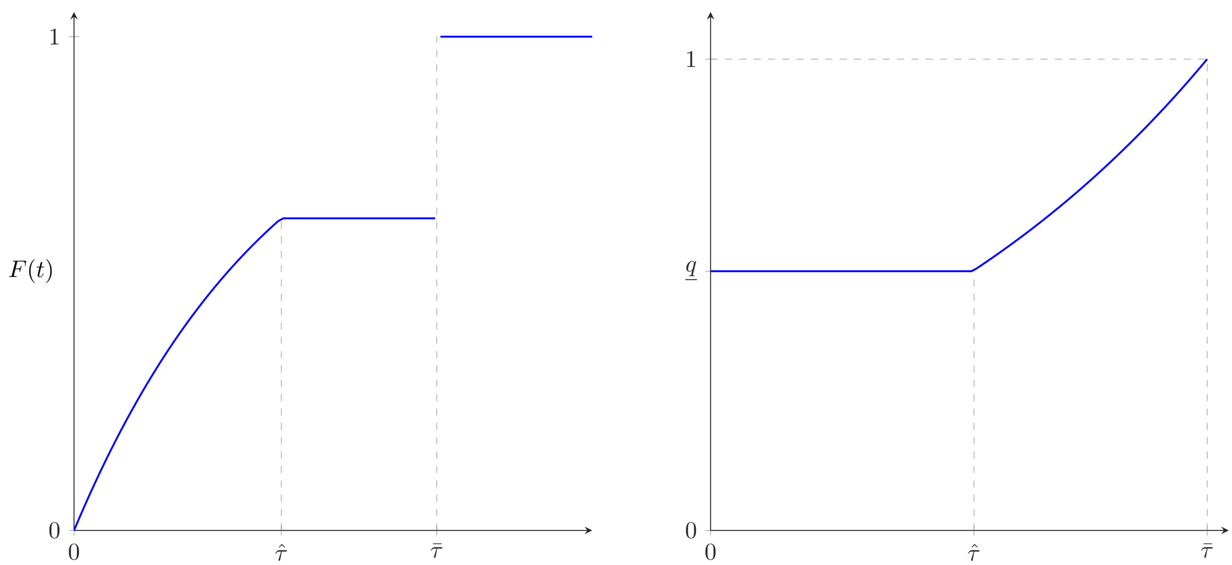
The next lemma proves that there cannot be atoms in the distribution of monitoring prior to the final review date $\bar{\tau}$.

Lemma 5. *Let $(m_\tau^*, p_\tau^*, \bar{\tau}^*)$ be an optimal policy, then $p_s^* = 0$ for all $s \in [0, \bar{\tau}^*)$.*

[15] 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.

Some inspections are driven by incentive provision –early ones– and others by learning. This lemma confirms the bang-bang nature of learning-driven inspections: as uncertainty accumulates, learning may become important to the principal, which is behind the jump in the CDF: however the jump must be such that the CDF attains 1 at that point, explaining why there can only be a single jump in a given cycle.

Because the optimal policy has no jumps before time τ^*, Lemma 4 immediately implies that q_τ is non-decreasing in time. Moreover, this also implies that if $m_\tau^* = 0$ at time τ, then it must be the case that $m_s^* = 0$ for all $s \in [\hat{\tau}, \bar{\tau}^*]$. Figure 2 illustrates the form of the optimal policy implied of Lemmas 4 and 5. Figure 2b shows the trajectory of q_τ: either the incentive compatibility is binding and q_τ is constant or q_τ increases until it reaches one. Figure 2a shows the monitoring policy associated to the trajectory of q_τ. Before time $\hat{\tau}$, the incentive compatibility constraint is binding, and this requires a monitoring rate equal to m^* (where m^* is the same as in Proposition 2). After time $\hat{\tau}$, the incentive compatibility constraint is slack so there is no monitoring up to time $\bar{\tau}$, at which point $q_\tau = 1$, and so the agent is monitored with probability one.



(a) Example CDF

(b) Path q_τ

Figure 2: Cumulative density function and path of q_τ implied by Lemmas 4 and 5.

Moreover, whenever $0 < \hat{\tau} < \bar{\tau} < \infty$ we have that $\hat{\tau} = \max\{\tau < \bar{\tau} : \text{IC constraint is binding}\}$, and so $\hat{\tau}$ and $\bar{\tau}$ are connected by the equation $q_{\hat{\tau}} = \underline{q} = e^{-(r+\lambda)(\bar{\tau}-\hat{\tau})}$: Hence, $\hat{\tau}$ is pinned down by $\bar{\tau}$ and is given by

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

This means that, given $\theta_{T_{n-1}}$, the optimal policy can be characterized by the single number $\bar{\tau}$ and takes the following form: 1) $m_\tau = m^*$ for some (possibly empty) time interval $[0, \hat{\tau}]$, and 2) $m_\tau = 0$ on the interval $(\hat{\tau}, \bar{\tau}]$. If $\hat{\tau} = 0$ then the optimal policy is purely deterministic, while if $\bar{\tau} = \infty$, then the optimal policy is purely random.

It turns out that the optimal policy always takes one of these two extremes. The following proposition provides a general characterization of the optimal monitoring policy.

Proposition 3 (Optimal Monitoring Policy)**.** *Suppose that $u(x)$ is strictly convex, then the optimal policy* *given $\theta_{T_{n-1}} = \theta$ is either:*

1. *deterministic with an inspection date*

$$\bar{\tau}^*_\theta \leq \tau^{bind} \equiv \frac{1}{r+\lambda} \log \frac{1}{\underline{q}},$$

where τ^{bind} is the review time that makes the incentive constraint binding at time zero.

2. *or fully random, with inspection time following an exponential distribution with mean arrival rate*

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

The optimal policy is remarkably simple. It's either deterministic, when the the principal's learning incentives are relatively strong or fully random otherwise. The constant monitoring rate arising in the random case may seem surprising because it shows that beliefs about quality don't affect monitoring rates.[16] The reason is that the monitoring intensity is chosen such that the incentive constraints bind at all times but, given our technology assumptions and the absence of news, the incentive constraints are the same for all beliefs and qualities.

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

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

and let, $\mathcal{G}^\theta_{\text{rand}}$ be the payoff associated with the fully random policy, as given by:

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

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 3 studies the effect of monitoring cost c and effort cost k on the optimal policy. The left panel shows that, as the cost of monitoring c increases, the policy shifts from relying on deterministic monitoring to random monitoring. This shift reflects the fact that, as c goes up, the policy becomes more concerned about incentive provision. The right panel shows that as the cost of effort k goes up, exacerbating moral hazard concerns, the policy relies more on random reviews.

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; there is inter-temporal smoothing. 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

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

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. The next proposition highlight some comparative statics.

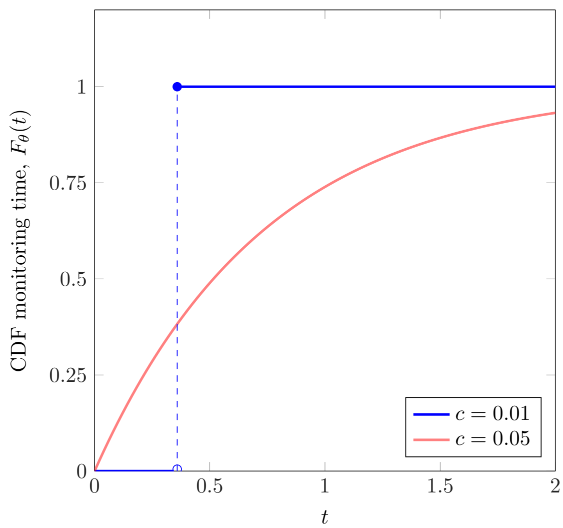
Proposition 4 (Comparative Statics)**.**

1. *There is* $c^\dagger > 0$ *such that, if* $c < c^\dagger$ *then the optimal policy is deterministic monitoring, and if* $c > c^\dagger$ *then the optimal policy is random monitoring.*
2. *There is* $k^\dagger < \lambda/(r+\lambda)$ *such that for any* $k > k^\dagger$ *the optimal policy is random monitoring.*
3. *There is* $\bar{a}^\dagger < 1$ *such that, for any* $\bar{a} \in (\bar{a}^\dagger, 1)$*, the optimal policy given* $\theta_{T_{n-1}} = H$ *is random monitoring. Similarly, there is* $\bar{a}_\dagger > 0$ *such that, for any* $\bar{a} \in (0, \bar{a}_\dagger)$*, the optimal policy given* $\theta_{T_{n-1}} = L$ *is random monitoring.*

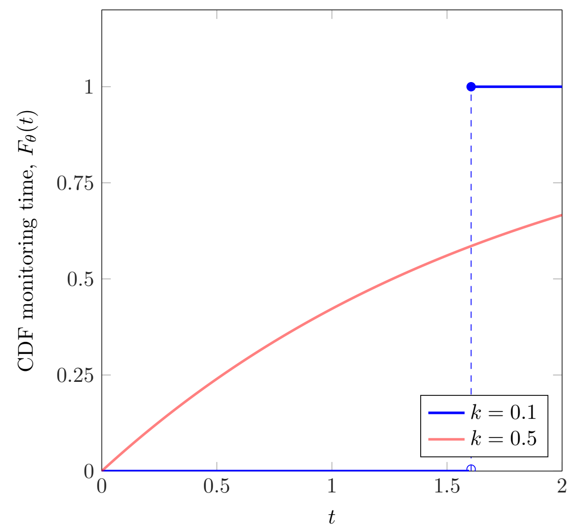
Figure 3a shows the monitoring distribution for low and high monitoring cost: When the cost of monitoring is low, the policy implements deterministic monitoring; in fact, if the cost of monitoring is sufficiently low then monitoring in the benchmark policy (the relaxed problem without incentive constraint) is sufficiently frequent, and so the incentive compatibility constraint is slack. Similarly, Figures 3b and 4b show the comparative statics for the cost of effort, k. The monitoring policy is random if k is high enough. On the other hand, the optimal policy is deterministic if k is low, and the cost of monitoring is sufficiently low (so there is some monitoring in the benchmark policy). This is intuitive, when the cost of effort is low there is little need for monitoring to incentivize effort because the incentive compatibility constraint is slack: so, the benchmark policy (deterministic monitoring) is optimal in the problem with incentives.

Finally, Figure 4a shows the comparative statics for $\bar{a}$. When $\bar{a}$ is low, the monitoring policy calls for random monitoring after negative outcomes and deterministic monitoring after positive outcomes. On the contrary, if $\bar{a}$ is high, then the optimal policy calls for deterministic monitoring after negative outcomes and random monitoring after positive outcomes. The intuition for these results is as follows: If $\bar{a}$ is high, then high quality is highly persistent, and so an agent who is exerting effort is very likely to sustain a high quality. This means that there is no much that we can learn from monitoring, and monitoring is driven by the provision of incentives, which means that random monitoring is optimal. On the other hand, low quality is transitory: beliefs quickly revert to the steady state $\bar{a}$ and learning becomes important, which means that deterministic monitoring is likely to be optimal.

