Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring*

Felipe Varas † Iván Marinovic‡ Andrzej Skrzypacz §

March 19, 2018

Abstract

This paper studies the design of monitoring policies in dynamic settings with moral hazard. The firm benefits from having a reputation for quality, and the principal can learn the firm's quality by conducting costly inspections. Monitoring plays two roles: An incentive role, because the outcome of inspections affects the firm's reputation, and an informational role because the principal values the information about the firm's quality. We characterize the optimal monitoring policy inducing full effort. It can be implemented by dividing firms into two types of lists: recently inspected and not, with random inspections of firms in the latter.

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-scheduled tests? Should a regulator inspect firms for compliance at pre-scheduled dates or should it use random inspections? For example, how often and how predictably should we test the quality of schools, health care providers, etc.? How should an industry self-regulate a voluntary licensing program, in particular when its members should be tested for compliance? What about the timing of internal audits to measure divisional performance and allocate capital within organizations?

Monitoring is fundamental for the implementation of any regulation. It is essential for enforcement and ultimately the optimal allocation of resources, yet it is costly. According to the OECD (2014), "regulators in many countries are increasingly under pressure to do 'more with less'. A

*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 Duke, Stanford, FTG (Princeton), MIT, Warwick, Columbia, 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

well-formulated enforcement strategy, providing correct incentives for regulated subjects can help reduce monitoring efforts and thus the cost for both business and the public sector, while increasing the efficiency and achieving better regulatory goals."

In practice, information serves multiple roles: on the one hand, it provides valuable information that can improve the allocation of resources in the economy. On the other hand, information is an important incentive device in many markets in which agents are concerned about reputation, and often substitutes monetary awards.[1] The role of information and reputation is particularly important in organizations where explicit monetary rewards that are contingent on performance are not feasible. As Dewatripont et al. (1999) point out in their study of incentives in bureaucracies, in many organizations incentives arise not through explicit formal contracts but rather implicitly through career concerns. This can be the case because formal performance based incentive schemes are difficult to implement due to legal, cultural or institutional constraints. Similarly, regulators may be limited in their power to impose financial penalties on firms and may try to use market reputation to discipline the firms, and fines might be a secondary concern for firms. We believe that our model captures optimal monitoring practices in these situations in which fines and transfers are of second order compared to reputation.

Most real-life monitoring policies fall into one of two classes: random inspections or deterministic inspections – namely inspections that take place at pre-announced dates, for example once a year. At first, neither of these policies seem optimal. A policy of deterministic inspections may induce "window dressing" by the firm: the firm has strong incentives to put in effort toward the inspection date, merely to pass the test, and weak incentives right after the inspection, since the firm knows that it will not be inspected in the near future. On the other hand, random inspection might be wasteful from an information acquisition standpoint. Random inspections are not targeted and may fail to identify cases in which the information acquired is more valuable.

In this paper, we study a model with investment in quality and costly inspections. The objective is to identify trade-offs involved in the design of optimal dynamic monitoring systems. Our main result (Theorem 1) is that when both incentive provision and learning are important, the optimal policy does not take one of these extreme forms. Nevertheless, we show that the optimal policy is simple. It can be implemented by dividing firms into two sets: the recently-inspected ones and the rest. Firms in the second set are inspected randomly, in an order that is independent of their time on the list (that is, with a constant hazard rate). Firms in the first set are not inspected at all. They remain in the first set for a deterministic amount of time (that may depend on the results of the last inspection). When that "holiday" period expires, the principal inspects a fraction of the firms and transfers the remaining fraction to the second set.

The pure deterministic and pure random policies are special cases of our policy. When all firms are inspected at the end of the "holiday" period, the policy is deterministic; when the duration

[1] For example, 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 captures first-order tradeoffs in such markets.

of the “holiday” period shrinks to zero, the policy becomes purely random. We show when these extreme policies can be optimal. When moral hazard is weak, the optimal policy tends to be deterministic. On the other hand, when information gathering has no direct value to the principal, the optimal policy is purely random.

In our model, an agent/firm provides a service and earns profits that are proportional to its reputation, defined as the public belief about the firm's underlying quality. Quality is random but persistent. It fluctuates over time with transitions that depend on the firm's private effort. A principal/regulator designs a dynamic monitoring policy, specifying the timing of costly inspections that fully reveal the firm's current quality. The regulator's flow payoff is convex in firm's reputation, capturing the possibility the regulator values information per se. We characterize the monitoring policy that maximizes the principal's expected payoff (that includes costs of inspections) subject to inducing full effort by the firm.

We extend our two-type benchmark model in two directions. First, to show robustness beyond binary types, we analyze a model where quality follows a mean-reverting Ornstein-Uhlenbeck process and the principal has mean-variance preferences over posterior beliefs. We show that the optimal policy belongs to the same family as that in the binary case and provide additional comparative statics. Second, we consider a model in which additional exogenous news process can reveal information about current quality (as provided, in practice, by consumer reviews or market analysts). In this extension, we assume that conditional on quality, good and bad news are exogenous, and following Board and Meyer-ter-Vehn (2013) allow good and bad news to arrive at different intensities. When preferences are linear, we show that the optimal random monitoring rate is no longer constant. The intuition is that when bad news arrives faster than good news, the moral hazard problem is more acute when the firm's reputation is low, and vice-versa. Since inspections substitute the direct incentives from news, if bad news arrives faster, the optimal monitoring policy calls for high monitoring intensity when the agent's reputation is low.

In some markets inspections play additional roles that our model does not capture. For example, regulators may want to test schools to identify the source of the success of the best performers in order to transfer that knowledge to other schools. Inspections could also be used as direct punishments or rewards – for example, a regulatory agency may punish a non-compliant firm by inspecting it more, or a restaurant guide may reward good restaurants by reviewing it more often. In the last section, we discuss how some of these other considerations could qualitatively affect our results. Our general intuition is that additional considerations (such as dynamic punishments, or direct monetary incentives) that make the moral hazard less severe or the direct value of information higher, lead the optimal policy to favor deterministic monitoring over randomization.

The rest of the paper is organized as follows. We finish this introduction discussing related literature. In Section 2 we introduce the general model and a few applications that could be used to micro-fund our payoffs. In Section 3 we describe the optimal policy when the moral hazard is mild. In Section 4 we describe necessary properties of all incentive-compatible policies (i.e. such that the agent chooses full effort after all histories). In Section 5 we show that when the principal's payoffs are linear in reputation, so that the principal has no direct value of information, the optimal policy is random with a constant hazard rate of inspections. In Section 6 we use optimal control methods to provide our main theorem: characterization of the optimal policy for general preferences. In Section 7 we study a model with quality driven by Brownian shocks. In Section 8 we introduce exogenous news to the model. In Section 9 we discuss how our intuitions would apply to several extensions and we conclude.

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. This 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. 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.[2]

Our 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), Polinsky and Shavell (1984), Reinganum and Wilde (1985), Mookherjee and Png (1989), Bassetto and Phelan (2008), Bond and Hagerty (2010). Kim (2015) compares the level of compliance with environmental norms induced by periodic and exponentially distributed inspections when firms that fail to comply with norms are subject to fines.

We build on the investment and reputation model of Board and Meyer-ter-Vehn (2013) where the firm's quality type changes stochastically. Unlike that paper, we analyze the optimal design

[2] Admittedly, while some existing studies provide evidence in support of the effectiveness of quality report cards, other studies have raised concerns by showing that report cards may induce sellers to game the system in ways that hurt consumers. 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.

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 a 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.[3] 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.

Our paper is also somewhat related to the literature that has explored design of rating mechanisms or reputation systems more broadly. For example, 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), Kovbasyuk and Spagnolo (2016), and Bhaskar and Thomas (2017) for studies of optimal design of rating systems in different environments.

2 Setting

Agents, Technology and Effort: There are two players: a principal and a firm. Time $t \in [0, \infty)$ is continuous. The firm sells a product whose quality changes over time. We model the evolution of quality as in Board and Meyer-ter-Vehn (2013): Initial quality is exogenous and commonly known. At time t, the quality of the product is $\theta_t \in \{L, H\}$, and we normalize $L = 0$ and $H = 1$. Quality changes over time and is affected by the firm's effort. At each time t, the firm makes a private effort choice $a_t \in [0, \bar{a}]$, $\bar{a} < 1$. 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 we illustrate how the analysis can be extended to the case in which quality θ_t can take on a continuum of values and effort affects the drift of the evolution of quality. Note that we bound a_t below one so unlike Board and Meyer-ter-Vehn (2013) quality is random even if the firm exerts full effort. The steady-state distribution of quality when the firm puts in full effort is $\Pr(\theta = H) = \bar{a}$.

Strategies and Information: At time t, the principal can inspect the quality of the product, in which case θ_t becomes public information (we can think of the regulator as disclosing the outcome of inspections to the public. A commitment to truthful disclosures by the regulator is optimal in our setting, given the linearity of the firm payoffs.)

[3] Board and Meyer-ter Vehn (2014) allows quality to be stochastic with full effort.

A monitoring policy specifies an increasing sequence of inspections $(T_n)_{n\geq 1}$ times.[4] Let $N_t \equiv \sup\{n : T_n \leq t\}$ be the counting process associated with $(T_n)_{n\geq 0}$, and denote the natural filtration $\sigma(\theta_s, N_s : s \leq t)$ by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. In addition, let $\mathbb{F}^P = (\mathcal{F}^P_t)_{t\geq 0}$ be the smaller filtration $\sigma(\theta_{T_n}, N_s : n \leq N_t, s \leq t)$ which represents the information available to the principal. The time elapsed between inspections is denoted by $\tau_n \equiv T_n - T_{n-1}$, so a monitoring policy can be represented by a sequence of cumulative density functions, $F_n : \mathbb{R}_+ \cup \{\infty\} \to [0, 1]$ measurable with respect to $\mathcal{F}^P_{T_{n-1}}$ specifying the distribution of τ_n conditional on the information at the inspection date T_{n-1}. The principal commits at time 0 to the full monitoring policy.

We assume that current quality is always privately known by the firm so its information is given by $\mathbb{F}$, but as discussed below, our results extend to the case where the firm does not 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 as driven by the firm's reputation. In particular, denote the market's conjecture about the firm's effort strategy by $\tilde{a} = (\tilde{a}_t)_{t\geq 0}$. Reputation at time t is given by $x_t \equiv E^{\tilde{a}}(\theta_t|\mathcal{F}^P_t)$ 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 the firm's strategy and inspection 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 reputation.[5] The marginal cost of effort is k, hence the firm's expected payoff at time t is

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

In absence of asymmetric information, effort is optimal for the firm if and only if $\lambda/(r+\lambda) \geq k$. We assume throughout the analysis that this condition is satisfied.

The principal discounts future payoffs at the same rate r as the firm. The principal's flow payoff is given by a strictly increasing and convex function of the firm's reputation, $u(\cdot)$. As mentioned previously, the convexity of u captures the possibility that the principal values the information about the firm's quality.

Also, monitoring is costly to the principal: the lump-sum cost of an inspection is c. Hence, the principal's payoff is

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

[4] 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.

[5] 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.

Note that the cost of effort is not part of 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 since it affects agent's effort incentives). Finally, we assume that, for any belief x_t, the principal values effort at least as much as the firm, which means that $u'(0) \geq 1$ which guarantees that full effort is optimal in the first best.

Some comments are in order. First, in some applications, the agent and principal might also care about true quality θ_t, in addition to reputation. For example, a school manager may care about how many students the school attracts thanks to its reputation and about the welfare of those students, which in turn depends on the school's actual quality. The current specification of the principal's payoff already incorporates this possibility.[6] When the agent's preferences are a quasilinear combination of θ_t and x_t the analysis extends directly to this more general case (see Remark 3). Second, we shall study both the case when the principal payoff $u(\cdot)$ is linear and that when it is strictly convex. Again, such convexity of the principal's flow payoff captures situations in which 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. Throughout the paper we ignore the use of monetary transfers –beyond transfers that are proportional to the current reputation.[7] In some settings, other forms of performance-based compensation can be 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. In addition, 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.

Third, we assume the principal can commit to a monitoring policy. There are many possible sources of such commitment. In some instances, commitment is achieved by regulation (for example, in case of aircraft safety, the FAA requires that an aircraft must undergo an annual inspection every 12 calendar months to be legal to operate). In other instances, commitment can be supported by relational contracts. That is, punishing the principal via inferior continuation equilibrium if he deviates. For example, it would call for no more inspections and hence induce no effort. Such commitment via relational concerns would be straightforward in case of deterministic inspections. In case of random inspections, if the principal interacts with many agents, it would be able to commit to inspecting a certain fraction of them in every period to approximate the optimal random

[6]If the principal payoff is $\tilde{u}(\theta_t, x_t)$ then the expected payoff is $u(x_t) = x_t\tilde{u}(H, x_t) + (1 - x_t)\tilde{u}(L, x_t)$.

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

policy we describe. The non-commitment case is beyond the scope of this paper.[8]

Incentive Compatibility and Optimal Policies. We seek to characterize monitoring policies that maximize the principal's payoff among those inducing full effort.[9] Since the firm's best response depends both on the monitoring policy and the principal's conjecture, $\tilde{a}$, incentive compatibility deserves some discussion.