The general lesson is that persistent states should be followed by random monitoring while transitory states should be followed by deterministic monitoring.In the case of schools, it is reasonable to assume high quality is persistent: Schools that have built the organization capital needed to provide a high quality education are likely to sustain it over time. If this is the case, then the previous results suggest that we should monitor schools with a high performance and schools with a low performance differently: Schools with a high performance less frequently (on average) and randomly, while low performing schools are reviewed periodically at pre-specified review dates.

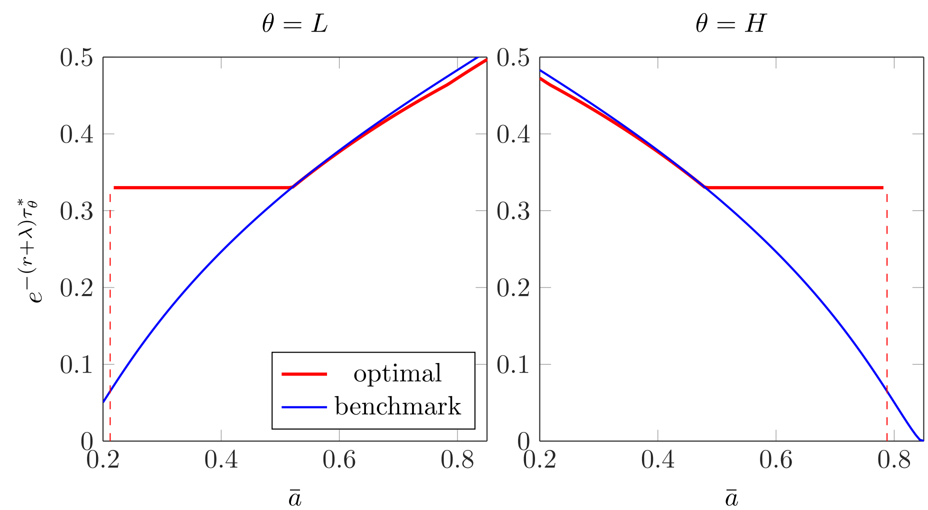


(a) Comparative statics for c. The optimal monitoring policy is given by $\hat{\tau}_\theta^* = \{0, \infty\}$ and $\tau_\theta^* = \{0.35, \infty\}$. The first best policy is $\tau_\theta^{FB} = \{0.35, 2.41\}$.

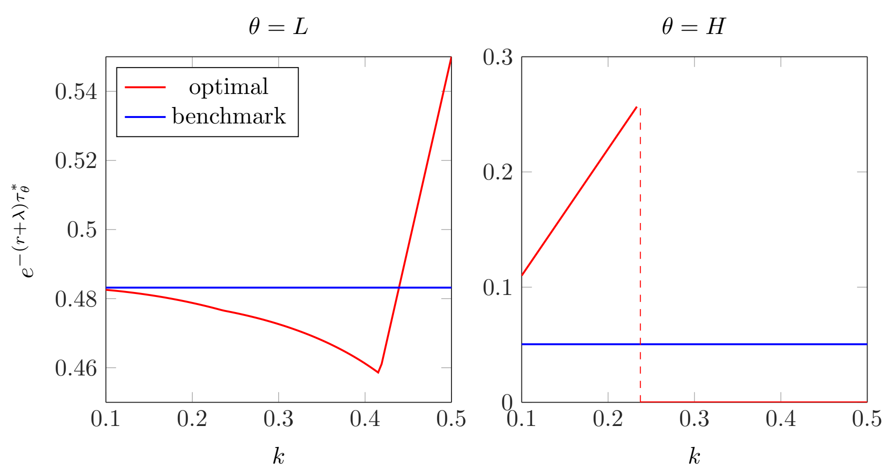


(b) Comparative statics for k. The optimal monitoring policy is given by $\hat{\tau}_\theta^* = \{0, \infty\}$ and $\tau_\theta^* = \{1.60, \infty\}$. The first best policy is $\tau_\theta^{FB} = 1.60$.

Figure 3: Comparative statics for optimal monitoring distribution. The figure shows the CDF of the monitoring time T_n when $u(x_\tau) = x_\tau - 0.5 \times x_\tau(1 - x_\tau)$ 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.



(a) Comparative statics with respect to $\bar{a}$.



(b) Comparative statics with respect to k.

Figure 4: Comparative statics for $\tau *_\theta$. Baseline parameters: $u(x_t) = x_t^2$ and $r = 0.05$, $\lambda = 0.5$, $k = 0.4$, $\bar{a} = 0.8$, $c = 0.15$. The left panel indicates the policy that follows $\theta_{T_{n-1}} = L$ while the right panel indicates the policy that follows $\theta_{T_{n-1}} = H$. The benchmark policy corresponds to the case without hidden effort while the optimal policy corresponds to the case with hidden effort. The dotted line in the optimal policy refer to the discontinuity when we change from a policy with deterministic inspections to a policy with random inspections.

4.6 Brownian Quality Shocks

One might think previous results are driven by the binary nature of quality. This binary specification makes the analysis more tractable (especially in the presence of exogenous news that we consider in the next section) but we think the structure of the optimal policy is driven by the nature of incentives and not the details of the quality process.

To support this intuition we show that the qualitative nature of the results holds when we look at a more general process of quality. As mentioned in Remark 2, the incentive compatibility characterization in Proposition 1 holds if quality follows any process such as

$$d\theta_t = \lambda(a_t - \theta_t)dt + \sigma dZ_t, \tag{28}$$

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, the principal's expected payoff given a monitoring policy M implementing $a_t = \bar{a}$ is just

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

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

Proposition 5. *Suppose the principal's preferences are linear in θ_τ. Then, the optimal monitoring policy is given by a Poisson process with mean arrival rate*

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

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 the analysis and computations 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.

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

$$x_t = \theta_0 e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right) \tag{29}$$

$$\Sigma_t = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right). \tag{30}$$

Moreover, using the law of iterated expectations, we can show that the principal's continuation payoff is linear in quality and given by

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

where $\mathcal{C}$ solves

$$\begin{aligned}
\mathcal{C} = & \min_{\bar{\tau}\geq 0, m_\tau \geq 0, q_0 \geq \underline{q}} \int_0^{\bar{\tau}} e^{-r\tau - M_{\tau^-}} \left(\gamma\Sigma_\tau + m_\tau\left(c + \mathcal{C}\right)\right)\mathrm{d}\tau + e^{-r\bar{\tau} - M_{\bar{\tau}}}\left(c + \mathcal{C}\right) \\
& \text{subject to} \\
& \dot{q}_\tau = (r + \lambda + m_\tau)q_\tau - m_\tau, \; q_\tau = 1 \\
& q_\tau \in [\underline{q}, 1], \; \forall \tau \in [0, \bar{\tau}].
\end{aligned}$$

The optimal policy is now formulated recursively as a cost minimization problem, where the cost borne by the principal has two sources, monitoring and uncertainty, as captured by the residual variance of quality Σ_τ. As before, the principal chooses the monitoring intensity m_τ , the probability of a review p_τ, and the date of the scheduled review $\bar{\tau}$. The main state variable is again the expected time until the next review, q_τ. 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 $(\hat{\tau}, \bar{\tau})$ solving

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

Given the symmetry in the linear-quadratic case, the optimal policy is independent of the outcome in the previous inspection. As in the binary case, we can reduce the problem to a simple one dimensional optimization problem, and we can show that the optimal policy consist on either periodic reviews or pure random monitoring. This means that the cost of monitoring is

$$\mathcal{C} = \min\left\{\min_{\bar{\tau}\in[0,\tau^{\text{bind}}]} \frac{\int_0^{\bar{\tau}} e^{-r\tau}\gamma\Sigma_\tau\mathrm{d}\tau + e^{-r\bar{\tau}}c}{1 - e^{-r\bar{\tau}}}, \frac{(r+m^*)\int_0^\infty e^{-(r+m^*)\tau}\gamma\Sigma_\tau\mathrm{d}\tau + m^* c}{r}\right\}, \tag{31}$$

and the optimal monitoring policy is given by:

Proposition 6. *Suppose that θ_t follows the Ornstein-Uhlenbeck process in equation* (28)*, and that the principal's expected payoff flow is $u(x_t, \Sigma_t) = x_t - \gamma\Sigma_t$, then the optimal monitoring policy is either deterministic monitoring at intervals of fixed length $\bar{\tau}^*$ or random with a constant Poisson arrival rate m^*.*

The comparative statics in the case of Brownian shocks are similar to those in Proposition 4. There are two new parameters in the model, γ and σ: However, after inspecting equations (30) and (31) we notice that the monitoring policy only depends on the therm $c/\gamma\sigma^2$, and thus any increment in γ or σ is equivalent to a reduction in the cost of monitoring: It follows immediately that the frequency of monitoring is increasing in γ and σ, and that random monitoring is optimal if either γ or σ are sufficiently low. Interestingly, the optimal policy is independent of $\bar{a}$ when the principal preferences are linear in the mean and the variance. This is the case because when quality follows because the variance of quality is independent of $\bar{a}$ if quality follows an Ornstein-Uhlenbeck process: in this case, $\bar{a}$ only affects the mean. However, this is a knife edge case as the monitoring policy would depend on $\bar{a}$ if the principal preferences are non-linear in reputation.

5 Exogenous News

Exogenous news such as media articles, customer reviews, and academic research provide information to the market that may complement or substitute the principal's own monitoring and disclosures. To provide some insights about the interaction between monitoring policy and exogenous news, we consider the presence of 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. We find that when the exogenous news convey negative states faster than positive states, then monitoring tends to intensify when the firm's reputation is low because the moral hazard is also more intense. By contrast, when exogenous news convey positive states faster than positive states, then monitoring is more intense in good times since effort incentives are weaker when the firm's reputation is high.

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_τ^θ be the firm's continuation value conditional on being type θ_τ and define $D_\tau \equiv \Pi_\tau^H - \Pi_\tau^L$. The continuation value must satisfy the Bellman equations

$$r\Pi_\tau^H = \max_{a\in[0,\bar{a}]} \left\{ x_\tau - ka_\tau - \lambda(1 - a_\tau)D_{\tau^-} + (\mu_H + m_\tau)(\Pi(H) - \Pi_\tau^H) + \dot{\Pi}_\tau^H \right\}$$
$$r\Pi_\tau^L = \max_{a\in[0,\bar{a}]} \left\{ x_\tau - ka_\tau + \lambda a_\tau D_{\tau^-} + (\mu_L + m_\tau)(\Pi(L) - \Pi_\tau^L) + \dot{\Pi}_\tau^L \right\},$$

where we use the fact that if $a_t = \bar{a}$ for any $t \geq T_n$ then – whenever $\theta_{T_n} \in \mathcal{F}^P_{T_n}$ – the continuation payoff is $\Pi^\theta_0 = \Pi(\theta)$, where $\Pi(\theta)$ was defined in (5). From here it follows that full effort $a_\tau = \bar{a}$ is incentive compatible if and only if:[18]

$$D_\tau \geq \frac{k}{\lambda}.$$

Accordingly, given any incentive-compatible policy, the firm's continuation value satisfies the following differential equation (in absence of news and monitoring):

$$\begin{aligned}
\dot{\Pi}^H_\tau &= (r + \mu_H + m_\tau)\Pi^H_\tau - x_\tau + k\bar{a} + D_\tau\lambda(1-\bar{a}) - (\mu_H + m_\tau)\Pi(H) \\
\dot{\Pi}^L_\tau &= (r + \mu_L + m_\tau)\Pi^L_\tau - x_\tau + k\bar{a} - D_\tau\lambda\bar{a} - (\mu_L + m_\tau)\Pi(L).
\end{aligned}$$

Taking the difference between $\dot{\Pi}^H_\tau$ and $\dot{\Pi}^L_\tau$ (and keeping $a_\tau = \bar{a}$) we get that D_τ satisfies the following differential equation

$$\dot{D}_\tau = (r + \lambda + m_\tau)D_\tau - \mu_H(\Pi(H) - \Pi^H_\tau) + \mu_L(\Pi(L) - \Pi^L_\tau) - m_\tau\Delta.$$

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

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 monitoring or news), and the type observed at that time. The optimal control problem (ignoring jumps in the monitoring distribution) becomes

$$\mathscr{G}^\theta(U_L, U_H) = \sup_{\bar{\tau}, m_\tau, \Pi^{\theta^c}_0} \int_0^{\bar{\tau}} e^{-r\tau - M_{\tau-}} \left(u(x^\theta_\tau) + \mu_H x^\theta_\tau U_H + \mu_L(1 - x^\theta_\tau)U_L + m_\tau \mathcal{M}_\theta(U, x_\tau)\right) \mathrm{d}\tau + e^{-r\bar{\tau} - M_{\bar{\tau}}}\mathcal{M}_\theta(U, x_{\bar{\tau}})$$

subject to

$$\begin{aligned}
&\dot{\Pi}^H_\tau = (r + \mu_H + m_\tau)\Pi^H_\tau - x_\tau + k\bar{a} + \lambda(1-\bar{a})(\Pi^H_\tau - \Pi^L_\tau) - (\mu_H + m_\tau)\Pi(H),\ \Pi^H_{\bar{\tau}} = \Pi(H) \\
&\dot{\Pi}^L_\tau = (r + \mu_L + m_\tau)\Pi^L_\tau - x_\tau + k\bar{a} - \lambda\bar{a}(\Pi^H_\tau - \Pi^L_\tau) - (\mu_L + m_\tau)\Pi(L),\ \Pi^L_{\bar{\tau}} = \Pi(L) \\
&\Pi^\theta_0 = \Pi(\theta) \\
&\frac{k}{\lambda} \leq \Pi^H_\tau - \Pi^L_\tau,\ \forall \tau \in [0, \bar{\tau}] \\
&0 \leq m_\tau.
\end{aligned}$$

Note that even though $\Pi^{\theta_0}_0 = \Pi(\theta_0)$ while if $x_0 = 0$, $\Pi^L_0 = \Pi(L)$ it is not true that $\Pi^{\neg\theta_0}_0 = \Pi(\neg\theta_0)$ if $x^{\theta_0}_0 \neq \theta_0$ because there is a divergence between the market and the firm's beliefs.

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

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_τ as

$$\dot{D}_\tau = (r + \mu + \lambda + m_\tau)D_\tau - (\mu + m_\tau)\Delta.$$

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

$$\dot{q}_\tau = (r + \mu + \lambda + m_\tau)q_\tau - \mu - m_\tau, \tag{32}$$

The only difference between equation (32) and the one in the case without news is that we have added μ to the monitoring rate m_τ. This means we can write the principal's problem as

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

subject to

$$\dot{q}_\tau = (r + \mu + \lambda + m_\tau)q_\tau - \mu - m_\tau$$
$$q_\tau \geq \underline{q}, \ \forall \tau \in [0, \bar{\tau}]$$
$$0 \leq m_\tau.$$

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 $[\tau_0, \tau_1]$, then it must be that $\dot{q}_\tau = 0$ 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 suffices 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 7. *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 Intensity

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 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, $\bar{\tau} = \infty$) 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) $\bar{\tau} = \infty$, and 2) $m_\tau > 0$ only if the incentive compatibility constraint is binding, that is if $\Pi^H_\tau - \Pi^L_\tau = k/\lambda$. Given the previous results, it is natural to conjecture that for linear $u(\cdot)$ this is indeed an optimal monitoring policy, at least for a wide range of parameters. Given this conjecture, we can follow the same steps as before, and derive the monitoring rate using the conditions $(\dot{\Pi}^H_\tau - \dot{\Pi}^L_\tau) = 0$ and $\Pi^H_\tau - \Pi^L_\tau = k/\lambda$. This conditions are necessary for the incentive compatibility constraint to be binding all the time, and using this necessary conditions, we can derive the rate boils which is given by

$$m_\tau = \alpha + \beta \Pi^L_\tau, \tag{33}$$

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 so in the bad news case the monitoring rate is positive only if $\Pi^L_\tau \leq -\alpha/\beta$, and in the good news case, the monitoring rate is positive only if $\Pi^L_\tau \geq -\alpha/\beta$. That is, with bad news, monitoring is needed only if the firm's continuation value is low, and with good news, monitoring is needed only if the firm's continuation value is high. The logic for these conditions follows the results in Board and Meyer-ter-Vehn (2013): With bad news, the incentives for effort increase in reputation, while with good news the incentives for effort decrease in reputation.

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. The details of the necessary conditions for optimality are relegated to the appendix. In this section we focus on the simplest case in which $m_\tau > 0$ for all $\tau \geq 0$; this simpler case 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 $\Pi^H_\tau = \Pi^L_\tau + D_\tau = \Pi^L_\tau + k/\lambda$ and the monitoring rate (33) we write the evolution of

[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). That is, our analysis focuses on the cases where news are not informative enough, and so some amount of monitoring is needed at all times to solve the agency problem.

the low quality firm continuation value as

$$\dot{\Pi}^L_\tau = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi^L_\tau + \beta(\Pi^L_\tau)^2 - x_\tau. \tag{34}$$

If $\theta_0 = L$ then the initial condition is $\Pi^L_0 = \Pi(L)$. If $\theta_0 = H$ (and the incentive compatibility is binding) the initial condition is $\Pi^L_0 = \Pi(H) - k/\lambda$.[20] We can analyze the evolution of monitoring by studying the phase diagram in the space (x_τ, Π^L_τ) in Figure 5.

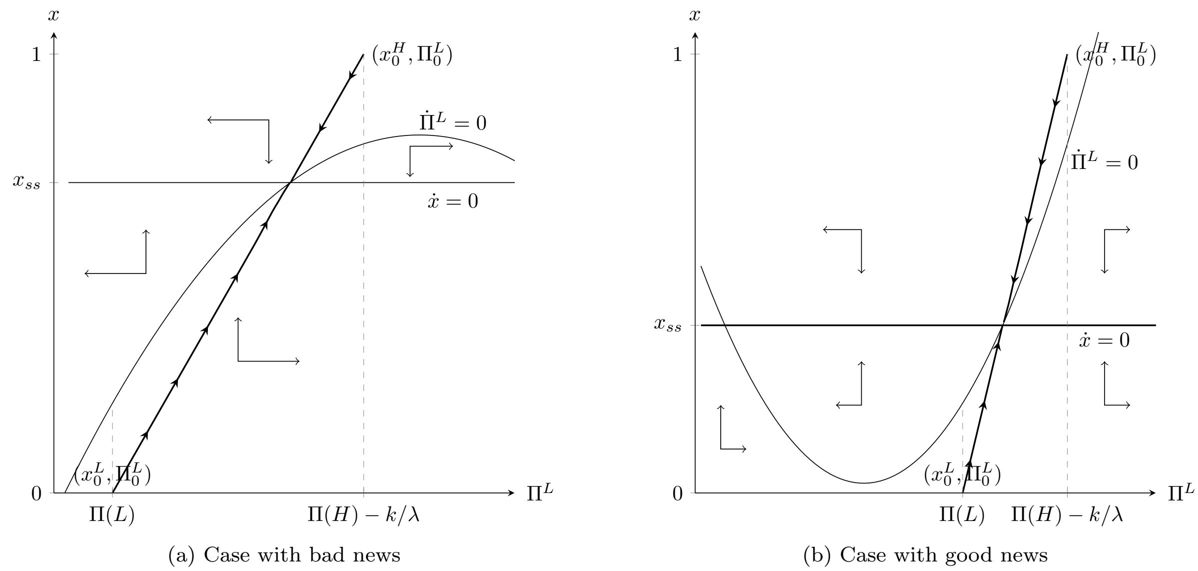


Figure 5: Phase diagram. The (x_τ, Π^L_τ) system has two steady states. In each case, one of the steady states is a saddle point. If the optimal solution is such $m_\tau > 0$ all $\tau \geq 0$, then the optimal solution corresponds to the trajectory converging to the saddle point. In this case, the analysis of the phase diagram reveals that the trajectory of Π^L_τ must be monotone between news arrivals. This immediately implies that the evolution of monitoring between news is monotone as well.

Using the ODE for Π^L_τ in equation (34) we get a quadratic equation for the steady state.

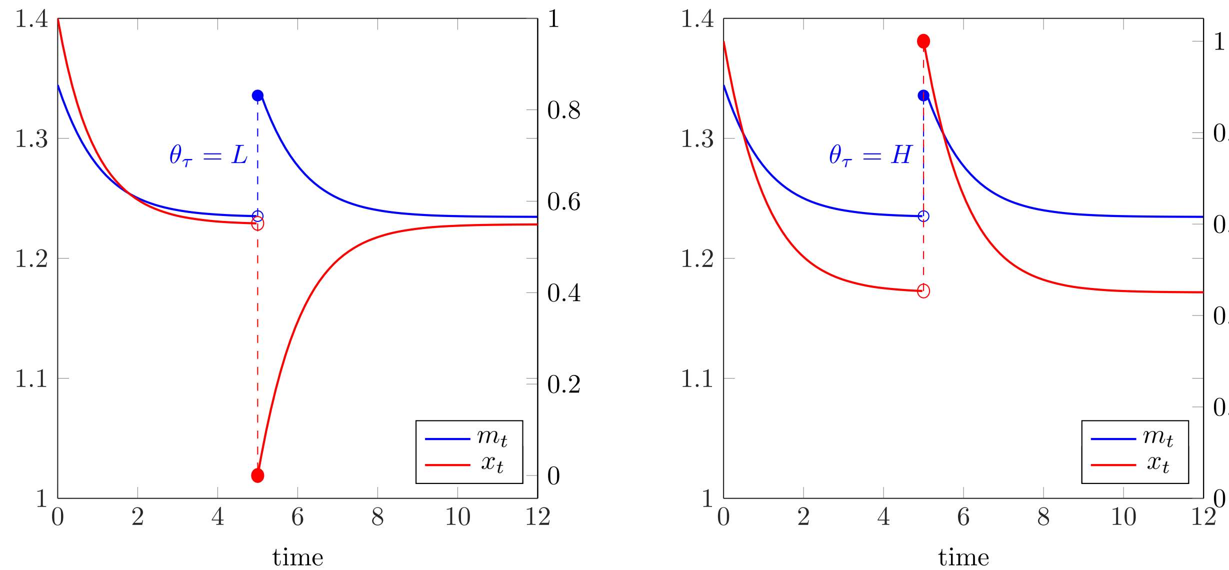
$$0 = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi^L + \beta(\Pi^L)^2 - x. \tag{35}$$

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 5.