First, we define what it means for an effort policy to be consistent with an equilibrium for a given monitoring policy:

Definition 1. *Fix a monitoring policy $(F_n)_{n\geq 1}$. An equilibrium is a pair of effort and conjectured effort $(\tilde{a}, a)$ such that for every history on the equilibrium path:*[10]

1. *x_t is consistent with Bayes' rule, given $(F_n)_{n\geq 1}$ and $\tilde{a}$.*

2. *a maximizes Π.*

3. *$\tilde{a} = a$.*

Second, we define incentive compatibility of the monitoring policy by requiring existence of an equilibrium with full effort for that policy, and define the optimal policy accordingly.

Definition 2. *A monitoring policy $(F_n)_{n\geq 1}$ is* incentive compatible *if under that policy there exists an equilibrium with $a_t = \bar{a}$. A monitoring policy is* optimal *if it maximizes U over all incentive compatible monitoring polices.*

In other words, we assume the firm chooses full effort whenever there exists an equilibrium given $(F_n)_{n\geq 1}$ that implements full effort (even if there are multiple equilibria).

An optimal policy faces the following trade-off: First, the policy seeks to minimize the cost of inspections subject to maintaining incentives for effort provision (one can always provide incentives for full effort by implementing very frequent inspections, but that would be too costly). Second, since the principal values information per se, 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.

This completes our description of the setting. To further motivate the model, we describe three applications in the next section.

2.1 Examples

Before we begin the analysis, we discuss three applications of the model. They illustrate how the firm and principal payoffs can be micro-founded.

[8] For analysis of costly disclosure without commitment that is triggered by the firm, see Marinovic et al. (2018).

[9] 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 after all histories. Another motivation for focusing on full effort is that in some applications, for example in the case of schools, punishing the firms by implementing low effort might not be practical. We discuss this assumption further in the end of the paper.

[10] 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: School Monitoring. Here we study monitoring of school quality in the presence of horizontal differentiation. Specifically, consider a Hotelling model of school choice with two schools located at opposite extremes of the unit line: School A, with a known constant quality and school B with unknown and evolving quality. The evolution of the quality of school B depends on the school's hidden investment and is unobservable to the public unless a regulator monitors it. Students are distributed uniformly over the unit line. Both schools charge the same tuition and students choose them based on location and perceived quality differences. Assume 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 normalize marginal costs to zero. Hence, the profit flows of schools A and B are

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

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

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

Finally, suppose that the principal's (i.e., the school regulator) 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 the schools' profits does not depend on it (since the two schools just split the subsidy per student, reputation only affects the distribution of profits). 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 2: Quality Certification. Consider a version of the classic problem of moral hazard in quality provision, as studied by the reputation literature (see e.g., Mailath and Samuelson (2001)). There are two firms. The product of firm 2 (good 2) has a known quality $x_2 \in (0,1)$, while the product of firm 1 (good 1) – which is the firm we analyze– has random quality that is either high or low, $\theta_1 \in \{0,1\}$ with reputation denoted by x_1. Each firm produces a unit (flow) of the good per period. There are $N \geq 3$ buyers with types q_j that represent a buyer's preference for quality: Each buyer j has type q_j with $q_1 > q_2 = q_3 = ... = q$, and if agent j gets the good with expected quality x and pays p their consumer surplus is

$$q_j x - p.$$

Prices and allocations are set competitively as follows. When $x_1 < x_2$ the efficient allocation is that buyer 1 gets good 2 and any of the other buyers gets good 1. Competition between the less-efficient buyers drives the price of good 1 to $p_1 = qx_1$ (these buyers get no surplus), while the price of good 2 is the smallest price such that agents $j \geq 2$ do not want to outbid agent 1 for it:

$$qx_1 - p_1 = qx_2 - p_2 \Rightarrow p_2 = qx_2.$$

When $x_1 > x_2$, then the efficient allocation is that agent 1 gets good 1, and, by an analogous reasoning, competition implies that prices are $p_2 = qx_2$ and $p_1 = qx_1$: Therefore, for all levels of x_1 the price of the output of firm 1 is $p_1 = qx_1$. Suppose the planner wants to maximize total social surplus. Because the less efficient buyers compete away all the surplus, the social surplus is

$$TS = p_1 + p_2 + CS_{1,}$$

where CS_1 is the surplus of agent 1, and so we have that

$$CS_1 = \begin{cases} q_1 x_2 - p_2 \text{ if } x_1 < x_2 \\ q_1 x_1 - p_1 \text{ if } x_1 \geq x_2 \end{cases}$$

which means that the surplus flow per period is

$$u(x_1) = \begin{cases} qx_1 + q_1 x_2 \text{ if } x_1 < x_2 \\ q_1 x_1 + qx_2 \text{ if } x_1 \geq x_2 \end{cases}$$

The surplus is a convex function because $q_1 > q$: Intuitively, while prices are linear in expected quality (reputation), consumer surplus is convex because reputation affects the allocation of goods – information about the true quality of product 1 allows to allocate it more efficiently among the agents.[11] The principal's preferences are linear if $q_1 = q$ because information has no allocative role. This corresponds to the setting in Mailath and Samuelson (2001) and Board and Meyer-ter-Vehn (2013) who consider a monopolist selling a product to a competitive mass of buyers.

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.[12] As in Stein (1997) and Harris and Raviv (1996), we assume managers enjoy a private benefit from larger investments. In particular, assume the manager enjoys a private benefit at time t of $b * \iota_t$ from investment ι_t.[13] Projects arrive according to a Poisson process $\tilde{N}_t$ with arrival intensity μ. The manager's expected payoff is

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

Similarly, the division's cash-flows follow a compound Poisson process $(Y_t)_{t\geq 0}$ given by

$$Y_t = \sum_{i=1}^{\tilde{N}_t} f(\theta_{t_i}, \iota_{t_i}),$$

where $f(\theta_t, \iota_t) = \theta_t - \gamma(\iota_t - \theta_t)^2$ is a quadratic production function similar to the one used in Jovanovic and Rousseau (2001). At each time t that a project arrives, the headquarters decides how much resources allocate to the division, and the optimal investment choice of the headquarter is to allocate $\iota_t = \arg\max_\iota E[f(\theta_t, \iota)|\mathcal{F}^P_{t^-}]$ resources to the division, so $\iota_t = x_t$.[14] Hence, the manager's expected flow payoff is

$$\pi_t = \mu b x_t - k a_t,$$

and the principal's expected flow payoff is

$$\begin{aligned} u(x_t) &= \mu \left(x_t - \gamma \mathrm{Var}\left[\theta_t | \mathcal{F}^P_t\right]\right) \\ &= \mu \left((1-\gamma) x_t + \gamma x_t^2 \right). \end{aligned}$$

In the baseline model, we assume that monitoring is the only source of information about θ available to the headquarter. In this application it is natural to assume that the headquarter also learns

[11] In this example $u(x)$ is piece-wise linear. It is an artifact of having two types of agents and two products since there are only two possible allocations. It is possible to construct a model with a continuum of agent types and continuum of goods where the allocation changes continuously in x and the resulting consumer surplus is strictly convex.

[12] Some examples are found in Hart and Moore (1995), Harris and Raviv (1996), and Harris and Raviv (1998). Motta (2003) studies a model of capital budgeting with empire building preferences and career concerns.

[13] 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.

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

about the current productivity once the cash-flows arrive. We study the possibility of exogenous news arrivals in Section 8.

3 Optimal Monitoring with weak Moral Hazard

As an intermediate step toward characterizing the optimal policy in the general model, we study a relaxed problem that ignores the agent's incentive constraint. When $u(\cdot)$ is convex and both the cost of monitoring c and effort k are small enough, the solution of such a relaxed problem satisfies the agent's incentive constraint being thus the optimal policy.[15] Moreover, even if moral hazard is severe, the trade-offs identified in the unconstrained problem influence the structure of the optimal policy.

Consider the evolution of reputation between two inspection dates. Given that the firm exerts full effort, $a = \bar{a}$, reputation evolves according to

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

between inspection dates (T_{n-1}, T_n). Therefore, given $\theta_{T_{n-1}} = \theta$, the firm's reputation at time $T_{n-1} + \tau < T_n$ is

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

In the relaxed problem (i.e., ignoring the incentive constraint) the principal solves the following stochastic control problem

$$\begin{cases} U(x_0) = \sup_{(T_n)_{n\geq 1}} E\left[\int_0^\infty e^{-rt}u(x_t)\mathrm{d}t - \sum e^{-rT_n}c \middle| \mathcal{F}_0^P\right] \\ \text{subject to:} \\ \dot{x}_t = \lambda(\bar{a} - x_t) \quad \forall t \in [T_{n-1}, T_n) \\ x_{T_{n-1}} = \theta_{T_{n-1}}. \end{cases} \tag{2}$$

The optimal policy is Markovian in reputation. Denoting by $\mathcal{A}$ the set of reputations that lead to immediate inspection, the value function solves the Hamilton-Jacobi-Bellman (HJB) equation

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

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

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

[15] The solution of this relaxed problem also characterizes the optimal policy when effort (but not quality) is observable (recall we assume $u'(0) \geq 1$ so full effort is optimal in the first best).

and the threshold $\hat{x} \in \{\underline{x}, \overline{x}\}$ satisfies the boundary conditions:

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

$$U'(\hat{x}) = U(1) - U(0). \tag{4b}$$

Hence, we have the following standard result:

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

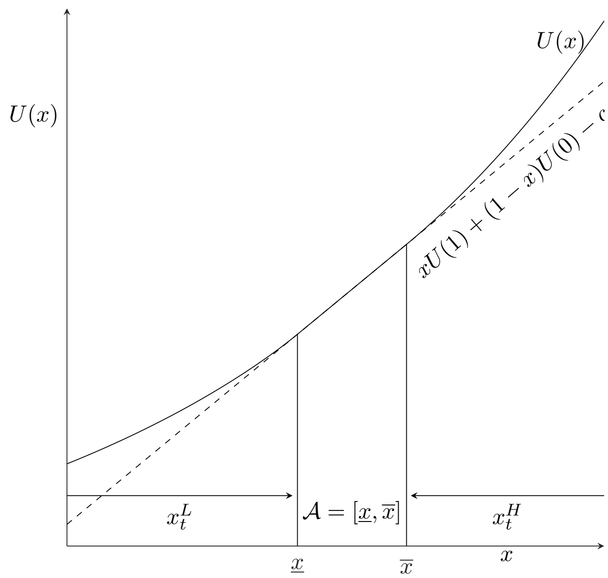


Figure 1: Value Function. The optimal policy requires to monitor whenever reputation enters the audit set, $x_t \in \mathcal{A}$.

Figure 1 illustrates the principal's payoff as a function of beliefs. Observe that after an inspection 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 the drift in beliefs is 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 too large) the optimal policy is to never monitor the firm but 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 monitoring). By contrast, as $u(\cdot)$ becomes more convex, the monitoring region widens leading to greater frequency of monitoring eventually leading to the incentive constraint being always slack which, as mentioned above, implies that the solution to the relaxed is the optimal monitoring policy.

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, $t - T_n$, and the outcome observed in the last review, θ_{T_n}. 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 $T_n = T_{n-1} + \tau_{\theta_{T_{n-1}}}$.[16]

Remark 1. *This representation of the optimal monitoring policy applies to the case in which both τ_L and τ_H are finite. Depending on the specific parameters of the model, either τ_L or τ_H can be infinite, or in other words there is no further monitoring after some outcomes. In terms of the policy specified as a function of beliefs 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 the threshold that is different from $\bar{a}$.*

4 Incentive Compatible Policies

The solution of the relaxed problem may violate incentive constraints when inspections are too costly hence infrequent. This would induce the firm to shirk some times particularly when the moral hazard issue is most severe, namely right after an inspection. In this section, we characterize necessary and sufficient conditions for a monitoring policy to satisfy the incentive compatibility constraints. In the following 2 sections, we use this characterization to derive optimal policies for $u(x)$ linear and convex, respectively.

To tackle this problem, let us begin by considering the firm's continuation payoff under full effort at time T_{n+1}, where T_{n+1} denotes the next review date:

$$\Pi_{T_{n+1}} = E^{\bar{a}}\left[\int_{T_{n+1}}^{\infty} e^{-r(t-T_{n+1})}(x_t - k\bar{a})\mathrm{d}t\Big|\mathcal{F}_{T_{n+1}}\right]$$
$$= \int_{T_{n+1}}^{\infty} e^{-r(t-T_{n+1})}\big(E^{\bar{a}}[x_t|\mathcal{F}_{T_{n+1}}] - k\bar{a}\big)\mathrm{d}t.$$

This expression corresponds to the firm's continuation value at time T_{n+1} assuming the firm exerts full effort thereafter. Simply put, it represents the expected present value of the firm future revenues net of effort costs. A key insight in the derivation of the incentive compatibility constraint is that

[16] 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.

the law of iterated expectations along with the Markov nature of the quality process, θ_t, imply that $E^{\bar{a}}[x_t|\mathcal{F}_{T_{n+1}}] = E^{\bar{a}}[\theta_t|\theta_{T_{n+1}}]$, and this is equal to

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

Therefore, in any incentive-compatible monitoring policy, if the quality at time T_n is public, the firm's continuation value at time T_{n+1} given $\theta_{T_{n+1}} = \theta$ is:

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

The first term is the NPV of revenue flows given steady-state reputation; the second is the deviation from the steady-state flows given that at time T_{n+1} the firm re-starts with an extreme reputation, and the last term is the NPV of effort costs. Importantly, since the firm's payoffs are linear in reputation and the firm incurs no direct cost of inspections, these continuation payoffs are independent of the future monitoring policy. That dramatically simplifies the characterization of incentive compatible policies as we show next. Moreover, because the continuation value at time T_{n+1} is independent of the previous history of effort (it depends on effort only indirectly via $\theta_{T_{n+1}}$), 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

$$\theta_t = e^{-\lambda t}\theta_0 + \int_0^t e^{-\lambda(t-s)}(\lambda a_s \mathrm{d}s + \mathrm{d}Z_s), \tag{6}$$

where Z_t is a martingale, corresponding to the compensated Poisson process of changes in quality.