From inspection of the phase diagram it is clear that Π^L_τ is monotone: it is decreasing after good news and increasing after bad news. This implies that in the bad news case, jumps up and after bad news and then decreases; while it jumps down after good and then increases. Figure 6 shows the dynamics of monitoring: In the bad news case, monitoring increases after news. The opposite trend is optimal in the good news

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

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.



(a) Bad news

(b) Good news

Figure 6: 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 blue curves represent optimal monitoring intensity, m_τ and the red curves the evolution of reputation, x_τ. 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. Parameters: $r = 0.1$, $k = 0.5$, $c = 0.1$, $\bar{a} = 0.5$, $\lambda = 1$. In the bad news case we take $\mu_H = 0$, and $\mu_L = 0.2$, and in the good news case we take $\mu_H = 0.2$, and $\mu_L = 0$

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 use (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 a dual role: 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.

Some caveats are in order. As mentioned above, we study policies that implement full effort. Full effort is efficient in the first best, but it is possible that an optimal policy would prescribe no inspections and implement no effort after some particularly bad histories, if there is little noise. The intuition for this comes from the case $\bar{a} = 1$, which is the case with no noise. Under those circumstances, if the firm achieves high quality it can prevent any future drops by putting full effort. 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 strong 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 the analysis to characterize optimal monitoring policy and effort choice is a good 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 contracting settings, 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

$$(r+\lambda)U'(x) = u'(x) + \lambda(\bar{a} - x)U''(x) \tag{36a}$$

$$(r+2\lambda)U''(x) = u''(x) + \lambda(\bar{a} - x)U'''(x) \tag{36b}$$

Using (36b) 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 (36a) 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

$$\frac{u'(\underline{x})}{r+\lambda} \leq U(1) - U(0) \leq \frac{u'(\overline{x})}{r+\lambda} \tag{37}$$

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 (37) is satisfied. The fact that U is increasing follows directly from the convexity of U and equation (36a).

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

$$xU(1) + (1-x)U(x) - U(x) \leq c. \tag{38}$$

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,

$$0 \geq u(x) + \lambda(\bar{a} - x)U'(x) - rU(x), \; x \in [0,1]. \tag{39}$$

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

$$
\begin{aligned}
e^{-rt}E[U(\tilde{x}_t)] &= U(x_0) + E\left[\int_0^t e^{-rs}(\lambda(\bar{a}-\tilde{x}_t)U'(\tilde{x}_t) - rU(\tilde{x}_t))\mathrm{d}s + \sum_{s\leq t} e^{-rs}\big(\tilde{x}_s U(1) + (1-\tilde{x}_s)U(0) - U(\tilde{x}_s)\big)\right] \\
&\leq U(x_0) - E\left[\int_0^t e^{-rs}u(\tilde{x}_t)\mathrm{d}s - \sum_{s\leq t} e^{-rs}c\right], \qquad (40)
\end{aligned}
$$

where we have used inequalities (38) and (39). Taking the limit when $t\to\infty$ we conclude that

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

The proof concludes noting that (40) holds with equality for the optimal policy. □

B Case without News

B.1 Incentive Compatibility: Proof of Lemma 1

Proof. Let $N_t^{LH} = \sum_{s\leq t}\mathbf{1}_{\{\theta_{s-}=L,\theta_s=H\}}$ and $N_t^{HL} = \sum_{s\leq t}\mathbf{1}_{\{\theta_{s-}=H,\theta_s=L\}}$ 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

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

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

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

For any deviation we have that

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

So, we can write the incentive compatibility condition as

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

The result in the lemma then 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 $\bar{\tau}$ solving $e^{-(r+\lambda)\bar{\tau}}=\underline{q}$ 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 $E(e^{-rT})=\int_0^\infty e^{-rt}dF(t)\leq e^{-r\epsilon}$. For any bounded functions f,g we have that $|\sup f-\sup g|\leq\sup|f-g|$. Clearly, the function $\mathscr{G}_\epsilon=(\mathscr{G}_\epsilon^L,\mathscr{G}_\epsilon^H)$ is bounded in R. Accordingly, we have that

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

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

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

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

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

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 $\bar{\tau}$ 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 in Section 4.2 is equivalent to the infinite horizon problem

$$\sup_{m_\tau,p_\tau,q_0} \int_0^\infty e^{-rt-M_{\tau-}} \Big(u(x_\tau^\theta) + m_\tau \mathcal{M}_\theta(U, x_\tau)\Big) \mathrm{d}\tau + \sum e^{-rs-M_{s-}} p_s \mathcal{M}_\theta(U, x_s^\theta)$$

subject to

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

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

$$q_\tau \in [\underline{q}, 1],\ \forall \tau \in [0, \bar\tau]$$

$$0 \le m_\tau$$

$$p_s \le q_{\bar\tau-}.$$

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

B.3.1 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 $q_\tau > \underline{q}$ then the Lagrange multiplier is $\psi_\tau = 0$. We show that if $\psi_\tau = 0$ then it can not be the case that $S(\tau) = 0$ along a singular arc, which is a necessary condition for $0 < m_\tau < \infty$.

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

$$-rU_\tau + u(x_\tau^\theta) + (rq_\tau + \lambda)\nu_\tau + \dot x_\tau^\theta (U_H - U_L) = 0 \tag{41}$$

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

$$-r\dot U_\tau + u'(x_\tau^\theta)\dot x_\tau^\theta + r\dot q_\tau \nu_\tau + (rq_\tau + \lambda)\dot\nu_\tau + \ddot x_\tau^\theta (U_H - U_L) = 0$$

Using the first order condition $m_\tau S(\tau) = 0$ we can write the evolution of U_τ

$$\dot U_\tau = rU_\tau - u(x_\tau^\theta) - m_\tau \nu_\tau (1 - q_\tau).$$

Replacing $\dot U_\tau$ and $\dot\nu_\tau$ we get

$$r(-rU_\tau + u(x_\tau^\theta) + m_\tau\nu_\tau(1-q_\tau)) + u'(x_\tau^\theta)\dot x_\tau^\theta + r\nu_\tau((r+\lambda+m_\tau)q_\tau - m_\tau) - (rq_\tau+\lambda)\lambda\nu_\tau + \ddot x_\tau^\theta(U_H - U_L) = 0,$$

which after some simplification yield

$$-rU_\tau + u(x_\tau^\theta) + r^{-1}u'(x_\tau^\theta)\dot x_\tau^\theta + r^{-1}(r^2 q_\tau - \lambda^2)\nu_\tau + r^{-1}\ddot x_\tau^\theta (U_H - U_L) = 0 \tag{42}$$

Combining equations (41) and (42) we get that along the circular arc we have

$$\nu_\tau = \frac{\dot x_\tau^\theta}{\lambda^2 + r\lambda}\Big(u'(x_\tau^\theta) - (r+\lambda)(U_H - U_L)\Big) \tag{43}$$

Differentiating (51) we get

$$\begin{aligned}
\dot{\nu}_\tau &= \frac{\ddot{x}^\theta_\tau}{\lambda^2 + r\lambda}\Big(u'(x^\theta_\tau) - (r+\lambda)(U_H - U_L)\Big) + \frac{(\dot{x}^\theta_\tau)^2}{\lambda^2 + r\lambda}u''(x^\theta_\tau) \\
&= \frac{\ddot{x}^\theta_\tau}{\dot{x}^\theta_\tau}\nu_\tau + \frac{(\dot{x}^\theta_\tau)^2}{\lambda^2 + r\lambda}u''(x^\theta_\tau) \\
&= -\lambda\nu_\tau + \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau)
\end{aligned}$$

Equation (18b) implies that

$$\dot{\nu}_\tau = -\lambda\nu_\tau - \psi_\tau = -\lambda\nu_\tau < -\lambda\nu_\tau + \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau),$$

which yields the contradiction as $u''(x^\theta_\tau) > 0$. This shows that it cannot be the case that $q_\tau > \underline{q}$ and $S(\tau) = 0$ 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 τ it must be the case that $p_\tau = 1$. First, we specify the necessary conditions for a jump to be optimal using the maximum principle for optimal control problems with jumps in the state variables in Seierstad and Sydsaeter (1986). Because the optimality conditions are stated in terms of the non-current-value co-state variables, and ν_τ and ζ_τ are current value co-states, we need to be careful of using the multipliers $\tilde{\nu}_\tau = e^{-r\tau - M_\tau}\nu_t$ and $\tilde{\zeta}_\tau = e^{-r\tau - M_\tau}\zeta_\tau$ when we apply the results in Seierstad and Sydsaeter (1986) to account for the effect of jumps in co-state variables.

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

$$e^{-rt - M_{\tau^+}}\nu_{\tau^+} - e^{-rt - M_{\tau^-}}\nu_{\tau^-} = -e^{-rt - M_{\tau^+}}\nu_{\tau^+}\frac{p_\tau}{1-p_\tau}.$$

Hence, we have that

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

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

$$\begin{aligned}
S(\tau^+) &= x^\theta_\tau U_H + (1 - x^\theta_\tau)U_L - c - U_{\tau^+} - (1-q_\tau)\nu_{\tau^+} \\
&= \mathcal{M}(U, x^\theta_\tau) - \frac{U_{\tau^-}}{1-p_\tau} + \frac{p_\tau}{1-p_\tau}\mathcal{M}(U, x^\theta_\tau) - (1-q_\tau)\nu_{\tau^-} \\
&= S(\tau^-) + \frac{p_\tau}{1-p_\tau}\big(\mathcal{M}(U, x^\theta_\tau) - U_{\tau^-}\big) \qquad (44)
\end{aligned}$$

Using the continuity of ν_τ, the jump size p_τ solves

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

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

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

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

$$\begin{aligned} \mathcal{H}(q_{\tau+},\zeta_{\tau+},\nu_{\tau+},\psi_{\tau+},m_{\tau+},\tau) - \frac{\mathcal{H}(q_{\tau-},\zeta_{\tau-},\nu_{\tau-},\psi_{\tau-},m_{\tau-},\tau)}{1-p_\tau} - \dot{x}^\theta_\tau(U_H - U_L)\frac{p_\tau}{1-p_\tau} &= \\ \frac{p_\tau}{1-p_\tau}\Big[u(x^\theta_\tau) - r\mathcal{M}(U,x^\theta_\tau) + (r+\lambda)\nu_{\tau-} - \dot{x}^\theta_\tau(U_H-U_L) - r(1-q_\tau)\nu_\tau\Big] &= \\ \frac{p_\tau}{1-p_\tau}\Big[u(x^\theta_\tau) + r\mathcal{M}(U,x^\theta_\tau) + (r+\lambda)\nu_{\tau-} - \dot{x}^\theta_\tau(U_H-U_L) - r(1-q_\tau)\nu_\tau\Big] & \end{aligned}$$

This means that it must be the case that

$$r\mathcal{M}(U,x^\theta_\tau) = u(x^\theta_\tau) + (r+\lambda)\nu_{\tau-} + \dot{x}^\theta_\tau(U_H-U_L) - r(1-q_\tau)\nu_\tau$$

Note that this is the same as condition (24) to determine $\bar{\tau}$ (in which case $q_{\bar{\tau}} = 1$). Hence, it is enough to show that there is at most one point satisfying this condition.

Uniqueness of atom at $\bar{\tau}$: Differentiating the switching function $S(\tau)$ we get that

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

where

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

Suppose that there is an atom at $\tilde{\tau} < \bar{\tau}$, then it must be the case that $S(\tilde{\tau}^-) = 0$ and $G(\tilde{\tau}^-) = r(1-q_{\tilde{\tau}-})\nu_{\tilde{\tau}-} \leq 0$ ($\psi_\tau \geq 0$ and $\nu_0 \leq 0$ implies that ν_τ 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 $G(\tau)$ we have that $G(\bar{\tau}) = 0$ 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 $G(\tilde{\tau}) < 0$: $G(\bar{\tau}) = 0$ implies that $\dot{S}(\bar{\tau}) = 0$. At time $\bar{\tau}$, it must be the case that there is some $\epsilon > 0$ such that for all $\tau \in (\bar{\tau}-\epsilon,\bar{\tau})$ we have: i) $q_\tau > \underline{q}$, and i) $S(\tau) < 0$. Point i) follows from the constraint $p_\tau \leq (q_{\tau-} - \underline{q})/(1-\underline{q})$ while ii) follows from i) together with Lemma 4. Because $S(\tau)$ is negative for $\tau \in (\bar{\tau}-\epsilon,\bar{\tau})$, $S(\tau^-) = 0$ and $\dot{S}(\bar{\tau}) = G(\bar{\tau}) = 0$ there must be some interval $(\bar{\tau}-\epsilon',\bar{\tau})$ such that $\dot{S}(\tau) > 0$ and $\ddot{S}(\tau) < 0$ for all $\tau \in (\bar{\tau}-\epsilon',\bar{\tau})$. This means that $\ddot{S}(\tau^-) = \dot{S}(\tau^-) + \dot{G}(\tau^-) = \dot{G}(\bar{\tau}) \leq 0$.

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

$$\ddot{G}(\tau) = -\lambda\dot{G}(\tau) + (\dot{x}^\theta_\tau)^2u''(x^\theta_\tau).$$

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

$$\dot{G}(\tau) = -\int_\tau^\tau e^{\lambda(s-\tau)}(\dot{x}^\theta_s)^2u''(x^\theta_s)\mathrm{d}s + e^{\lambda(\tau-\tau)}\dot{G}(\tau).$$

Because $\dot{G}(\tau) \leq 0$ and $(\dot{x}^\theta_s)^2u''(x^\theta_s) \geq 0$ we get that $\dot{G}(\tau) \leq 0$ for all $\tau \in (\tilde{\tau},\tau)$ which yields a contradiction to the fact that $\dot{G}(\tau) > 0$ for all $\tau \in (\tau-\epsilon,\tau)$.

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 $\tilde{\tau}$, $\ddot{S}(\tau)\leq 0$ 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 $\dot{G}(\tau)>0$ 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 $\tilde{\tau}>\tilde{\tau}$ such that $G(\tilde{\tau})<0$. We can then replicate the same argument as in the case with $G(\tilde{\tau})<0$ starting from time $\tilde{\tau}$.

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 (44)). Accordingly, it is the case that $m_\tau=0$ and so $q_{\tau+\epsilon}>\underline{q}$ for small $\epsilon>0$. Lemma 4 implies then that $m_{\tau+\epsilon}=0$ and so $m_\tau=0$ and $\dot{q}_\tau>0$ for all $\tau>\bar{\tau}$. 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 $F(\tau)$ has at most one atom.

□

Proof of Proposition 3

Proof. Lemmas 4 and 5 imply that the optimal policy must take the following form:

$$F(\tau)=\begin{cases}1-e^{-\min\{\tau,\hat{\tau}\}m^*} & \text{if } \tau<\bar{\tau}\\ 1 & \text{if } \tau\geq\bar{\tau}\end{cases}, \tag{45}$$

where

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

The only step left is to show that $\hat{\tau}\in\{0,\infty\}$. We consider the following optimization problem:

$$\begin{aligned}U_\theta&=\max_{\bar{\tau}\geq 0}\int_0^\tau e^{-r\tau\ -m^*\tau\wedge\hat{\tau}}\left(u(x_\tau^\theta)+m^*\mathcal{M}(U,x_\tau^\theta)\mathbf{1}_{\tau<\hat{\tau}}\right)\mathrm{d}\tau+e^{-r\bar{\tau}-m^*\bar{\tau}\wedge\hat{\tau}}\mathcal{M}(U,x_{\bar{\tau}}^\theta)\\ \hat{\tau}&=\max\left(\tau+\frac{\log\underline{q}}{r+\lambda},0\right).\end{aligned}$$

The objective function is

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

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

$$\begin{aligned}\mathcal{U}'(\bar{\tau})&=e^{-(r+m^*)\hat{\tau}}\left(u(x_{\hat{\tau}}^\theta)+m^*\mathcal{M}(U,x_{\hat{\tau}}^\theta)\right)-m^*e^{-m^*\hat{\tau}}\int_{\hat{\tau}}^{\bar{\tau}}e^{-r\tau}u(x_\tau^\theta)\mathrm{d}\tau\\ &+e^{-r\bar{\tau}-m^*\hat{\tau}}u(x_{\bar{\tau}}^\theta)-e^{-(r+m^*)\hat{\tau}}u(x_{\hat{\tau}}^\theta)-(r+m^*)e^{-r\bar{\tau}-m^*\hat{\tau}}\mathcal{M}(U,x_{\bar{\tau}}^\theta)\\ &+e^{-r\bar{\tau}-m^*\hat{\tau}}\dot{x}_{\bar{\tau}}^\theta(U_H-U_L)\\ &=e^{-m^*\hat{\tau}}\Bigg[m^*e^{-r\hat{\tau}}\mathcal{M}(U,x_{\hat{\tau}}^\theta)-m^*\int_{\hat{\tau}}^{\bar{\tau}}e^{-r\tau}u(x_\tau^\theta)\mathrm{d}\tau\\ &+e^{-r\bar{\tau}}u(x_{\bar{\tau}}^\theta)-(r+m^*)e^{-r\bar{\tau}}\mathcal{M}(U,x_{\bar{\tau}}^\theta)+e^{-r\tau}\dot{x}_{\bar{\tau}}^\theta(U_H-U_L)\Bigg]\end{aligned}$$

We define the function

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

so we can write the first derivative as

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

and the second derivative as

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

where

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

We show that any candidate interior solution $\bar{\tau}$ 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 $\mathcal{U}'(\bar{\tau}) = 0 \Rightarrow \mathcal{U}''(\bar{\tau}) \geq 0$, which is equivalent to show that $H(\bar{\tau}) = 0 \Rightarrow H'(\bar{\tau}) \geq 0$. If $H(\bar{\tau}) = 0$, then we can write

$$
\begin{aligned}
H'(\bar{\tau}) = e^{-r\bar{\tau}} \Big[& \Big(m^* \underline{q}^{-\frac{r}{r+\lambda}} \dot{x}^\theta_{\hat{\tau}} - (r+m^*)\dot{x}^\theta_{\bar{\tau}} - \lambda \dot{x}^\theta_{\bar{\tau}} \Big)(U_H - U_L) \\
& + u'(x^\theta_{\bar{\tau}})\dot{x}^\theta_{\bar{\tau}} + m^* \left(q^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - u(x^\theta_{\bar{\tau}}) - e^{r\bar{\tau}} \int_{\hat{\tau}}^{\bar{\tau}} r e^{-r\tau} u(x^\theta_\tau)\mathrm{d}\tau \right) \Big]
\end{aligned}
$$

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

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

Hence, we get that

$$
H'(\bar{\tau}) = e^{-r\bar{\tau}} \left[u'(x^\theta_{\bar{\tau}})\dot{x}^\theta_{\bar{\tau}} + m^* \left(\underline{q}^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - u(x^\theta_{\bar{\tau}}) + e^{r\tau} \int_{\hat{\tau}}^{\bar{\tau}} (-re^{-r\tau}) u(x^\theta_\tau)\mathrm{d}\tau \right) \right]
$$

Using integration by parts

$$
\begin{aligned}
e^{r\tau} \int_{\hat{\tau}}^{\tau} (-re^{-rt}) u(x^\theta_\tau)\mathrm{d}\tau &= u(x^\theta_\tau) - e^{r(\tau - \hat{\tau})} u(x^\theta_{\hat{\tau}}) - e^{r\tau} \int_{\hat{\tau}}^{\tau} e^{-rt} u'(x^\theta_\tau)\dot{x}^\theta_\tau \mathrm{d}\tau \\
&= u(x^\theta_\tau) - \underline{q}^{-\frac{r}{r+\lambda}} u(x^\theta_{\hat{\tau}}) - e^{r\tau} \int_{\hat{\tau}}^{\tau} e^{-rt} u'(x^\theta_\tau)\dot{x}^\theta_\tau \mathrm{d}\tau,
\end{aligned}
$$

and so we get that

$$\begin{aligned} H'(\tau) &= e^{-r\tau}u'(x^\theta_{\bar\tau})\dot x^\theta_\tau - (r+\lambda)\frac{\underline{q}}{1-\underline{q}}\int_{\hat\tau}^{\tau} e^{-rt}u'(x^\theta_\tau)\dot x^\theta_\tau \mathrm{d}\tau \\ &= \lambda(\bar a-\theta)\left[e^{-(r+\lambda)\tau}u'(x^\theta_{\bar\tau}) - \frac{\underline{q}}{1-\underline{q}}\int_{\hat\tau}^{\tau}(r+\lambda)e^{-(r+\lambda)\tau}u'(x^\theta_\tau)\mathrm{d}\tau\right], \end{aligned}$$

where the second line comes from replacing $\dot x^\theta_\tau = \lambda(\bar a-\theta)e^{-\lambda\tau}$. If $\theta=0$, then we have that $u'(x^\theta_\tau) > u'(x^\theta_\tau)$ for all $\tau \le \bar\tau$ so

$$\begin{aligned} H'(\tau) &> \lambda\bar a u'(x^\theta_{\bar\tau})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}(r+\lambda)\int_{\hat\tau}^{\tau} e^{-(r+\lambda)\tau}\mathrm{d}\tau\right] \\ &= \lambda\bar a u'(x^\theta_{\bar\tau})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}\left(e^{-(r+\lambda)\hat\tau} - e^{-(r+\lambda)\tau}\right)\right] \\ &= \lambda\bar a u'(x^\theta_{\bar\tau})e^{-(r+\lambda)\tau}\left[1 - \frac{1}{1-\underline{q}}\left(1-\underline{q}\right)\right] \\ &= 0. \end{aligned}$$

On the other hand, if $\theta=1$, then we have that $u'(x^\theta_{\bar\tau}) < u'(x^\theta_\tau)$ for all $\tau < \bar\tau$ so

$$\begin{aligned} H'(\tau) &> \lambda(1-\bar a)u'(x^\theta_{\bar\tau})\left[-e^{-(r+\lambda)\tau} + \frac{\underline{q}}{1-\underline{q}}(r+\lambda)\int_{\hat\tau}^{\tau} e^{-(r+\lambda)\tau}\mathrm{d}\tau\right] \\ &= \lambda(1-\bar a)u'(x^\theta_{\bar\tau})\left[e^{-(r+\lambda)\tau} - \frac{\underline{q}}{1-\underline{q}}\left(e^{-(r+\lambda)\hat\tau} - e^{-(r+\lambda)\tau}\right)\right] \\ &= \lambda(1-\bar a)u'(x^\theta_{\bar\tau})e^{-(r+\lambda)\tau}\left[1 - \frac{1}{1-\underline{q}}\left(1-\underline{q}\right)\right] \\ &= 0. \end{aligned}$$

Hence, $H'(\tau) \ge 0$ so $\mathcal{U}''(\tau) \ge 0$ for any τ such that $\mathcal{U}'(\tau)=0$, and this means that the maximum cannot be an interior point.