Consider the agent's effort incentives. Effort may affect the firm payoff by changing its quality and thereby the outcome of future inspections. Informally, the effect of effort on future quality is $\partial\theta_{T_{n+1}}/\partial a_t = \lambda e^{-\lambda(T_{n+1}-t)}\mathrm{d}t$, so the marginal benefit of exerting effort over an interval of size $\mathrm{d}t$ is

$$E[\lambda e^{-(r+\lambda)(T_{n+1}-t)}|\mathcal{F}_t](\Pi(H) - \Pi(L))\mathrm{d}t.$$

This is intuitive: having high quality rather than low quality at the inspection time yields the firm a benefit $\Pi(H) - \Pi(L)$. Also, a marginal increase in effort leads to higher quality with probability (flow) $\lambda\mathrm{d}t$. However, to reap the benefits of high quality, the firm must wait till the next review date, T_{n+1}, facing the risk of an interim (i.e., before the inspection takes place) drop in quality. Hence, the benefit of having high quality at a given time must be discounted according to the interest rate r and the quality depreciation rate λ. On the other hand, the marginal cost of effort is $k\mathrm{d}t$. Combining these observations we can express the incentive compatibility condition for full effort as follows.

Proposition 1. *Full effort is incentive compatible if and only if for all $n \geq 0$,*

$$\frac{1}{r+\lambda} E\left[e^{-(r+\lambda)(T_{n+1}-t)}\middle|\mathcal{F}_t\right] \geq \frac{k}{\lambda} \qquad \forall t \in [T_n, T_{n+1})$$

Proof. The first step is to define the martingale Z_t in equation (6). Let $N_t^{LH} = \sum_{s\le t} \mathbf{1}_{\{\theta_{s-}=L,\theta_s=H\}}$ and $N_t^{HL} = \sum_{s\le t} \mathbf{1}_{\{\theta_{s-}=H,\theta_s=L\}}$ be counting processes indicating the number of switches from L to H and from H to L, respectively. The processes

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

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

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

which leads to equation (6). Full effort is incentive compatible if and only if for any deviation $\hat{a}_t$ (with an associated process for quality $\hat{\theta}_t$)

$$
\begin{aligned}
E^{\bar{a}}\left[\int_t^{T_{n+1}} e^{-r(s-t)}(x_s - k\bar{a})\mathrm{d}s + e^{-r(T_{n+1}-t)}\big(\theta_{T_{n+1}}\Pi(H) + (1-\theta_{T_{n+1}})\Pi(L)\big)\big|\mathcal{F}_t\right] \ge \\
E^{\hat{a}}\left[\int_t^{T_{n+1}} e^{-r(s-t)}(x_s - k\hat{a}_s)\mathrm{d}s + e^{-r(T_{n+1}-t)}\big(\hat{\theta}_{T_{n+1}}\Pi(H) + (1-\hat{\theta}_{T_{n+1}})\Pi(L)\big)\big|\mathcal{F}_t\right]
\end{aligned}
$$

Letting $\Delta \equiv \Pi(H) - \Pi(L)$ and replacing the solution for θ_t in (6), we can write the incentive compatibility condition as

$$
E^{\hat{a}}\left[\int_t^{T_{n+1}} e^{-r(s-t)}\big(\lambda e^{-(r+\lambda)(T_{n+1}-s)}\Delta - k\big)(\bar{a} - \hat{a}_s)\mathrm{d}s\big|\mathcal{F}_t\right] \ge 0.
$$

For any deviation we have that

$$
\begin{aligned}
E^{\hat{a}}\left[\int_t^{T_{n+1}} e^{-r(s-t)}\big(\lambda e^{-(r+\lambda)(T_{n+1}-s)}\Delta - k\big)(\bar{a} - \hat{a}_s)\mathrm{d}s\big|\mathcal{F}_t\right] = \\
E^{\hat{a}}\left[\int_t^{\infty} \mathbf{1}_{\{T_{n+1}>s\}} e^{-r(s-t)}\big(\lambda E_s[e^{-(r+\lambda)(T_{n+1}-s)}]\Delta - k\big)(\bar{a} - \hat{a}_s)\mathrm{d}s\big|\mathcal{F}_t\right]
\end{aligned}
$$

So, we can write the incentive compatibility condition as

$$
E^{\hat{a}}\left[\int_t^{T_{n+1}} e^{-r(s-t)}\big(\lambda E_s[e^{-(r+\lambda)(T_{n+1}-s)}\big|\mathcal{F}_t]\Delta - k\big)(\bar{a} - \hat{a}_s)\mathrm{d}s\right] \ge 0
$$

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

In essence, this condition says that for a monitoring policy to be incentive compatible the next inspection must be sufficiently close, in expectation. Incentive compatibility imposes a lower bound on the expected discounted inspection date $E_t\left[e^{-(r+\lambda)(T_n - t)}\right]$. What matters for incentives at a given point in time is not necessarily the monitoring intensity at that point but the cumulative discounted likelihood of monitoring in the near future. Future monitoring affects incentives today because effort has a persistent effect on quality, hence shirking today can lead to a persistent drop in quality that can be detected by the principal in the near future.

As mentioned above, future inspections are discounted both by r and the switching intensity λ because effort today matters insofar as quality is persistent (or switching is not so frequent). Notice that the incentive compatibility constraint is independent of the true quality of the firm at time t, so the incentive compatibility condition is the same if the firm does not observe the quality process, which is why the optimal monitoring policy is the same wether the firm observes quality or not. The incentive compatibility constraint is independent of θ_t because effort enters linearly in the law of motion of θ_t and the cost of effort is independent of θ_t, which means that the marginal benefit and cost of effort are independent of θ_t, and so is the incentive compatibility constraint. In fact, we have the incentive compatibility constraint in Proposition 1 apply to more general processes for quality.

Remark 2. *Notice that Proposition 1 holds for any process for quality satisfying the stochastic differential equation*

$$\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \mathrm{d}Z_t, \tag{7}$$

where Z_t is a martingale. In particular, it holds when Z_t is a Brownian motion, so quality follows an Ornstein-Uhlenbeck process. This case is considered later in Section 7. The binary setting is a particular case of (7) in which Z_t is the compensated Poisson process defined in the proof of Proposition 1.

Remark 3. *Proposition 1 can be extended to the case in which the agent also cares about quality and has a quasilinear flow payoff $v(\theta_t) + x_t$. In this case, the incentive compatibility constraint becomes*

$$\frac{1}{r+\lambda}E\left[e^{-(r+\lambda)(T_{n+1}-t)}\middle|\mathcal{F}_t\right] \geq \frac{k}{\lambda} - \frac{v(1)-v(0)}{r+\lambda} \qquad \forall t \in [T_n, T_{n+1})$$

All the results extend to this case by setting the cost of effort equal to $k - \lambda(v(1) - v(0))/(r+\lambda)$.

5 Linear Payoffs: Information without Direct Social Value

In this section we analyze the case in which the principal's flow payoff $u(\cdot)$ is linear. As discussed above, this case captures applications where the principal is say an industry self-regulatory organization that is not directly concerned about consumer surplus but wishes to maximize the industry's expected profits.

Under linear payoffs, information has no direct value to the regulator. Hence, the principal's problem boils down to minimizing the expected monitoring costs, subject to the incentive compatibility constraints. Accordingly, using Proposition 1, we can reduce the principal's problem to the following cost minimization problem:

$$\begin{cases} C_0 = \inf_{(T_n)_{n\geq 1}} E\left[\sum_{n\geq 1} e^{-rT_n} c \middle| \mathcal{F}_0^P\right] \\ \text{subject to:} \\ \frac{k}{\lambda} \leq \frac{1}{r+\lambda} E\left[e^{-(r+\lambda)(T_{n+1}-t)} | \mathcal{F}_t\right] \ \forall t \in [T_n, T_{n+1}). \end{cases} \tag{8}$$

The principal aims to minimize expected monitoring costs subject to the agent always having an incentive to exert effort. The optimal monitoring policy in this case is simple, consisting of random inspections with a constant hazard rate:

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

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

where

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

The intuition for Proposition 2 follows from the fact that in (8) future monitoring is discounted by r in the objective function and by $r+\lambda$ in the constraints (as previously discussed, inspections have a discounted effect on incentives because quality depreciates over time.) As a result, the optimal monitoring policy front-loads inspections in a way that the incentive compatibility constraints bind in all periods. This implies that the optimal intensity of monitoring is constant at $m_\tau = m^* = (r+\lambda)\underline{q}/(1-\underline{q})$, and that there are no deterministic reviews nor atom, or else the incentive constraint would be slack some time prior to the review, in which case the principal could save some monitoring expenses without violating the firm's incentive to exert full effort.

Remark 4. *As mentioned in Remark 2, the incentive compatibility characterization in Proposition 1 holds for general processes of the form*

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

Proposition 2 extends to the case in which Z_t is Brownian motion (that is, Ornstein Uhlenbeck process) or a more general general Levy process. The key assumption here is that the drift of quality is linear in effort a_t.

Remark 5. *It follows directly from the proof of Proposition 2 that the result extends to the case in which the principal and the firm 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.*

6 Convex Payoffs: Information with Direct Social Value

In many applications information has direct value to the principal. We capture this possibility by assuming that the principal's utility flow $u(x)$ is convex in beliefs. In this case, the principal designs the monitoring policy facing a trade-off between cost minimization and information acquisition. Because of this dual role of monitoring, we cannot use as simple an argument as in the linear case; we need to analyze the full optimization problem.

According to Proposition 1, the incentive compatibility constraint depends only on the distribution of $\tau_{n+1} \equiv T_{n+1} - t$ and is independent of the distribution of monitoring times during future monitoring cycles, $\{T_{n+k}\}_{k\geq 2}$. Let

$$\mathcal{M}(\mathbf{U}, x) \equiv xU_H + (1-x)U_L - c$$

be the expected payoff at the inspection date given beliefs x and continuation payoffs $\mathbf{U} \equiv (U_L, U_H)$. We can write the principal problem recursively using θ_{T_n} as a state variable at time T_n as follows[17]

$$\begin{cases} \mathscr{G}^\theta(\mathbf{U}) = \sup_F \int_0^\infty \left(\int_0^\tau e^{-rs} u(x_s^\theta) \mathrm{d}s + e^{-r\tau} \mathcal{M}(\mathbf{U}, x_\tau^\theta) \right) \mathrm{d}F(\tau) \\ \text{subject to} \\ \underline{q} \leq E[e^{-(r+\lambda)(\tau_{n+1}-\tau)} \mid \tau_{n+1} \geq \tau, \theta_0 = \theta] \;\; \forall \tau, \end{cases} \tag{9}$$

where $x_\tau^\theta \equiv \theta e^{-\lambda\tau} + \bar{a}\left(1 - e^{-\lambda\tau}\right)$.

The principal payoff is given by the fixed point $\mathscr{G}^\theta(\mathbf{U}) = \mathbf{U}$, $\theta \in \{L, H\}$. The following lemma establishes that a fixed point exists, is unique, and that the supremum in (9) is attained.

Lemma 1 (Existence)**.** *There exists a unique fixed point $\mathscr{G}^\theta(\mathbf{U}) = \mathbf{U}$, $\theta \in \{L, H\}$. Furthermore, for any continuation payoff $\mathbf{U}$, there exists a monitoring policy F^* solving the maximization problem in* (9)*.*

We reformulate the problem as an optimal control problem with state constraints. For this, we define the state variable

$$q_\tau \equiv E\left[e^{-(r+\lambda)(\tau_{n+1}-\tau)} \middle| \tau_{n+1} \geq \tau, \theta_0 \right],$$

where the expectation is taken over the next monitoring time, τ_{n+1}.

That is, q_τ, represents the expected discounted time till the next review, where the effective discount rate incorporates the depreciation rate λ. The incentive compatibility constraint in Proposition 1 becomes $q_\tau \geq \underline{q}$. The next step is to derive the law of motion of $(q_\tau)_{\tau\geq 0}$ to use it as a state variable in the principal's optimization problem. It is convenient to express the optimization problem in terms of the hazard measure $M : \mathbb{R}_+ \cup \{\infty\} \to \mathbb{R}_+ \cup \{\infty\}$ defined by $1 - F(\tau) = e^{-M_\tau}$. M_τ is a non-decreasing function and by the Lebesgue decomposition theorem, it can be decomposed

[17] Notice that because θ_t is a Markov process and the Principal problem is Markovian, we can wlog reset the time to zero after every inspection and denote the value of θ_t at time T_n by θ_0.

into its continuous and its discrete part[18]

$$M_\tau = M^c_\tau + M^d_\tau.$$

Thus, we can write

$$\begin{aligned} q_\tau &= \int_\tau^\infty e^{-(r+\lambda)(s-\tau)} \frac{\mathrm{d}F(s)}{1-F(\tau-)} \\ &= \int_\tau^\infty e^{-(r+\lambda)(s-\tau)-(M_{s-}-M_\tau)} \mathrm{d}M^c(s) + \sum_{s>\tau} e^{-(r+\lambda)(s-\tau)-(M_{s-}-M_\tau)}(1-e^{-\Delta M^d_s}) \end{aligned}$$

At any point of continuity we have that

$$\mathrm{d}q_\tau = (r+\lambda)q_\tau \mathrm{d}\tau - (1-q_\tau)\mathrm{d}M^c_\tau, \tag{10}$$

while at any point of discontinuity we have that

$$q_{\tau-} = e^{-\Delta M^d_\tau} q_\tau + (1-e^{-\Delta M^d_\tau}). \tag{11}$$

The next lemma summarizes the recursive formulation for the incentive compatibility constraint.