□

Proof of Proposition 4

Before proving the statements in the proposition, we start proving the following Lemma

Lemma 6. *For any $\underline{q} \in (0,1)$,*

$$e^{-r\tau^{bind}} > \frac{m^*}{r+m^*}$$

Proof. If we let $\beta \equiv r/(r+\lambda)$, then by replacing τ^{bind} and m^* we can verify that it is enough to show that

$$\underline{q}^\beta - \frac{\underline{q}}{\beta(1-\underline{q})+\underline{q}} > 0.$$

Consider the function

$$H(q) \equiv \beta q^{\beta-1} + (1-\beta)q^\beta - 1,$$

so we need to show that $H(q) > 0$ for all $q \in (0,1)$. The function H is such $H(0) > 0$ and $H(1) = 0$. Moreover, the derivative of H is given by

$$H'(q) = \beta(\beta-1)q^{\beta-2} + (1-\beta)\beta q^{\beta-1} = -\beta(1-\beta)q^{\beta-2}(1-q) < 0,$$

and so it follows that $H(q) > 0$ for all $q \in (0,1)$. $\square$

Comparative statics c

Fix an arbitrary continuation value U. Define

$$\Delta(c,U) \equiv \mathcal{G}^{\theta}_{\text{det}}(U) - \mathcal{G}^{\theta}_{\text{rand}}(U).$$

We want to show that $\Delta(c,U) > 0$ if and only if $c < c^{+}$, where $\Delta\left(c^{+},U\right) = 0$. For any $\tau \in \left[0, \tau^{bind}\right]$ we have

$$e^{-r\tau} - \frac{m^*}{r+m^*} \geq e^{-r\tau^{\text{bind}}} - \frac{m^*}{r+m^*} > 0,$$

where the second inequality follows from Lemma 6. Therefore,

$$\left.\frac{\partial\Delta(c,U)}{\partial c}\right|_{c^{+}} = -\left(e^{-r\tau} - \frac{m^*}{r+m^*}\right) < 0,$$

and so for any U, we have

$$\Delta(c,U) > 0 \leftrightarrow c < c^{+}.$$

Optimality of random monitoring for large k

We show that when k converges to its upper bound, $\lambda/(r+\lambda)$, the difference between the benefit of using random and deterministic monitoring converge to zero while the difference in their cost remains bounded away of zero. For large k, we can restrict attention to monitoring policies in which the IC constraint is binding, and it is enough to compare policies that rely exclusively on deterministic or random monitoring (the argument to rule out policies that alternate between random and deterministic depending on $\theta_{T_{n-1}}$ is analogous).

First, we look at the difference in the cost. The cost of deterministic policy is

$$C^{\text{det}} = \frac{e^{-r\bar{\tau}}}{1-e^{-r\bar{\tau}}} = \frac{\underline{q}^{\beta}}{1-\underline{q}^{\beta}}$$

while the cost of the random policy is

$$C^{\text{rand}} = \frac{m^*}{r} = \frac{1}{\beta}\frac{\underline{q}}{1-\underline{q}}.$$

The difference in the cost is

$$C^{\text{det}} - C^{\text{rand}} = \frac{\underline{q}^{\beta}}{1-\underline{q}^{\beta}} - \frac{1}{\beta}\frac{\underline{q}}{1-\underline{q}} = \frac{1}{\beta}\frac{\beta\underline{q}^{\beta} - \underline{q} + (1-\beta)\underline{q}^{\beta+1}}{1-\underline{q}-\underline{q}^{\beta}+\underline{q}^{\beta+1}},$$

and applying L'Hopital's rule twice we find that

$$\begin{aligned}
\lim_{\underline{q}\to 1}\frac{\beta\underline{q}^{\beta} - \underline{q} + (1-\beta)\underline{q}^{\beta+1}}{1-\underline{q}-\underline{q}^{\beta}+\underline{q}^{\beta+1}} &= \lim_{\underline{q}\to 1}\frac{\beta^2\underline{q}^{\beta-1} - 1 + (1-\beta)(1+\beta)\underline{q}^{\beta}}{-1-\beta\underline{q}^{\beta-1}+(\beta+1)\underline{q}^{\beta}} \\
&= \lim_{\underline{q}\to 1}\frac{\beta(\beta-1)+(1-\beta^2)\underline{q}}{(1-\beta)+(\beta+1)\underline{q}} \\
&= \frac{1-\beta}{2} > 0
\end{aligned}$$

Next, we look at the benefit of monitoring (excluding its cost). First, we compute the benefit of a deterministic policy. The benefit of the deterministic policy, B_θ^{det}, solves the system of equations

$$\begin{aligned} B_L^{\text{det}} &= \int_0^{\bar\tau} e^{-r\tau} u(x_\tau^L)\mathrm{d}\tau + e^{-r\bar\tau}(x_{\bar\tau}^L B_H^{\text{det}} + (1-x_{\bar\tau}^L)B_L^{\text{det}}) \\ B_H^{\text{det}} &= \int_0^{\bar\tau} e^{-r\tau} u(x_\tau^H)\mathrm{d}\tau + e^{-r\bar\tau}(x_{\bar\tau}^H B_H^{\text{det}} + (1-x_{\bar\tau}^H)B_L^{\text{det}}). \end{aligned}$$

Solving this system we get that the payoff is given by

$$\begin{aligned} B_L^{\text{det}} &= \frac{\int_0^{\bar\tau} e^{-r\tau} u(x_\tau^L)\mathrm{d}\tau}{1-e^{-r\bar\tau}} + \frac{e^{-r\bar\tau}x_{\bar\tau}^L}{1-e^{-r\bar\tau}(x_{\bar\tau}^H - x_{\bar\tau}^L)} \frac{\int_0^{\bar\tau} e^{-rt}(u(x_{\bar\tau}^H)-u(x_{\bar\tau}^L))\mathrm{d}\tau}{1-e^{-r\bar\tau}} \\ B_H^{\text{det}} &= \frac{\int_0^{\bar\tau} e^{-rt} u(x_{\bar\tau}^H)\mathrm{d}\tau}{1-e^{-r\bar\tau}} - \frac{e^{-r\bar\tau}(1-x_{\bar\tau}^H)}{1-e^{-r\bar\tau}(x_{\bar\tau}^H - x_{\bar\tau}^L)} \frac{\int_0^{\bar\tau} e^{-rt}(u(x_{\bar\tau}^H)-u(x_{\bar\tau}^L))\mathrm{d}\tau}{1-e^{-r\bar\tau}}, \end{aligned}$$

and taking the limit when $\bar\tau \to 0$ (which is equivalent to taking the limit when $k \to \lambda/(r+\lambda)$) we get that

$$\begin{aligned} B_L^{\text{det}} &\to \frac{1}{r}\left(\frac{r+\lambda(1-\bar a)}{r+\lambda}u(0) + \frac{\lambda\bar a}{r+\lambda}u(1)\right) \\ B_H^{\text{det}} &\to \frac{1}{r}\left(\frac{\lambda(1-\bar a)}{r+\lambda}u(0) + \frac{r+\lambda\bar a}{r+\lambda}u(1)\right) \end{aligned}$$

On the other hand, the benefit of the random policy is

$$\begin{aligned} B_L^{\text{rand}} &= \int_0^\infty e^{-(r+m^*)\tau}(u(x_\tau^L) + m^*(x_\tau^L B_H^{\text{rand}} + (1-x_\tau^L)B_L^{\text{rand}}))\mathrm{d}\tau \\ B_H^{\text{rand}} &= \int_0^\infty e^{-(r+m^*)\tau}(u(x_\tau^H) + m^*(x_\tau^H B_H^{\text{rand}} + (1-x_\tau^H)B_L^{\text{rand}}))\mathrm{d}\tau, \end{aligned}$$

where

$$\begin{aligned} B_L^{\text{rand}} &= \int_0^\infty e^{-(r+m^*)\tau}u(x_\tau^L)\mathrm{d}\tau + \frac{m^*}{r+m^*}B_L^{\text{rand}} + \frac{m^*\lambda\bar a}{(r+m^*)(r+\lambda+m^*)}(B_H^{\text{rand}} - B_L^{\text{rand}}) \\ B_H^{\text{rand}} &= \int_0^\infty e^{-(r+m^*)\tau}u(x_\tau^H)\mathrm{d}\tau + \frac{m^*}{r+m^*}B_L^{\text{rand}} + \left[\frac{m^*}{r+\lambda+m^*} + \frac{m^*\lambda\bar a}{(r+m^*)(r+\lambda+m^*)}\right](B_H^{\text{rand}} - B_L^{\text{rand}}) \end{aligned}$$

From here we get

$$B_H^{\text{rand}} - B_L^{\text{rand}} = \frac{r+\lambda+m^*}{r+\lambda}\int_0^\infty e^{-(r+m^*)\tau}(u(x_\tau^H) - u(x_\tau^L))\mathrm{d}\tau$$

So, replacing in the previous equations

$$B_L^{\text{rand}} = \frac{r+m^*}{r}\int_0^\infty e^{-(r+m^*)\tau}u(x_\tau^L)\mathrm{d}\tau + \frac{m^*\lambda\bar a}{r(r+\lambda)(r+m^*)}\int_0^\infty (r+m^*)e^{-(r+m^*)\tau}(u(x_\tau^H) - u(x_\tau^L))\mathrm{d}\tau.$$

We can also write

$$B_H^{\text{rand}} - B_L^{\text{rand}} = \frac{r+\lambda+m^*}{(r+\lambda)(r+m^*)}\int_0^\infty (r+m^*)e^{-(r+m^*)\tau}(u(x_\tau^H) - u(x_\tau^L))\mathrm{d}\tau$$

From here we get that when $m^* \to \infty$ the benefit converges to

$$B_L^{\text{rand}} \to \frac{1}{r}\left(\frac{r+\lambda(1-\bar a)}{r+\lambda}u(0) + \frac{\lambda\bar a}{r+\lambda}u(1)\right),$$

and

$$B_H^{\text{rand}} - B_L^{\text{rand}} \to \frac{1}{r+\lambda}(u(1) - u(0))$$

so

$$B_H^{\text{rand}} \to \frac{1}{r}\left(\frac{\lambda(1-\bar{a})}{r+\lambda}u(0) + \frac{r+\lambda\bar{a}}{r+\lambda}u(1)\right)$$

Comparing the limit of the deterministic and random policy we verify that both yield the same benefit in the limit of $C^{\text{det}} - C^{\text{rand}}$ is strictly positive, which means that the random policy dominates.