Lemma 2 (Incentive Compatibility)**.** *For any monitoring policy M_τ, let $\bar{\tau} = \inf\{\tau \in \mathbb{R}_+ \cup \{\infty\} : F(\tau) = 1\}$. For any $\tau \in [0, \bar{\tau}]$, let q_τ be the solution to equations (10) and (11) with terminal condition $q_{\bar{\tau}} = 1$. Full effort is incentive compatible if and only if $q_\tau \geq \underline{q}$, for all $\tau \in [0, \bar{\tau}]$.*

The significance of Lemma 2 is that it allows us to represent the optimal monitoring policy recursively, with q_τ being the state variable, and use the tools of optimal control theory to study the optimal policy. To formulate the principal problem as an optimal control with state constraints, it is convenient to use the principal's continuation value, U^θ_τ, as an additional state variable (in the rest of the paper, we omit the dependence of M and U on that last outcome θ). The continuation payoff for the principal given a monitoring policy M_τ and continuation payoffs $\mathbf{U}$ is

$$\begin{aligned} U_\tau = \int_\tau^\infty e^{-r(s-\tau)-(M_{s-}-M_\tau)} u(x^\theta_s)\mathrm{d}s + \int_\tau^\infty e^{-r(s-\tau)-(M_{s-}-M_\tau)} \mathcal{M}(\mathbf{U}, x^\theta_s)\mathrm{d}M^c_s \\ + \sum_{s>\tau} e^{-r(s-\tau)-(M_{s-}-M_\tau)}(1-e^{-\Delta M^d_s})\mathcal{M}(\mathbf{U}, x^\theta_s) \end{aligned}$$

[18]With some abuse of notation, we are allowing $M_\tau = \infty$ to incorporate the event that there is monitoring with probability 1 at time τ. Technically, this means that M is not a σ-finite measure so the Lebesgue decomposition does not follow directly. The definition $1 - F(\tau) = e^{-M_\tau}$ is convenient in terms of notation, and the decomposition of M_τ is valid for $\tau < \bar{\tau} = \inf\{\tau > 0 : F(\tau) = 1\}$. Thus, the definition $1 - F(\tau) = e^{-M_\tau}$ should be interpreted as a shorthand for

$$1 - F(\tau) = \begin{cases} e^{-M^c_\tau} \prod_{0<s<\tau} e^{-\Delta M^d_s} & \text{if } \tau < \bar{\tau} \\ 0 & \text{if } \tau \geq \bar{\tau} \end{cases}$$

At any point of continuity, the continuation value satisfies the differential equation

$$\mathrm{d}U_\tau = \Big(rU_\tau - u(x_\tau^\theta)\Big)\,\mathrm{d}\tau + \Big(U_\tau - \mathcal{M}(\mathbf{U}, x_\tau^\theta)\Big)\,\mathrm{d}M_\tau^c, \tag{12}$$

while at any point of discontinuity, the jump in the continuation value is given by

$$U_{\tau-} = (1 - e^{-\Delta M_\tau^d})\mathcal{M}(\mathbf{U}, x_\tau^\theta) + e^{-\Delta M_\tau^d} U_\tau. \tag{13}$$

We can now state the optimal control problem associated with (9) as follows:

$$\begin{cases} \mathscr{G}^\theta(\mathbf{U}) = \max_{M_\tau}\ U_0 \\ \text{subject to} \\ \mathrm{d}U_\tau = \big(rU_\tau - u(x_\tau^\theta)\big)\,\mathrm{d}\tau + \big(U_\tau - \mathcal{M}(\mathbf{U}, x_\tau^\theta)\big)\,\mathrm{d}M_\tau^c,\ U_{\bar\tau} = \mathcal{M}(\mathbf{U}, x_{\bar\tau}^\theta) \\ U_{\tau-} = (1 - e^{-\Delta M_\tau^d})\mathcal{M}(\mathbf{U}, x_\tau^\theta) + e^{-\Delta M_\tau^d} U_\tau \\ \mathrm{d}q_\tau = (r+\lambda) q_\tau \mathrm{d}t - (1 - q_\tau)\mathrm{d}M_\tau^c,\ q_{\bar\tau} = 1 \\ q_{\tau-} = e^{-\Delta M_\tau^d} q_\tau + (1 - e^{-\Delta M_\tau^d}) \\ q_\tau \in [\underline{q}, 1] \end{cases} \tag{14}$$

Next, we provide the general characterization of the optimal monitoring policy.

Theorem 1 (Optimal Monitoring Policy)**.** *Suppose that $u(x)$ is strictly convex. Let F_θ^* be an optimal policy given $\theta_{T_n} = \theta$, then F_θ^* is either:*

1. *Deterministic with an inspection date at time*

$$\hat\tau_\theta^* \le \tau^{bind} \equiv \frac{1}{r+\lambda} \log \frac{1}{\underline{q}},$$

   where τ^{bind} is the review time that makes the incentive constraint bind at time zero. So the monitoring distribution is $F_\theta^*(\tau) = \mathbf{1}_{\{\tau \ge \hat\tau_\theta^*\}}$.

2. *Random with a monitoring distribution*

$$F_\theta^*(\tau) = \begin{cases} 0 & \textit{if } \tau \in [0, \hat\tau_\theta^*) \\ 1 - p_\theta^* e^{-m^*(\tau - \hat\tau_\theta^*)} & \textit{if } \tau \in [\hat\tau_\theta^*, \infty] \end{cases}$$

   where $\hat\tau_\theta^* \le \tau^{bind}$ and

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

$$p_\theta^* = \frac{1 - e^{(r+\lambda)\hat\tau_\theta^*}\underline{q}}{1-\underline{q}}.$$

Theorem 1 states that the optimal policy belongs to the following simple family of monitoring policies. For a given result of the last inspection, there is a time $\hat{\tau}^*$ such that the optimal policy calls for no monitoring until that time, a strictly positive probability (an atom) at that time, and then monitoring with a constant hazard rate. One extreme policy in that family is to inspect for sure at $\hat{\tau}^*$: the timing of the next inspection is deterministic and the incentive constraints bind at most right after an inspection (so that $\tau_\theta^* \leq \tau^{\text{bind}}$). On the other extreme, the optimal policy specifies $\hat{\tau}^* = 0$ and no atom, so that we obtain monitoring with a constant hazard rate, as in the linear $u(x)$ case. In between, the atom at $\hat{\tau}^*$ is such that the incentive constraints hold exactly at $\tau = 0$ (and then are slack till $\hat{\tau}^*$ and bind forever after).

Such a simple policy can be implemented by a principal who monitors many firms by dividing them into two sets: the recently-inspected firms and the rest. Firms in the second set are inspected randomly, in an order that is independent of their time on the list. Firms in the first set are not inspected at all. They remain in the first set for a deterministic amount of time that may depend on the results of the last inspection. When the time in the first set ends, the principal inspects a fraction of the firms (and resets their clock in the first set). The remaining fraction of firms is moved to the second set. This policy is described by two parameters: times in the first set after the good and bad results. Given those times, the fractions of firms inspected from each of the sets are uniquely pinned down by incentive constraints.

The economic intuition for why the optimal policy takes this form is as follows. In choosing the optimal timing, there are two trade-offs that echo the two benchmark cases we have analyzed above. On one hand, as we learned in the linear case, to minimize costs subject to satisfying incentive constraints, it is optimal to front-load incentives and hence monitor with a constant hazard rate. On the other hand, when $u(x)$ is convex, as reputation moves from one of the extremes towards the steady-state, inspections generate additional value from learning. As we saw in the unconstrained case, that value of learning is zero at the extreme reputations and grows fast (because beliefs move the fastest then).

If $u(x)$ is sufficiently convex, the benefit of delaying inspections to increase the value of learning is greater than the associated increase in cost, caused by departure from the constant hazard rate policy. However, over time, the value from learning grows slower and slower. For example, as reputation gets closer to the steady-state, it moves very little and at that time the tradeoffs are approximately the same as in the linear case. That explains why the optimal policy eventually implements a constant hazard rate. Convexity of $u(x)$ implies that there is a unique time when the benefits of delaying inspections balance the increased cost of inspections.

Solving this problem is challenging due to the presence of state constraints along with the associated fixed point problem. The formal proof relies on necessary conditions from Pontryagin's maximum principle for problem with state constraints (see Hartl et al. (1995) for a survey and Seierstad and Sydsaeter (1986) for a textbook treatment). Using these necessary conditions, we show that the optimal policy belongs to the family of distributions characterized in Theorem 1. This reduces the problem of finding the optimal policy to a one-dimensional maximization problem.

This simplifies the analysis significantly and reduces the problem of finding the optimal policy to solving the simpler fixed point problem in (22).

Remark 6. *The optimal control problem in* (14) *is nonstandard. Traditional optimal control theory restricts attention to trajectories of the state variables that are absolutely continuous; however, the set of admissible trajectories in* (14) *corresponds to trajectories of bounded variation. Moreover, we must allow for policies in which* M^c_τ *is singular to guarantee existence of a solution.*[19] *We analyze the Principal's optimization problem in* (14) *using the theory of optimal control of differential equations driven by measures developed by Rishel (1965) and extended to problems with state constraints by Arutyunov, Karamzin, and Pereira (2005). That being said, in our problem, the optimality conditions from this general formulation coincides with the ones obtained using the more traditional maximum principle in Seierstad and Sydsaeter (1986).*

An alternative approach is to rewrite the optimization problem in (9) *as a linear program and use weak duality to verify that the policy in Theorem 1 is optimal. This approach works well with hindsight, once we already know the form of the optimal policy. This is the approach that we use in the online appendix, where we solve the model in discrete time, and we show that the optimal policy converges to the one in Theorem 1.*[20] *However, it is difficult to uncover the main features of the optimal policy directly from the linear program. The variational approach in optimal control, allows us to uncover some key aspects of the optimal policy (lemmas 3 and 4), which lead to the characterization in Theorem 1.*

The analysis of the Principal problem follows 5 steps. In the first two steps, we derive necessary conditions that the optimal monitoring policy must satisfy. In Step 3, we show that the principal never monitors using a positive hazard rate if the incentive compatibility constraint is slack. In Step 4, we show that the monitoring distribution has at most one atom. In Step 5 we show that Steps 3 and 4 imply that the optimal policy belongs to the family characterized in Theorem 1. Using dynamic programming to solve the principal problem is difficult because it requires solving a non-linear partial differential equation. It is easier to analyze the problem using Pontryagin's maximum principle as in this case we only need to analyze incentives along the optimal path. However, even though the formal analysis relies on the maximum principle, next we provide a heuristic derivation of the main optimality conditions in Steps 1 and 2 using dynamic programming.[21] Let $U(x, q)$ be the value function of the principal, let $U_L = \max_{q \geq \underline{q}} U(0, q)$ and $U_H = \max_{q \geq \underline{q}} U(1, q)$ be the value just after the inspection, and let m and p be the monitoring rate and atom, respectively. **Step 1:**

[19] A function is singular if it is continuous, differentiable almost everywhere with a derivative of zero, and non-constant. The traditional example of a singular function is the Cantor function.

[20] The optimal policy in discrete time is qualitatively similar to the one in Theorem 1. The main difference occurs when the optimal policy in continuous time requires deterministic monitoring at time τ^{bind}. In discrete time, it might not be possible to make the IC constraint binding using deterministic monitoring, and this means that in some cases the support of the optimal policy can be concentrated in two consecutive periods. In continuous time, these two consecutive periods collapse to one, and all probability mass is concentrated in τ^{bind}. The analysis of the discrete time model can be found at `http://sites.duke.edu/fvaras/files/2017/11/discrete-time-appendix.pdf`.

[21] A formal analysis using dynamic programming would require to use the theory of constrained viscosity solutions (Fleming and Soner, 2006, Chapter II.12).

In the region without atoms, the value function satisfies the following HJB equation[22]

$$rU = u(x) + \lambda(\bar{a} - x)U_x + (r + \lambda)qU_q + \max_{m \geq 0} \left\{ m(\mathcal{M}(\mathbf{U}, x) - U - (1-q)U_q) \right\}. \tag{15}$$

Considering the first order condition and the complementary slackness constraint in the optimization problem in (15) we find that

$$m(\mathcal{M}(\mathbf{U}, x) - U - (1-q)U_q) = 0 \tag{16a}$$

$$\mathcal{M}(\mathbf{U}, x) - U - (1-q)U_q \leq 0. \tag{16b}$$

The presence of the incentive compatibility constraint introduces a wedge in the optimality conditions. In the absence of incentive constraints, the value function satisfies the value matching condition in Equation (4a), $U = \mathcal{M}(\mathbf{U}, x)$. Once we introduce incentive constraints, the value matching condition might not be satisfied. We show that $U_q(q_\tau, x_\tau) \leq 0$ through the optimal path, which means that the monitoring probability can be positive even though $\mathcal{M}(\mathbf{U}, x_\tau) < U(q_\tau, x_\tau)$. This captures the fact that monitoring is driven by incentive considerations in this case as the payoff if the principal monitors is lower than the continuation value if the principal does not monitor.

Step 2: Whenever there is an atom in the monitoring distribution, the value of q must jump. In particular, if q^+ is the value of q_τ just after the atom (conditionally on not monitoring), then it follows directly from Equation (13) that the monitoring probability is given by

$$p = \frac{q - q^+}{1 - q^+}.$$

The probability of monitoring is determined by q^+ so we can consider the maximization with respect to q^+ instead of p. Hence, the value function at the atom satisfies

$$U(q, x) = \max_{q^+ \in [\underline{q}, q]} \left(\frac{q - q^+}{1 - q^+} \right) \mathcal{M}(\mathbf{U}, x) + \left(\frac{1 - q}{1 - q^+} \right) U(q^+, x) \tag{17}$$

Looking at the first order conditions in (17), we find that the value of q^+ must satisfy

$$\mathcal{M}(\mathbf{U}, x) - U(q^+, x) - (1 - q^+)U_q(q^+, x) = 0, \ \ q^+ \in (\underline{q}, q) \tag{18a}$$

$$\mathcal{M}(\mathbf{U}, x) - U(q^+, x) - (1 - q^+)U_q(q^+, x) \leq 0, \ \ q^+ = \underline{q}. \tag{18b}$$

Unsurprisingly, the first order condition for the atom is analogous to Equation (16) (any atom can be approximated by an arbitrary large hazard rate). Using the envelope condition in equation (17) we get that

$$U_x(q, x) = \left(\frac{q - q^+}{1 - q^+} \right) (U_H - U_L) + \left(\frac{1 - q}{1 - q^+} \right) U_x(q^+, x). \tag{19}$$

[22] This equation holds for $q > \underline{q}$. For the purpose of this heuristic derivation, we assume that U is continuously differentiable so the equation holds at $\underline{q}$ when we consider the right limit of U_q.

Notice that Equation (19) is the analogous to the smooth pasting condition in the benchmark case (Equation (4b)). Combining Equations (15) and (16) just before and just after the atom (that is, it is satisfied at q and q^+) the value function satisfies

$$rU(q,x) = u(x) + \lambda(\bar{a} - x)U_x(q,x) + (r+\lambda)qU_q(q,x) \tag{20a}$$

$$rU(q^+,x) = u(x) + \lambda(\bar{a} - x)U_x(q^+,x) + (r+\lambda)q^+U_q(q^+,x). \tag{20b}$$

If we assume that U satisfies the smooth pasting condition $U_q(q,x) = U_q(q^+,x)$, then, by combining Equations (17), (19) and (20), we arrive to the following optimality condition that must be satisfied at any time τ_k with an atom:

$$r\mathcal{M}(\mathbf{U}, x_{\tau_k}) = u(x_{\tau_k}) + \lambda(\bar{a} - x_{\tau_k})(U_H - U_L) + (r+\lambda)U_q(q_{\tau_k}, x_{\tau_k}). \tag{21}$$

Equation (21) is the main optimality condition that we use to study the presence of atoms in the monitoring distribution. Once again, it is instructive to compare Equation (21) with the analogous condition in the benchmark without incentives constraints. In the benchmark, the HJB equation together with value matching and smooth pasting implies that at the monitoring boundary we have

$$r\mathcal{M}(\mathbf{U}, x_{\tau_k}) = u(x_{\tau_k}) + \lambda(\bar{a} - x_{\tau_k})(U_H - U_L).$$

The intuition is that the principal is indifferent between monitoring right now and delaying monitoring by $\mathrm{d}\tau$. The incentive constraint introduces a wedge (just as in the case of equations (16) and (18)). On the optimal path, this wedge is negative which means that at the time of the atom

$$r\mathcal{M}(\mathbf{U}, x_{\tau_k}) < u(x_{\tau_k}) + \lambda(\bar{a} - x_{\tau_k})(U_H - U_L),$$

which means that the principal would prefer to delay monitoring if it were not by the effect on incentive provision.

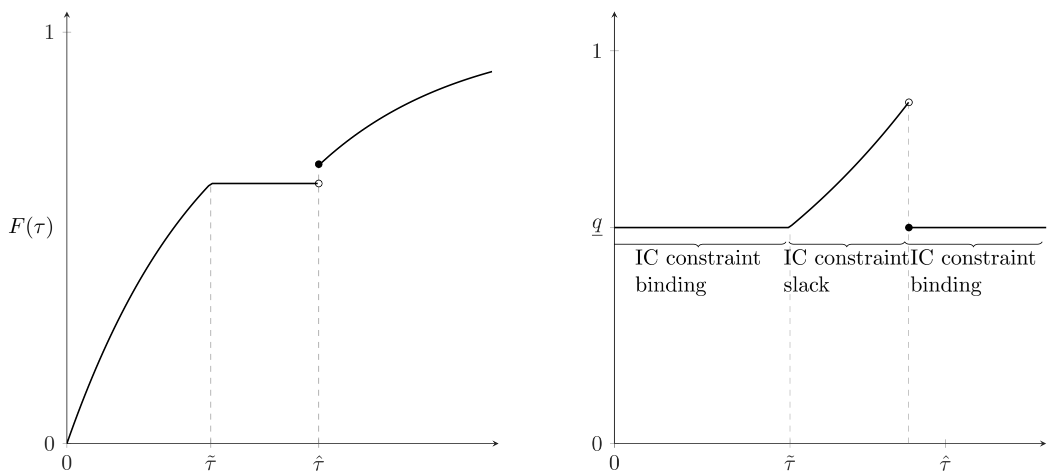
Step 3: If the incentive compatibility constraint is binding over a period of time, then q_τ is constant and the constant monitoring rate is determined by the condition that $q_\tau = \underline{q}$. On the other hand, if the incentive compatibility constraint is slack and $m_\tau > 0$, then the necessary condition (16a) requires that $\mathcal{M}(\mathbf{U}, x_\tau) - U(q_\tau, x_\tau) - (1-q_\tau)U_q(q_\tau, x_\tau) = 0$. We show that this condition cannot hold in a time interval over which and the incentive compatibility constraint is slack due to the convexity of $u(x)$, which means that the monitoring rate must be zero in this case. Formally, we show that

Lemma 3. *Let M^*_τ be an optimal policy, M^{c*}_τ its continuous part, and $B = \{\tau \in [0, \bar{\tau}^*] : q_\tau > \underline{q}\}$ the set of dates at which the IC constraint is slack. Then, $\int_B \mathrm{d}M^{c*}_\tau = 0$.*

Step 4: Once again, using the the convexity of $u(x)$, we show that Equation (21) satisfies a single crossing condition that implies that (21) holds at most at one point on the optimal path of q_τ, so the optimal policy can have at most one atom. Formally, we show that

Lemma 4. *There is at most one time $\hat{\tau}$ such that $\Delta M^d_{\hat{\tau}} > 0$.*

Step 5: The final step is to verify that Lemmas 3 and 4 imply that any optimal policy must take the form in Theorem 1. Figure 2 illustrates the form of the policies consistent with Lemmas 3 and 4. 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 either (1) reaches one or (2) there is an atom and q_τ jumps down to $\underline{q}$, and the incentive compatibility constraint is binding thereafter. Figure 2a shows the monitoring policy associated with the trajectory of q_τ. Before time $\tilde{\tau}$, 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 $\tilde{\tau}$, there is no monitoring and the incentive compatibility constraint is slack. At time $\hat{\tau}$, either there is monitoring with probability 1, so $q_{\hat{\tau}} = 1$ and $\hat{\tau} = \bar{\tau}$, or there is an interior atom so conditional on not monitoring, the monitoring distribution is exponential thereafter. We show in the the proof sthat $\tilde{\tau}$ is either zero or infinity, which allow us to conclude that the optimal policy must take the form in Theorem 1.



(a) Example of CDF.

(b) Path of q_τ conditional on not monitoring.

Figure 2: Cumulative density function and path of q_τ implied by Lemmas 3 and 4.

Theorem 1 allows to write the principal's problem as a one dimensional problem in which we choose the date of the atom in the monitoring distribution. Let $\mathcal{G}^\theta_{\text{det}}$ be the best incentive compatible deterministic policy given continuation payoffs $\mathbf{U}$:

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

and let $\mathcal{G}^{\theta}_{\text{rand}}$ be the payoff of best random policy, as given by:

$$\mathcal{G}^{\theta}_{\text{rand}}(\mathbf{U}) \equiv \max_{\hat{\tau}\in[0,\tau^{\text{bind}}]} \int_0^{\hat{\tau}} e^{-r\tau} u(x^{\theta}_{\tau})\mathrm{d}\tau + e^{-r\hat{\tau}} \Bigg[\left(\frac{e^{(r+\lambda)\hat{\tau}} - 1}{1-\underline{q}} \right) \underline{q}\mathcal{M}(\mathbf{U}, x^{\theta}_{\hat{\tau}}) + \left(\frac{1 - e^{(r+\lambda)\hat{\tau}}\underline{q}}{1-\underline{q}} \right) \int_{\hat{\tau}}^{\infty} e^{-(r+m)(\tau-\hat{\tau})} \left(u(x^{\theta}_{\tau}) + m\mathcal{M}(\mathbf{U}, x^{\theta}_{\tau}) \right) \mathrm{d}\tau \Bigg]$$

The optimal random policy $\mathcal{G}^{\theta}_{\text{rand}}(\mathbf{U})$ is fully described by the monitoring rate m^* and the length of the quiet period, captured by $\hat{\tau}^*_{\theta}$, which, by Theorem 1, pins down the size of the atom that initializes the random monitoring phase. The solution to the principal's problem is thus given by the fixed point:

$$U_{\theta} = \max\{\mathcal{G}^{\theta}_{\text{det}}(\mathbf{U}), \mathcal{G}^{\theta}_{\text{rand}}(\mathbf{U})\},\ \theta \in \{L, H\}. \tag{22}$$

To build some intuition about the form of the optimal policy, it is instructive to look at the first order conditions for the time of the atom in $\mathcal{G}^{\theta}_{\text{rand}}(\mathbf{U})$. Let $V(\tau)$ be the (ex-post) payoff of monitoring at time τ, which is given by

$$V(\tau) = \int_0^{\tau} e^{-rs} u(x^{\theta}_s)\mathrm{d}s + e^{-r\tau}\mathcal{M}(\mathbf{U}, x^{\theta}_{\tau})$$

It can be verified that the first order condition for the maximization problem in $\mathcal{G}^{\theta}_{\text{rand}}(\mathbf{U})$ can be written as

$$V'(\hat{\tau}) = \frac{r+\lambda}{1-\underline{q}} \left(E[V(\tau)|\tau > \hat{\tau}] - V(\hat{\tau}) \right) \tag{23}$$

The first order condition in equation (23) highlights the trade-off between deterministic and random monitoring. On the one hand, the marginal benefit of delaying monitoring at time τ is given by $V'(\tau)$. Because preferences are convex and beliefs move faster at the beginning, $V'(\tau)$ is the highest just after the last inspection, implying that the incentives to delay monitoring are the highest at this time. On the other hand, incentive compatibility requires that any delay in monitoring is compensated by a higher probability of monitoring (bigger atom). This effect is captured by the second term in (23): the marginal effect of $\hat{\tau}$ on the size of the atom is given by $(r+\lambda)/(1-\underline{q})$ while the net benefit/cost of using random monitoring after time $\hat{\tau}$ is given by $E[V(\tau)|\tau > \hat{\tau}] - V(\tau)$. Figure 3 shows that the first effect dominates at the beginning so the optimal policy specifies no monitoring at this time.

Having characterized the structure of the optimal policy, we can discuss the conditions under which random monitoring dominates deterministic monitoring. The next proposition considers how parameters affect the form of the optimal policy.