So, the difference in the benefit converges to zero while the limit of

Optimality of random monitoring following $\theta_{T_{n-1}} = H$ for large $\bar{a}$.

First, we find an upper bound for payoff of following a deterministic policy

$$\begin{aligned}
\mathcal{G}_{\text{det}}^{\theta}(U) &= \int_0^{\bar{\tau}} e^{-rt} u\left(x_{\tau}^{\theta}\right) \mathrm{d}\tau + e^{-r\bar{\tau}}\left[U_L - c + \bar{a}\Delta U + (\theta - \bar{a}) e^{-\lambda\bar{\tau}} \Delta U\right] \\
&< \frac{u(1)}{r}(1 - e^{-r\bar{\tau}}) + e^{-r\bar{\tau}}(U_H - c) \\
&\leq \frac{u(1)}{r}(1 - e^{-r\tau^{\text{bin}}}) + e^{-r\tau^{\text{bin}}}(U_H - c) \\
&= \frac{u(1)}{r}(1 - \underline{q}^{\frac{r}{r+\lambda}}) + \underline{q}^{\frac{r}{r+\lambda}}(U_H - c)
\end{aligned}$$

Next, we find a lower bound for the payoff of following a random policy

$$\begin{aligned}
\mathcal{G}_{\text{rand}}^{\theta}(U) &= \int_0^{\infty} e^{-(r+m^*)\tau}\left[u\left(x_{\tau}^{\theta}\right) + m^* \mathcal{M}\left(U, x_{\tau}^{\theta}\right)\right] \mathrm{d}\tau \\
&> \int_0^{\infty} e^{-(r+m^*)\tau} \mathrm{d}\tau \left[u(\bar{a}) + m^*(\bar{a}U_H + (1-\bar{a}) U_L - c)\right] \\
&= \frac{u(\bar{a})}{r+m^*} + \frac{m^*(\bar{a}U_H + (1-\bar{a}) U_L - c)}{r+m^*}
\end{aligned}$$

Finally, we show that if $\bar{a}$ is large enough, then the upper bound for $\mathcal{G}_{\text{det}}^{\theta}(U)$ is below the lower bound for $\mathcal{G}_{\text{rand}}^{\theta}$. This requires that for any U we have

$$\frac{u(1)}{r}(1 - \underline{q}^{\frac{r}{r+\lambda}}) + \underline{q}^{\frac{r}{r+\lambda}}(U_H - c) \leq \frac{u(\bar{a})}{r+m^*} + \frac{m^*(\bar{a}U_H + (1-\bar{a}) U_L)}{r+m^*}$$

Following the proof in Lemma 6, we let $\beta \equiv \frac{r}{r+\lambda}$ so we can write

$$\begin{aligned}
\frac{u(\bar{a})}{r+m^*} + \frac{m^*(\bar{a}U_H + (1-\bar{a}) U_L)}{r+m^*} &= \frac{u(\bar{a})}{r}(1 - \underline{q}^{\beta}) + u(\bar{a})\left(\frac{\underline{q}^{\beta} - 1}{r} + \frac{1}{r+m^*}\right) \\
&\quad + q^{\beta}(\bar{a}U_H + (1-\bar{a}) U_L - c) \\
&\quad + \left(\frac{m^*}{r+m^*} - q^{\beta}\right)(\bar{a}U_H + (1-\bar{a}) U_L - c)
\end{aligned}$$

Letting $\Delta U \equiv U_H - U_L$, we write our required inequality as

$$\left(\frac{u(1)}{r} - \frac{u(\bar{a})}{r}\right)(1 - \underline{q}^{\beta}) \leq \frac{u(\bar{a})}{r}\left(\underline{q}^{\beta} - \frac{m^*}{r+m^*}\right) + \left(\frac{m^*}{r+m^*} - q^{\beta}\right)(U_H - c) - \frac{m^*}{r+m^*}(1-\bar{a})\Delta U,$$

and after replacing m^* we reduce it to

$$\left(\frac{u(1)}{r}-\frac{u(\bar{a})}{r}\right)(1-\underline{q}^{\beta})\leq\left(\frac{u(\bar{a})}{r}+c-U_H\right)\left(\underline{q}^{\beta}-\frac{\underline{q}}{\beta(1-\underline{q})+\underline{q}}\right)-\frac{\underline{q}(1-\bar{a})\Delta U}{\beta(1-\underline{q})+\underline{q}}$$

Clearly, it must be the case that $\frac{u(1)}{r}>U_H$, which means that

$$\begin{aligned}\lim_{\bar{a}\to 1}\left(\frac{u(1)}{r}-\frac{u(\bar{a})}{r}\right)(1-\underline{q}^{\beta})=0&\\ &<\left(\frac{u(1)}{r}+c-U_H\right)\left(\underline{q}^{\beta}-\frac{\underline{q}}{\beta(1-\underline{q})+\underline{q}}\right)\\ &=\lim_{\bar{a}\to 1}\left\{\left(\frac{u(\bar{a})}{r}+c-U_H\right)\left(\underline{q}^{\beta}-\frac{\underline{q}}{\beta(1-\underline{q})+\underline{q}}\right)-\frac{\underline{q}(1-\bar{a})\Delta U}{\beta(1-\underline{q})+\underline{q}}\right\},\end{aligned}$$

and so there is $\epsilon>0$ such that for all $\bar{a}\in(1-\epsilon,1)$ we have that $\mathcal{G}^{\theta}_{\text{det}}(U)<\mathcal{G}^{\theta}_{\text{rand}}(U)$

Optimality of random monitoring following $\theta_{T_{n-1}}=L$ for small $\bar{a}$

The proof follows a similar argument as the one for large $\bar{a}$. The payoff of the deterministic policy satisfies the inequality

$$\begin{aligned}\mathcal{G}^{\theta}_{\text{det}}(U) &< \int_0^{\bar{\tau}}e^{-r\tau}u(\bar{a})\,\mathrm{d}\tau+e^{-r\bar{\tau}}[U_L-c+\bar{a}\Delta U+(\theta-\bar{a})\,e^{-\lambda\bar{\tau}}\Delta U]\\ &= \frac{u(\bar{a})}{r}\left(1-e^{-r\bar{\tau}}\right)+e^{-r\bar{\tau}}[U_L-c+\bar{a}\Delta U\left[1-e^{-\lambda\bar{\tau}}\right]]\end{aligned}$$

Replacing τ_{bind} and taking the limit when $\bar{a}$ goes to zero we find

$$\lim_{\bar{a}\to 0}\mathcal{G}^{\theta}_{\text{det}}(U)<\frac{u(0)}{r}\left(1-e^{-r\tau_{\text{bind}}}\right)+e^{-r\tau_{\text{bind}}}\lim_{\bar{a}\to 0}[U_L-c]$$

Similarly, the payoff of the random policy satisfies

$$\begin{aligned}\mathcal{G}^{\theta}_{\text{rand}}(U) &= \int_0^{\infty}e^{-\left(r+m^*\right)\tau}\left[u\left(x^{\theta}_{\tau}\right)+m^*\mathcal{M}\left(U,x^{\theta}_{\tau}\right)\right]\mathrm{d}\tau\\ &= \frac{1}{r+m^*}\int_0^{\infty}(r+m^*)\,e^{-\left(r+m^*\right)\tau}\left[u\left(x^{\theta}_{\tau}\right)+m^*\mathcal{M}\left(U,x^{\theta}_{\tau}\right)\right]\mathrm{d}\tau\\ &> \frac{[u\left(\frac{a\lambda}{r+m+\lambda}\right)+\frac{a\lambda}{r+m+\lambda}m^*U_H+m^*(1-\frac{a\lambda}{r+m+\lambda})U_L-m^*c]}{r+m^*},\end{aligned}$$

and so the limit when $\bar{a}$ goes to zero is

$$\lim_{\bar{a}\to 0}\mathcal{G}^{\theta}_{\text{rand}}(U)>\frac{r\frac{u(0)}{r}+m^*\lim_{\bar{a}\to 0}(U_L-c)}{r+m^*}$$

In the limit, it must be the case that $\frac{u(0)}{r}\geq\lim_{\bar{a}\to 0}(U_L-c)$: If fact

$$\lim_{\bar{a}\to 0}U_L<\lim_{\bar{a}\to 0}E\left[\int_0^{\infty}e^{-r\tau}u(\theta_{\tau})\,d\tau|\theta_0=L\right],$$

and by dominated convergence

$$\begin{aligned}\lim_{\bar{a}\to 0}E\left[\int_0^{\infty}e^{-r\tau}u(\theta_{\tau})\,\mathrm{d}\tau|L\right] &= \int_0^{\infty}e^{-r\tau}\lim_{\bar{a}\to 0}E\left[u(\theta_{\tau})\,|\theta_0=L\right]\mathrm{d}\tau\\ &= \frac{u(0)}{r}.\end{aligned}$$

From Lemma 6 we have that $e^{-r\tau_{\text{bind}}} > \frac{m^*}{r+m^*}$, and so it follows that

$$\lim_{\bar{a}\to 0} \mathcal{G}^{\theta}_{\text{rand}}(U) - \lim_{\bar{a}\to 0} \mathcal{G}^{\theta}_{\text{det}}(U) > 0.$$

This means that there is $\epsilon > 0$ such that the random policy dominates the deterministic policy for any $\bar{a} \in (0, \epsilon)$

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 $U_\tau = -C_\tau$. The ODE for U_τ is

$$\dot{U}_\tau = (r + m_\tau)U_\tau + \gamma\Sigma_\tau + m_\tau(c + \mathcal{C}). \tag{46}$$

The Hamiltonian is

$$\mathcal{H}(q_\tau, \zeta_\tau, \nu_\tau, \psi_\tau, m_\tau, \tau) = \zeta_\tau\big((r + m_\tau)U_\tau + \gamma\Sigma_\tau + m_\tau(c + \mathcal{C})\big) + \nu_\tau\left((r + \lambda + m_\tau)q_\tau - m_\tau\right) + \psi_\tau\left(q_\tau - \underline{q}\right).$$

The switching function is

$$S(\tau) = (c + \mathcal{C}) - U_\tau - (1 - q_\tau)\nu_\tau. \tag{47}$$

The co-state ν_τ evolves according to

$$\dot{\nu}_\tau = -\lambda\nu_\tau - \psi_\tau. \tag{48}$$

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

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

$$\dot{S}(\tau) = -\dot{U}_\tau + \dot{q}_\tau\nu_\tau - (1 - q_\tau)\dot{\nu}_\tau.$$

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

$$-rU_\tau - \gamma\Sigma_\tau + (rq_\tau + \lambda)\nu_\tau = 0 \tag{49}$$

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

$$-r\dot{U}_\tau - \gamma\dot{\Sigma}_\tau + r\dot{q}_\tau\nu_\tau + (rq_\tau + \lambda)\dot{\nu}_\tau = 0$$

Using the first order condition $m_\tau S(\tau) = 0$ we can write the evolution of C_τ

$$\dot{U}_\tau = rU_\tau + \gamma\Sigma_\tau - m_\tau\nu_\tau(1 - q_\tau).$$

Replacing $\dot{U}_\tau$ and $\dot{\nu}_\tau$ we get

$$r(-rU_\tau - \gamma\Sigma_\tau + m_\tau\nu_\tau(1 - q_\tau)) - \gamma\dot{\Sigma}_\tau + r\nu_\tau((r + \lambda + m_\tau)q_\tau - m_\tau) - (rq_\tau + \lambda)\lambda\nu_\tau = 0,$$

which after some simplification yield

$$-rU_\tau - \gamma\Sigma_\tau - r^{-1}\gamma\dot{\Sigma}_\tau + r^{-1}(r^2q_\tau - \lambda^2)\nu_\tau = 0 \tag{50}$$

Combining equations (49) and (50) we get that along the circular arc we have

$$\nu_\tau = \frac{\gamma\dot{\Sigma}_\tau}{\lambda(r+\lambda)}, \tag{51}$$

where

$$\begin{aligned}\dot{\Sigma}_\tau &= \sigma^2 - 2\lambda\Sigma_\tau \\ \ddot{\Sigma}_\tau &- 2\lambda\Sigma_\tau\end{aligned}$$

Differentiating (51) we get

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

Equation (18b) implies that

$$\dot{\nu}_\tau = -\lambda\nu_\tau - \psi_\tau = -\lambda\nu_\tau < -2\lambda_\tau\nu_\tau,$$

as $\nu_\tau < 0$ if the IC constraint is binding. We have a contradiction; it cannot be the case that $q_\tau > \underline{q}$ and $S(\tau) = 0$ over a singular arc.

Proof of Proposition 6

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

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

where

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

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

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

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

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

To simplify the previous expressions, we define the constants

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

so we can write

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

with derivatives

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

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 $\bar{\tau}^* \in \{-\log(\underline{q})/(r+\lambda), \infty\}$.

D Exogenous News

Proof of Proposition 7

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

$$\begin{aligned}
\mathscr{G}^\theta U = \sup_{\tau, \tilde{m}_\tau} & \int_0^\tau e^{-rt - \tilde{M}_\tau -} \left(u(x_\tau^\theta) - \mu c + \tilde{m}_\tau \mathcal{M}_\theta(U, x_\tau) \right) \mathrm{d}\tau + e^{-r\tau - \tilde{M}_\tau} \mathcal{M}_\theta(U, x_\tau) \\
& \text{subject to} \\
\dot{\tilde{M}}_\tau &= \tilde{m}_\tau \\
\dot{q}_\tau &= (r + \lambda + \tilde{m}_\tau) q_\tau - \tilde{m}_\tau \\
q_\tau &\geq \underline{q}, \ \forall \tau \in [0, \tau] \\
\mu &\leq m_\tau.
\end{aligned}$$

This problem is essentially the same as in the one in the absence of exogenous news with the exception that the non-negativity constraint $m_\tau \geq 0$ is replaced by the lower bound $\tilde{m}_\tau \geq \mu$. 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 $\dot{q}_\tau = 0\big|_{q_\tau = \underline{q}}$ requires that $\tilde{m}_\tau > \mu$. This means that the solution for $\tilde{m}_\tau = m_\tau + \mu$ 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 $m_\tau = 0$ we get $q_\tau = \mu/(r+\lambda+\mu) > \underline{q}$. 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 $q_\tau > \underline{q}$, all $\tau \geq 0$, and so by Lemma 4 $m_\tau = 0$. Suppose that this is not the case and there is a time $\tau^\dagger$ such that $q_{\tau^\dagger} = \underline{q}$, then we have that

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

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

D.1 Necessary Conditions with Asymmetric News

The Hamiltonian for the optimal control problem is

$$\begin{aligned}
\mathcal{H}(\Pi_\tau^L, \Pi_\tau^H, \zeta_\tau, \nu_\tau^L, \nu_\tau^H, \psi_\tau, m_\tau, \tau) = {} & \zeta_\tau((r + m_\tau)U_\tau - x_\tau^\theta - \mu_H x_t^\theta U_H - \mu_L(1 - x_\tau^\theta)U_L - m_\tau \mathcal{M}_\theta(U, x_\tau) \\
& + \psi_\tau(\Pi_\tau^H - \Pi_\tau^L - k/\lambda) + \nu_\tau^H \big((r + \mu_H + m_\tau)\Pi_\tau^H - x_\tau + k\bar{a} + \lambda(1-\bar{a})(\Pi_\tau^H - \Pi_\tau^L) \\
& - (\mu_H + m_\tau)\Pi(H)\big) + \nu_\tau^L\big((r + \mu_L + m_\tau)\Pi_\tau^L - x_\tau + k\bar{a} - \lambda\bar{a}(\Pi_\tau^H - \Pi_\tau^L) - (\mu_L + m_\tau)\Pi(L)\big)
\end{aligned}$$

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

$$\dot{\nu}_\tau^H = -(\mu_H + \lambda(1-\bar{a}))\nu_\tau^H - \psi_\tau + \lambda\bar{a}\nu_\tau^L$$
$$\dot{\nu}_\tau^L = -(\mu_L + \lambda\bar{a})\nu_\tau^L + \psi_\tau + \lambda(1-\bar{a})\nu_\tau^H.$$

The switching function $S(\tau)$ is given by

$$S(\tau) = \mathcal{M}_\theta(U, x_\tau) + \nu_\tau^H(\Pi_\tau^H - \Pi(H)) + \nu_\tau^L(\Pi_\tau^L - \Pi(L)) - U_\tau.$$

We pin-down the boundary condition for the co-state variables ν_τ^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 $\lim_{\tau\to\infty}\nu_\tau^L = \nu_{ss}^L$. 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_τ^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 $\lim_{\tau\to\infty}\nu_\tau^H = \nu_{ss}^H$. 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, $\dot{S}_\tau = 0$, to back out the value of the Lagrange multiplier ψ_τ

$$\psi_\tau\big((\Pi_\tau^H - \Pi_\tau^L) - (\Pi(H) - \Pi(L))\big) = \dot{x}_\tau^\theta(U_H - U_L) - \dot{U}_\tau + (-(\mu_H + \lambda(1-\bar{a}))\nu_\tau^H + \lambda\bar{a}\nu_\tau^L)(\Pi_\tau^H - \Pi(H)) + \nu_\tau^H\dot{\Pi}_\tau^H$$
$$+ (-(\mu_L + \lambda\bar{a})\nu_\tau^L + \lambda(1-\bar{a})\nu_\tau^H)(\Pi_\tau^L - \Pi(L)) + \nu_\tau^L\dot{\Pi}_\tau^L$$

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

$$\psi_\tau = \frac{1}{k/\lambda - \Delta}\Big[\dot{x}_\tau^\theta(U_H - U_L) - \dot{U}_\tau - \big(\mu_H\nu_\tau^H + \mu_L\nu_\tau^L\big)(\Pi_\tau^L - \Pi(L)) + \big((\mu_H + \lambda(1-\bar{a}))\nu_\tau^H - \lambda\bar{a}\nu_\tau^L\big)\left(\frac{1}{r+\lambda} - \frac{k}{\lambda}\right)$$
$$+(\nu_\tau^L + \nu_\tau^H)\dot{\Pi}_\tau^L\Big].$$

Given that the maximized Hamiltonian is linear in the sate variables, a sufficient condition for our conjectured monitoring policy m_τ to be optimal is that the Lagrange multiplier ψ_τ is non-negative whenever the incentive compatibility constraint is binding. The monitoring policy m_τ is positive if and only if this constraint is binding; hence, the sufficiency condition reduces to verify that $\psi_\tau m_\tau \geq 0$. 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.

D.2 Monotonicity of Monitoring Policy with Asymmetric News

Proof. Looking at the phase diagram in Figure 5 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_τ is decreasing in x_τ. 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 (35), respectively. We show next that only one of this roots is consistent with $m_\tau \geq 0$.

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, $m_\tau > 0$ only if $\Pi_\tau < -\alpha/\beta$. 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 $m_\tau > 0$.

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

$$\alpha + \beta\Pi^L_- < 0.$$

In the good news case, $m_\tau > 0$ only if $\Pi_\tau > -\alpha/\beta$. 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 (2010). A reputational theory of firm dynamics. Technical report, working paper, UCLA.

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.

Bundorf, M. K., N. Chun, G. S. Goda, and D. P. Kessler (2009). Do markets respond to quality information? the case of fertility clinics. *Journal of health economics 28*(3), 718–727.

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

Dafny, L. and D. Dranove (2008). Do report cards tell consumers anything they don't already know? the case of medicare hmos. *The Rand journal of economics 39*(3), 790–821.

Dellarocas, C. (2006). How often should reputation mechanisms update a trader's reputation profile? *Information Systems Research 17*(3), 271–285.

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.

Dilmé, F., D. Garrett, et al. (2015). Residual deterrence. Technical report, CEPR Discussion Papers.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003). Is more information better? the effects of "report cards" on health care providers. *The Journal of Political Economy 111*(3), 555–588.

Eccles, R. G., S. C. Newquist, and R. Schatz (2007). Reputation and its risks. *Harvard Business Review 85*(2), 104.

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.

Epstein, A. M. (2000). Public release of performance data: a progress report from the front. *JAMA 283*(14), 1884–1886.

Feng Lu, S. (2012). Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics & Management Strategy 21*(3), 673–705.

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

Figlio, D. N. and M. E. Lucas (2004). What's in a grade? school report cards and the housing market. *The American Economic Review 94*(3), 591–604.

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.

Hoffman, J. V., L. C. Assaf, and S. G. Paris (2001). High-stakes testing in reading: Today in texas, tomorrow? *The reading teacher 54*(5), 482–492.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

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

Jin, G. Z. and P. Leslie (2009). Reputational incentives for restaurant hygiene. *American Economic Journal: Microeconomics 1*(1), 237–267.

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.

Pope, D. G. (2009). Reacting to rankings: evidence from america's best hospitals. *Journal of health economics 28*(6), 1154–1165.

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 (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)*.