Proposition 3 (Comparative Statics)**.** *Suppose that $u(x)$ is strictly convex, then:*

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.*

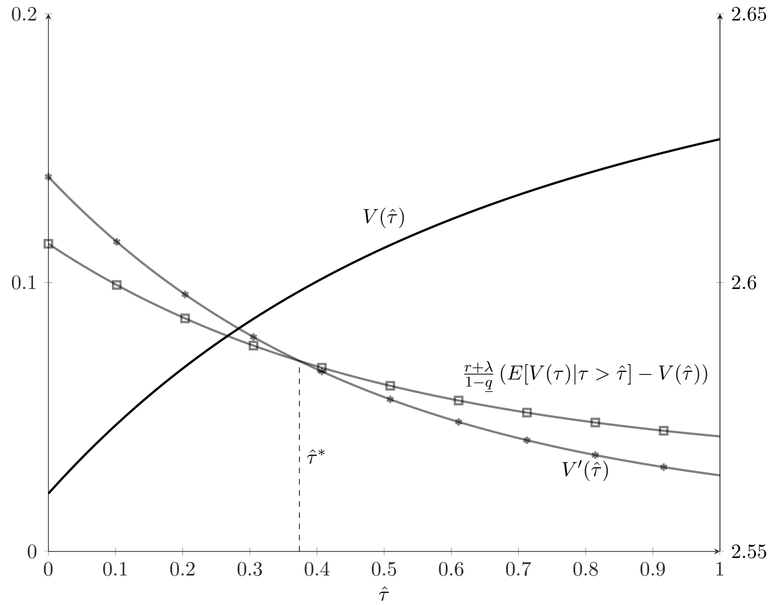
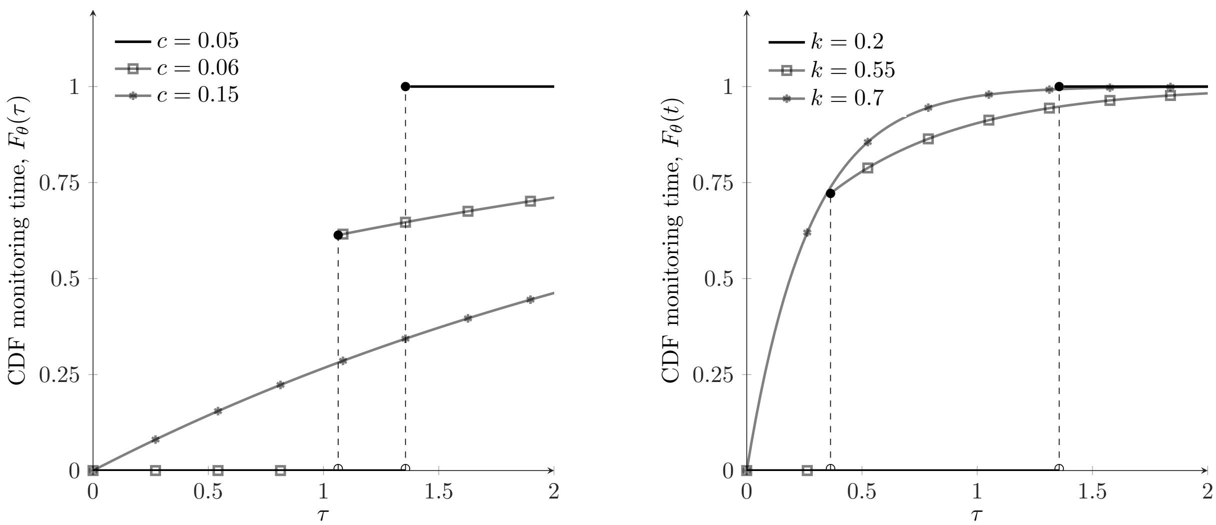


Figure 3: First order condition and ex-post payoff $V(\tau)$. In this picture, the right axis corresponds to the function $V(\tau)$ while the left axis corresponds to the functions $V'(\hat{\tau})$ and $\frac{r+\lambda}{1-\underline{q}}\left(E[V(\tau)|\tau > \hat{\tau}] - V(\hat{\tau})\right)$.

2. *There is $k^\dagger < \lambda/(r+\lambda)$ such that for any $k > k^\dagger$ the optimal policy has random monitoring.*

3. *There is $\bar{a}^\dagger < 1$ such that, for any $\bar{a} \in (\bar{a}^\dagger, 1)$, the optimal policy given $\theta_{T_{n-1}} = H$ has random monitoring. Similarly, there is $\bar{a}_\dagger > 0$ such that, for any $\bar{a} \in (0, \bar{a}_\dagger)$, the optimal policy given $\theta_{T_{n-1}} = L$ has random monitoring.*

Figure 4a 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 the benchmark policy (the relaxed problem without incentive constraint) prescribes frequent monitoring, and accordingly the incentive compatibility constraint is slack. When the cost is at an intermediate level, the optimal policy is a mixture of deterministic and random monitoring with a constant hazard rate, while when the cost of monitoring is high, the optimal policy specifies constant random monitoring starting at time zero. Similarly, Figure 4b show the comparative statics for the cost of effort, k. The monitoring policy is random if k is high enough, deterministic if k is low, and a mixture of both when k is in between. We provide a more detailed analysis of the comparative statics in Section 7 in the context of a model with linear quadratic preferences and quality driven by Brownian motion.



(a) Comparative statics for c. The cost of effort is $k = 0.2$.

(b) Comparative statics for k. The cost of monitoring is $c = 0.05$.

Figure 4: Comparative statics for the 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: When the cost of monitoring is very high the monitoring policy consist on random monitoring at all times and at a constant rate; on the other hand, if the cost of monitoring is an intermediate range, the optimal monitoring policy entails a first period without monitoring followed by an atom and constant random monitoring thereafter. In this example the payoff function and the technology are symmetric so the optimal monitoring policy is independent of θ_0.

7 Quality Driven by Brownian Motion

Our baseline model assumes that quality can take on two values. Such binary specification makes the analysis tractable but is not strictly needed: the economics of the problem is not driven by the details of the quality process. As mentioned in Remark 4, the policy in the linear case remains optimal for a general class of quality process. In this section, we analyze the optimal policy when information is valuable and quality follows the Ornstein-Uhlenbeck process

$$d\theta_t = \lambda(a_t - \theta_t)dt + \sigma dB_t, \tag{24}$$

where B_t is a Brownian motion.

Whenever the principal's payoff is not linear in quality one needs to specify the principal's preferences as a function of the firm reputation. With non-linear preferences, the optimal policy generally depends on the last inspection's outcome (which in this case has a continuum of outcomes). While this fact does not seem to change the core economic forces, it makes the analysis and computations more involved so we do not have a general characterization of the optimal policy for the convex case. However, we can get a clean characterization of the optimal policy when the principal's preferences are linear-quadratic. The linear-quadratic case is common in applications of costly information acquisition for its tractability (Jovanovic and Rousseau, 2001; Sims, 2003; Hellwig and Veldkamp, 2009; Alvarez et al., 2011; Amador and Weill, 2012).

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 $\Sigma_t \equiv \text{Var}(\theta_t|\mathcal{F}_t^M)$. For example, this preference specification corresponds to the case in which the evolution of quality is driven by Brownian motion in Example 3 in Section 2.1.

For the Ornstein-Uhlenbeck process in (24), the distribution of θ_t is Gaussian with moments

$$x_t = \theta_0 e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right) \tag{25}$$

$$\Sigma_t = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right). \tag{26}$$

Using the law of iterated expectations, we see that the principal's continuation payoff at the time of an inspection 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

$$\begin{cases} \mathcal{C} = \min_{M_\tau} \int_0^\infty e^{-r\tau - M_{\tau-}} \left(\gamma \Sigma_\tau \mathrm{d}\tau + (c+\mathcal{C})\, \mathrm{d}M_\tau^c\right) + \sum e^{-r\tau - M_{\tau-}} (1 - e^{-\Delta M_\tau^d})\,(c+\mathcal{C}) \\ \text{subject to} \\ \mathrm{d}q_\tau = (r+\lambda) q_\tau \mathrm{d}t - (1-q_\tau)\mathrm{d}M_\tau^c, \;\; q_{\bar{\tau}} = 1 \\ q_{\tau-} = e^{-\Delta M_\tau^d} q_\tau + (1 - e^{-\Delta M_\tau^d}) \\ q_\tau \in [\underline{q}, 1] \end{cases}$$

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^c, and the atoms corresponding to a discrete probability of an inspection, M_τ^d. The main state variable is again the expected discount factor until the next review, q_τ.

Given the symmetry in the linear-quadratic case, the optimal policy is independent of the outcome in the previous inspection, and using the previous results from the binary case, we can show that the optimal monitoring policy takes the same form as in the binary case. This means that the optimal monitoring policy and the cost of monitoring is given by

$$\mathcal{C} = \min \left\{ \min_{\bar{\tau} \in [0, \tau^{\text{bind}}]} \frac{\int_0^{\bar{\tau}} e^{-r\tau} \gamma \Sigma_\tau \mathrm{d}\tau + e^{-r\bar{\tau}} c}{1 - e^{-r\bar{\tau}}}, \right. \\ \left. \min_{\hat{\tau} \in [0, \tau^{\text{bind}}]} \frac{\int_0^{\hat{\tau}} e^{-r\tau} \gamma \Sigma_\tau \mathrm{d}\tau + e^{-r\hat{\tau}} \left(\frac{1 - e^{(r+\lambda)\hat{\tau}} \underline{q}}{1 - \underline{q}} \right) \int_{\hat{\tau}}^\infty e^{-(r+m^*)(\tau - \hat{\tau})} \gamma \Sigma_\tau \mathrm{d}\tau + \delta(\hat{\tau}) c}{1 - \delta(\hat{\tau})} \right\}, \quad (27)$$

where

$$\delta(\hat{\tau}) \equiv \left(\frac{e^{\lambda \hat{\tau}} - e^{-r\hat{\tau}}}{1 - \underline{q}} \right) \underline{q} + \left(\frac{e^{-r\hat{\tau}} - e^{\lambda \hat{\tau}} \underline{q}}{1 - \underline{q}} \right) \frac{m^*}{r + m^*}$$

and the optimal monitoring policy is given by:

Proposition 4. *Suppose that θ_t follows the Ornstein-Uhlenbeck process in (24), 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 given by the distribution*

$$F^*(\tau) = \begin{cases} 0 & \text{if } \tau \in [0, \hat{\tau}^*) \\ 1 - p^* e^{-m^*(\tau - \hat{\tau}^*)} & \text{if } \tau \in [\hat{\tau}^*, \infty] \end{cases}$$

where

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

and $\hat{\tau}^* \leq \tau^{bind}$. *If* $p^* > 0$, *then it is given by*

$$p^* = \frac{1 - e^{(r+\lambda)\hat{\tau}^*}\underline{q}}{1 - \underline{q}}.$$

As before the distribution of monitoring is characterized by two numbers, the size of the atom p^* and the monitoring rate m^*. As a special cases, the policy prescribes deterministic monitoring when $p^*_\theta = 0$, and purely random monitoring with constant rate m^* when $p^*_\theta = 1$.

The comparative statics in the case of Brownian shocks are similar to those in Proposition 3: The optimal policy is deterministic if the cost of monitoring is low and random if the cost of monitoring is high. There are two new parameters in the model, γ and σ: However, after inspecting equations (26) and (27) we see that the monitoring policy only depends on the cost of monitoring per unit or risk, $c/\gamma\sigma^2$, so increasing γ/σ is equivalent to reducing the cost of monitoring. We have the following proposition characterizing the comparative statics in the linear quadratic case.

Proposition 5 (Comparative Statics)**.** *Suppose that* θ_t *follows the Ornstein-Uhlenbeck process in* (24)*, and that the principal's expected payoff flow is* $u(x_t, \Sigma_t) = x_t - \gamma\Sigma_t$. *If we let* $\tilde{c} \equiv c/\gamma\sigma^2$ *then*

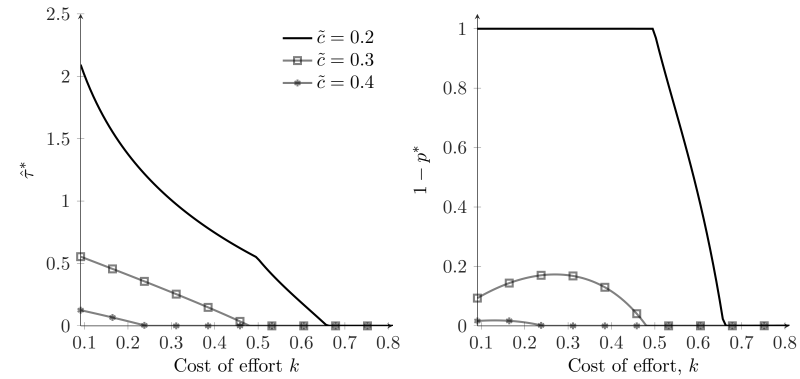
1. *There is* $\tilde{c}^\dagger > 0$ *such that the optimal policy is deterministic if* $\tilde{c} \leq \tilde{c}^\dagger$ *and random if* $\tilde{c} > \tilde{c}^\dagger$.

2. $\hat{\tau}^*$ *is increasing in* $\tilde{c}$ *for* $\tilde{c} \leq \tilde{c}^\dagger$ *and decreasing for* $\tilde{c} > \tilde{c}^\dagger$. *This means that the atom* p^* *is increasing in* $\tilde{c}$ *so the probability of monitoring at* $\hat{\tau}^*$ *is decreasing in* $\tilde{c}$.

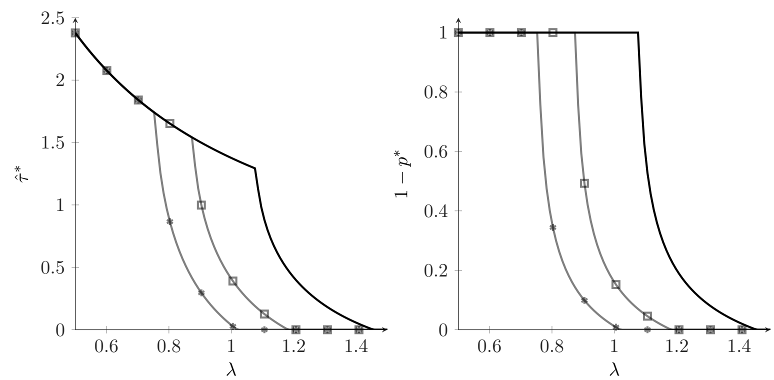
3. *If* $\tilde{c} \leq \frac{1}{2\lambda(r+2\lambda)}$ *then there is* $k^\dagger > 0$ *such that the optimal policy is deterministic if* $k \leq k^\dagger$ *and random if* $k > k^\dagger$. *For* $k > k^\dagger$, $\hat{\tau}^*$ *is decreasing in* k.

4. *Consider the i.i.d limit when* $\lambda_n \to \infty$, $\sigma_n = \sigma\sqrt{\lambda_n}$, $\tilde{c}_n = c\lambda_n/\gamma\sigma_n^2$. *In this limit, the optimal monitoring policy is random.*

Consistent with the notion that the principal faces two types of costs –the cost of inspections, captured by c, and the cost of uncertainty, captured by $\gamma\sigma^2$– the structure of the optimal policy (i.e., deterministic vs random) depends on the cost of inspection per unit of uncertainty, or $c/\gamma\sigma^2$. Intuitively, a low $\tilde{c}$ captures the case when the principal has little tolerance to uncertainty, characterized by frequent inspections and the absence of moral hazard issues (the incentive constraint is slack). By contrast, the high $\tilde{c}$ captures the case when inspections are too costly relative to the cost of uncertainty, leading to rather infrequent inspections and random monitoring. Finally, the result that the optimal policy is random in i.i.d limit, where quality shock are highly transitory, shows how the possibility of window dressing moves the optimal policy towards random monitoring.



(a) Comparative statics for k. Parameters values are $r = 0.1$, $\bar{a} = 0.5$ and $\lambda = 1$.



(b) Comparative statics for λ. Parameters values are $r = 0.1$, $\bar{a} = 0.5$ and $k = 0.2$.

Figure 5: Comparative statics linear quadratic model with Brownian shocks. In the top panel, the left figure shows the date of atom $\hat{\tau}^*$ while the right figure shows the probability of monitoring at $\hat{\tau}^*$, which is given by $1 - p^*$. That is, $1 - p^* = 1$ implies a deterministic monitoring date while $1 - p^* = 0$ implies random monitoring with constant rate m^* starting at time zero. In the bottom panel, the left figure shows the date of monitoring in the benchmark without incentive constraints, the center figure shows $\hat{\tau}^*$, while the right figure shows $1 - p^*$.

8 Exogenous News

Thus far we have ignored alternative sources of information, beside monitoring. In this section we explore the effect of having exogenous news on the optimal monitoring policy. We show that exogenous news, not only crowd-out monitoring, but by altering the severity of the moral hazard issue across states may modify the monitoring policy in a significant way.

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 efforts. To provide some insights about the interaction between monitoring and news, we consider the presence of an exogenous news process that may reveal current quality to the market. More specifically, we consider the case in which the quality of product is revealed to the market at a Poisson arrival rate.

Assume there are two Poisson processes $(N_t^L)_{t\geq 0}$ and $(N_t^H)_{t\geq 0}$. The process N_t^L is a bad news process with mean arrival rate $\theta_t = \mu_L \mathbf{1}_{\{\theta_t = L\}}$, and N_t^H is a good news process with mean arrival rate $\mu_H \mathbf{1}_{\{\theta_t = H\}}$. When $\mu_L \neq \mu_H$ we say that news are *asymmetric*, in which case, the absence of news is informative about the firm quality. On the other hand, if $\mu_L = \mu_H$ the lack of news arrival is uninformative. 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

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

The second term cancels if $\mu_H = \mu_L$ and the dynamics of beliefs (in the absence of any monitoring by the principal and arrival of exogenous news) is the same as in the case without news. On the other hand, if $\mu_L \neq \mu_H$, 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. The market learns from the absence of news since no news are informative when the news processes have asymmetric arrival rates.

Let's first consider the case with symmetric news arrival, i.e. $\mu_L = \mu_H = \mu$. From the firms's point of view, it does not matter if the state is learned due to monitoring or exogenous news. The only difference is that now, there is an extra arrival rate that reveal the state. If we denote the date at which quality is revealed, either by monitoring or exogenous news, by $\tilde{T}_n$, then we can still write the incentive compatibility constraint as

$$E\left[e^{-(r+\lambda)(\tilde{T}_n - t)} \middle| \mathcal{F}_t\right] \geq \underline{q}.$$

This means that we can still use q_τ as our main state variable, and the dynamics of q_τ are given by

$$\mathrm{d}q_\tau = (r+\lambda)q_\tau \mathrm{d}\tau - (1 - q_\tau)(\mathrm{d}M_\tau^c + \mu \mathrm{d}t). \tag{28}$$

Notice that the only difference between equations (10) and (28) is that the monitoring rate $\mathrm{d}M_\tau^c$ is incremented by $\mu \mathrm{d}\tau$ due to the exogenous news. Similarly, because the problem of the principal is the same going forward no matter if quality was learnt due to monitoring or exogenous news, we can still write the problem recursively based on 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 principal's continuation value now evolves according to

$$\mathrm{d}U_\tau = \left(rU_\tau - u(x^\theta_\tau)\right)\mathrm{d}\tau + \left(U_\tau - \mathcal{M}(\mathbf{U}, x^\theta_\tau)\right)\mathrm{d}M^c_\tau + \mu\left(U_\tau - x^\theta_\tau U_H - (1 - x^\theta_\tau)U_L\right)\mathrm{d}\tau. \tag{29}$$

Hence, the Principal's problem has the same structure as before with the exception that now the principal gets some monitoring with intensity μ *for free*. When news are symmetric, exogenous news are a perfect substitute of monitoring. Lemmas 3 and 5 still apply, and the monitoring rate is positive only if the incentive compatibility constraint is binding, in which case $\mathrm{d}q_\tau = 0$ so the monitoring rate is

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

Clearly, the monitoring rate to keep the incentive constraint binding needs to be positive only if μ is low enough. Otherwise, exogenous news suffices for incentive purposes. In this latter case, exogenous news are enough to discipline the firm, and the only purpose of monitoring is to learn the sate. 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 1.

Proposition 6. *Suppose that $\mu_L = \mu_H$. If $(r + \lambda)\frac{\underline{q}}{1-\underline{q}} \geq \mu$ then the optimal monitoring policy is the one characterized in Propositions 2 and 1 with a Poisson monitoring rate.*

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

On the other hand, if $(r + \lambda)\frac{\underline{q}}{1-\underline{q}} < \mu$, then the optimal monitoring policy is deterministic.

Proof. Letting $\tilde{M}^c_\tau = M^c_\tau + \mu\tau$ and $\tilde{u}(x) = u(x) + \mu c$, we can write

$$\begin{aligned}
\mathrm{d}q_\tau &= (r + \lambda)q_\tau \mathrm{d}\tau - (1 - q_\tau)\mathrm{d}\tilde{M}^c_\tau \\
\mathrm{d}U_\tau &= \left(rU_\tau - \tilde{u}(x^\theta_\tau)\right)\mathrm{d}\tau + \left(U_\tau - \mathcal{M}(\mathbf{U}, x^\theta_\tau)\right)\mathrm{d}\tilde{M}^c_\tau,
\end{aligned}$$

so the optimal control problem follows the same structure as before with two diferences: (1) now $\mathrm{d}\tilde{M}^c_\tau$ must be greater or equal than $\mu\mathrm{d}\tau$, and (2) q_τ is bounded below by $\frac{\mu}{r+\lambda+\mu}$. If $(r + \lambda)\frac{\underline{q}}{1-\underline{q}} \geq \mu$ then (1) and (2) are not binding. On the other hand, if $(r + \lambda)\frac{\underline{q}}{1-\underline{q}} < \mu$ then $q_\tau > \underline{q}$. Hence, the incentive compatibility constraint is slack at all times, so the solution to the Principal problem corresponds to the one in Section 3, which means that monitoring is deterministic. □

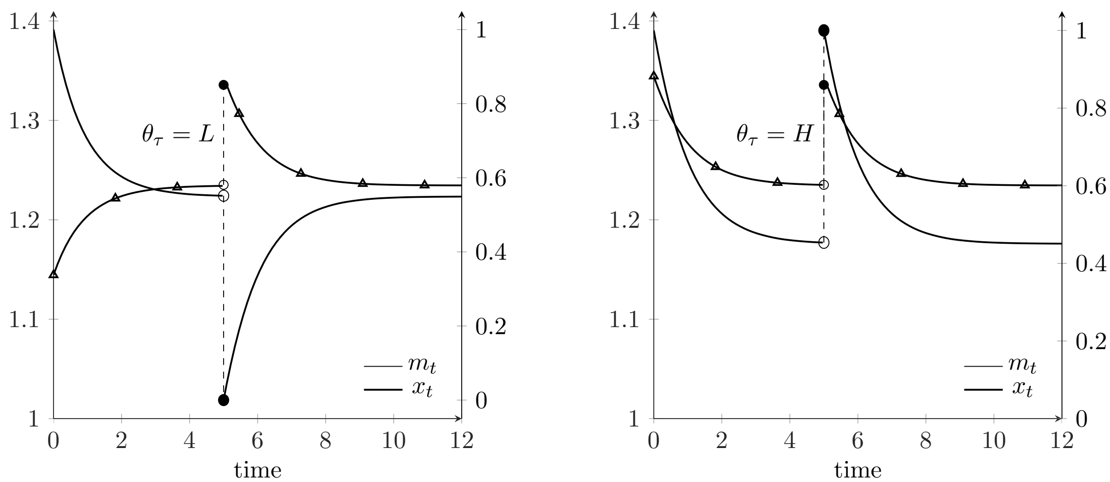
8.1 Asymmetric News Intensity

The qualitative results are different if $\mu_H \neq \mu_L$. In this case, the presence of news changes the dynamics of incentives: the monitoring rate changes over time and is dependent of the outcome of the outcome in the last review. We do not solve the full problem here, and instead focus on the case in which the principal's preferences are linear. Based on our previous analysis of the linear, it is natural to conjecture that the optimal policy has (1) no atoms and that (2) the monitoring rate is positive only if the incentive compatibility constraint is binding. We can use the maximum principle to verify if our conjectured policy is optimal. We relegate a detailed discussion of the solution to the appendix.

We focus on the simplest case with parameters such that the optimal policy has $m_\tau > 0$ for all $\tau \geq 0$; this case illustrates the effect of introducing exogenous news on the optimal monitoring policy at the lowest cost of technical complications.[23]

The dynamics of optimal monitoring are described in Figure 6. In the bad news case, monitoring increases after (bad) news. The opposite is optimal in the good news case. The dynamics of monitoring are 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, inspections are most needed 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, inspections are thus most needed when reputation is high. Accordingly monitoring complements exogenous news, being used when exogenous news are ineffective at providing incentives.

[23]Such policy is optimal when the rates of exogenous news arrivals are low. When those rates 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.



(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 $\mu_L = 0$

9 Final Remarks

In our model, a firm exerts hidden effort to affect a persistent quality process and has reputation concerns. The monitoring policy plays 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 the optimal policy is very simple: depending on the result of the last inspection, it either prescribes random inspections with a constant hazard rate, or a deterministic review, or a mixture between a deterministic review and random monitoring with constant hazard rate and fixed delay. Since in practice, monitoring is often triggered by the revelation of some news, we also provide some results for the interaction of monitoring and exogenous news.

We conclude with a discussion of possible extensions and caveats to our analysis that could capture some aspects of the design of a monitoring system that we have not incorporated in our analysis.

Fines, bonuses and transfers. We have considered settings without monetary transfers where incentives are driven purely by reputation concerns. With arbitrary fines, the problem becomes trivial since a very small intensity of monitoring combined with a large fine would bring us to first-best. A more realistic setting entails limited fines. Based on our analysis, we conjecture that limited fines would push the optimal policy away from random monitoring towards deterministic reviews, because the solution to the relaxed problem (that ignores incentive constraints) would be more likely to satisfy incentive constraints when low-quality firms pay fines.[24]

Full Effort We focus on policies that induce full effort, but for some parameters we conjecture that an optimal policy would prescribe no effort at all, after some bad histories. That could be optimal for the principal (even if effort is optimal in the first-best), especially if conditional on full effort the probability of maintaining superior quality is very high. The intuition comes from the case $\bar{a} = 1$, where once the firm achieves high quality it can maintain it forever. An inspection revealing low quality would then be off-the-equilibrium path and, to relax incentive constraints, it would be optimal to use the worst possible punishment for the firm after that outcome. This can be implemented by stopping monitoring altogether thereafter, leading to no effort. This is akin to revoking the firm's license. If $\bar{a}$ is very close to 1, we conjuncture that using such strong punishments with some probability may remain optimal. However, we also conjecture that if the cost of inspections is not too large and $\bar{a}$ is sufficiently away from 1, the optimal policy would indeed induce full effort.

Another reason to focus on full effort is that in many applications there are institutional constraints that may make punishments via no future effort unfeasible. For example, in the case of

[24]It is not immediately obvious to us what is a satisfactory model of limited fines. For example, if we only bound the fee charged per inspection, then upon finding the firm to be low quality, the regulator could perform many additional inspections in a short time interval and fine the firm multiple times. A similar issue arises if the firm incurs part of the physical cost of inspection: running additional inspections could expose the firm effectively to a large fine.

public schools, neighbors would probably not allow a policy that implements perpetual low quality if their local school has failed in the past. In this case, a policy that looks for high effort after any history might be the only thing that is politically feasible to implement.

Asymmetric Switching Intensities In the model, we have assumed that the switching intensity from the low to the high state is λa and the switching intensity from the high to the low state is $\lambda(1-a)$. Having the λ's symmetric turns out to simplify the IC constraints because the marginal return to effort is the same in the high and low state. In consequence, the set of incentive-compatible policies is the same whether the agent observes or not the current level of quality, and they are the same on and off the equilibrium path. If λ's in the two states are asymmetric and the agent observes current quality, then our analysis can be applied with almost no changes, as long as we want to maintain full effort in both states. We just need to use the lower of the two λ (since the marginal return to effort is increasing in λ and both levels of quality are on the equilibrium path at almost all times).

In the case where the agent does not observe the current state, the analysis is potentially more complicated because the agent's beliefs will diverge from the principal's if the agent deviates from the recommended effort. A policy that assures that the agent has incentives to put full effort in both states at all times would still be incentive compatible but not necessarily optimal. The intuition is that with different λ's the agent's incentive constraints would change with his beliefs about the state and the optimal policy could economize on inspection costs because of that. Our intuition about the optimal policy, in that case, is as follows. After a deviation to lower effort, the agent assigns a lower probability to the high state than the principal. If λ in the low state is higher than in the high state, the IC constraints become slack off-path. In that case, the optimal policy can be characterized using our current methods (for example, when principal's payoff is linear, the optimal policy would have random inspections with an intensity that is high and decreasing after a high-quality result and vice versa in case of low quality). If λ in the low state were lower than in the high state, then the analysis would get somewhat more complicated because, if the IC constraint is just binding on the equilibrium path, then it would be violated after a deviation. Preventing such "double deviations" may require a higher intensity of monitoring after a good inspection outcome (and IC constraints that are a bit slack on-path).

Multiple Dimensions In many applications quality is a multi-dimensional attribute and the principal inspects only a few dimensions at a time. In other words, the principal decides when and what to inspect. For example, when the Federal Aviation Administration inspects planes, it is impractical to test all components at every inspection. The FAA sometimes announces a focus on a particular part (like a hatch door) and sometimes not. As a result, even if the timing of inspections is deterministic (for example, every 400-600 flight hours), when the airline does not know which parts will be inspected, the timing of inspections is random from the perspective of each part. Similar considerations apply to accounting audits (where it often is uncertain which aspects of the firms financial statements and its internal control system will be scrutinized by the auditor) or FDA inspections of drug manufacturers (if there is uncertainty over which elements of the production process will be inspected).

Communication In our model we did not allow the firm to communicate with the principal. In some certification systems, certified firms are supposed to self-report any problems and communicate their resolution.[25] Safe-harbor provisions often protect and promote self-reporting. In our model the bad news process analyzed in Section 8 can be interpreted as capturing the self-reporting of problems detected by the firm, at least in settings where reporting is compulsory. Analysis of voluntary self-reporting and the effects of safe-harbor provisions (such as those included the Private Securities Litigation Act of 1995, which shelters managers disclosing forward looking information from litigation arising from unattained projections) within our model could provide additional insights about the effects of those provisions.

Another form of communication in some markets is that a firm that fails an inspection needs to request a new inspection after it solves its problems (for example, in case of hygiene inspections of restaurants). Such self-reporting could improve the performance of the optimal policy we identified in this paper by avoiding unnecessary inspections. To see this, suppose that the optimal policy is random monitoring. If we allowed firms who failed the last inspection to self-report improvements, the reputation would remain constant at 0 until the firm requested re-certification. That would improve the principals payoff because of learning and the possibly of lower certification costs (although the second effect is ambiguous). While we do not provide a characterization of the optimal policy with self-reporting, we expect the tradeoffs between random and deterministic inspections we stress in this paper would remain relevant in such a model, while new insights are likely to emerge (for example, a characterization of when immediate re-certification upon request is optimal and when it is not).

Non-linearitiy of Firm Payoffs In some markets the firm's payoffs are non-linear in the firm's reputation. For example, some consumers may be willing to buy only from firms with reputation levels above a certain threshold, making the firm payoffs convex in reputation. In these cases, monitoring could have an additional effect of providing direct value to the firm. Based on our analysis, we conjecture that this possibility would push the optimal monitoring policy towards deterministic reviews, which could explain why many Self-regulatory-organizations controlled by firms in the industry provide certification on a deterministic schedule (as in the Doctor Board Certification program). Analyzing the optimal policy in this case would be more difficult than in our model because when information has direct value to the firm, inspections provide additional incentives and a future increase or decrease in the inspection frequency could be used by the regulator to reward or punish the firm. The same complications and tradeoffs would appear if inspections were costly to the firms. In this case, we conjecture that the regulator would find

[25] For example the National Association for the Education of Young Children requires accredited child care centers to notify NAEYC within 72 hours of any critical incident that may impact program quality http://www.naeyc.org/academy/update accessed 2/28/2017.

it optimal to reward firms by less frequent future inspections and punish them by more frequent ones. That would increase the value of good reputation and hence relax moral hazard constraints, presumably making deterministic monitoring more attractive when information has a direct value.

References

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

Alvarez, F. E., F. Lippi, and L. Paciello (2011). Optimal price setting with observation and menu costs. *Quarterly Journal of Economics 126*(4), 1909–1960.

Amador, M. and P.-O. Weill (2012). Learning from private and public observations of othersê actions. *Journal of Economic Theory 147*(3), 910–940.

Arutyunov, A., D. Karamzin, and F. Pereira (2005). A nondegenerate maximum principle for the impulse control problem with state constraints. *SIAM Journal on Control and Optimization 43*(5), 1812–1843.

Bassetto, M. and C. Phelan (2008). Tax riots. *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.

Bhaskar, V. and C. Thomas (2017). The design of credit information systems. Working Paper, UT Austin.

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

Board, S. and M. Meyer-ter Vehn (2014). A reputational theory of firm dynamics. Working Paper, UCLA.

Bogachev, V. I. (2007). *Measure Theory*, Volume 1. Springer.

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

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. *Review of Economic Studies 66*(1), 199–217.

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. *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.

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. *American Economic Review 94*(3), 591–604.

Fleming, W. H. and H. M. Soner (2006). *Controlled Markov Processes and Viscosity Solutions*, Volume 25. Springer Science & Business Media.

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. *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. *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.

Hellwig, C. and L. Veldkamp (2009). Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies 76*(1), 223–251.

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. *Review of Economic Studies 66*(1), 169–182.

Horner, J. and N. S. Lambert (2016). Motivational ratings. Working Paper, Yale University.

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

Jovanovic, B. and P. L. Rousseau (2001). Why wait? a century of life before IPO. *American Economic Review 91*(2), 336–341.

Kim, S.-H. (2015). Time to come clean? disclosure and inspection policies for green production. *Operations Research 63*(1), 1–20.

Kovbasyuk, S. and G. Spagnolo (2016). Memory and markets. Available at SSRN.

Lansdowne, Z. F. (1970). The theory and applications of generalized linear control processes. Technical report, DTIC Document.

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

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

Marinovic, I., A. Skrzypacz, and F. Varas (2018). Dynamic certification and reputation for quality. *American Economic Journal: Microeconomics Forthcoming*.

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

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

OECD (2014). Regulatory enforcement and inspections, OECD best practice and principles for regulatory policy. Technical report, OECD.

Pereira, F. L. and G. N. Silva (2011). A maximum principle for constrained infinite horizon dynamic control systems. *IFAC Proceedings Volumes 44*(1), 10207–10212.

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

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

Rishel, R. W. (1965). An extended pontryagin principle for control systems whose control laws contain measures. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control 3*(2), 191–205.

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

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

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

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)*.

Appendix (Not for Publication)

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

$$(r+\lambda)U'(x) = u'(x) + \lambda(\bar{a} - x)U''(x) \tag{30a}$$

$$(r+2\lambda)U''(x) = u''(x) + \lambda(\bar{a} - x)U'''(x) \tag{30b}$$

Using (30b) 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 (30a) 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

$$\frac{u'(\underline{x})}{r+\lambda} \leq U(1) - U(0) \leq \frac{u'(\overline{x})}{r+\lambda} \tag{31}$$

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 (31) is satisfied. The fact that U is increasing follows directly from the convexity of U and equation (30a).

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

$$xU(1) + (1-x)U(x) - U(x) \leq c. \tag{32}$$

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,

$$0 \geq u(x) + \lambda(\bar{a} - x)U'(x) - rU(x), \; x \in [0,1]. \tag{33}$$

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

$$e^{-rt}E[U(\tilde{x}_t)] = U(x_0) + E\left[\int_0^t e^{-rs}(\lambda(\bar{a}-\tilde{x}_t)U'(\tilde{x}_t) - rU(\tilde{x}_t))\mathrm{d}s + \sum_{s\leq t} e^{-rs}\big(\tilde{x}_s U(1) + (1-\tilde{x}_s)U(0) - U(\tilde{x}_s)\big)\right]$$
$$\leq U(x_0) - E\left[\int_0^t e^{-rs}u(\tilde{x}_t)\mathrm{d}s - \sum_{s\leq t} e^{-rs}c\right], \tag{34}$$

where we have used inequalities (32) and (33). 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 (34) holds with equality for the optimal policy. □

B Principal's Problem

B.1 Existence: Proof of Lemma 1

Proof. The first step in the proof is to show that the operator $\mathcal{G}^\theta$ has a unique fixed point. 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 $|\sup f - \sup g| \leq \sup |f-g|$, and so because the function $\mathscr{G}_\epsilon = (\mathscr{G}_\epsilon^L, \mathscr{G}_\epsilon^H)$ is bounded in R, we have that

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

Hence, 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 have that the sequence U_{ϵ_k} is increasing and bounded above by $U^{\max}$: Accordingly, U_{ϵ_k} converges to some limit $\overline{U}$, and because $\mathscr{G}_\epsilon$ is lower semicontinuous as a function of ϵ (Aliprantis and Border, 2006, Lemma 17.29) we also have that

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

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

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

The next step is to show that a solution to the maximization problem exists. To prove existence, we consider the space of probability measures over $\mathbb{R}_+\cup\{\infty\}$, which we denote by $\mathcal{P}$, endowed with the weak* topology. The extended reals $\mathbb{R}_+\cup\{\infty\}$ are a metrizable compact space so by Theorem 15.11 in Aliprantis and Border (2006) the space $\mathcal{P}$ is compact in the weak* topology. The incentive compatibility constraint can be written $\int_\tau^\infty e^{-(r+\lambda)(s-\tau)}\mathrm{d}F(s) \geq \underline{q}(1-F(\tau-))$ for all $\tau \in \mathbb{R}_+ \cup \{\infty\}$ which means that the set of incentive compatible monitoring policies is a closed subset of $\mathcal{P}$, and so a compact set. Finally, the objective function is a bounded linear functional on $C(\mathbb{R}_+ \cup \{\infty\})$ so it is continuous in the weak* topology, and thus is maximized by some incentive compatible policy F^*. □

B.2 Linear Case: Proof of Proposition 2

Proof. Let T be the first monitoring time so the principal's 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 IC constraint and yields a lower expected monitoring cost to the principal. In fact, it is enough to show that if 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 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 postpone somewhat all inspection times before $\tilde{\tau}$ and still satisfy all IC constraints (obviously saving costs). 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 $\tau \in (0, \epsilon_0)$ it is

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

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

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

Define the random variable $z \equiv e^{-(r+\lambda)\tau}$, and let G and $\hat{G}$ be the respective CDFs of z. We have 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, hence it is given by a constant monitoring rate m^*. □

B.3 Convex Case: Necessary Conditions

We start the analysis by deriving some necessary conditions for optimality using Pontryagin's maximum principle for problems with state constraints. In order to guarantee existence, we rely in the general formulation in Arutyunov et al. (2005) for free-time impulse control problem with state constraints that allows for general measures. That being said, this general formulation leads to the same optimality conditions as the ones in the standard maximum principle presented in Seierstad and Sydsaeter (1986). While the results in Arutyunov et al. (2005) covers the case with a finite time horizon, Pereira and Silva (2011) extends the results to consider the infinite horizon case. In addition, because we are considering distributions over the extended real numbers, which are homeomorphic to the unit interval, it is possible to reparameterize the independent variable and work using distributions on discounted times rather than calendar time. In general, the main problem with an infinite horizon is to find the right transversality conditions to pin down a unique candidate for the solution. This is not a problem in our analysis because we do not use the maximum principle to pin down the unique solution. Instead, we use the maximum principle to identify some properties that any candidate policy must satisfy. This allows to restrict the candidate policies to a simple family of distributions. The final solution is found maximizing over this family, which is done in Equation (22). At this point, we only need to solve a one dimensional optimization problem to find the optimal policy.

The statement of Theorem 4.1 in Arutyunov et al. (2005) is quite convoluted. Next, we present the set of conditions in Theorem 4.1 that will be used in the analysis. Let's define

$$\tilde{H}(\tau) = \tilde{\zeta}_\tau \left(rU_\tau - u(x_\tau^\theta) \right) + \tilde{\nu}_\tau (r+\lambda) q_\tau \tag{35a}$$

$$\tilde{S}(\tau) = \tilde{\zeta}_\tau \left(U_\tau - \mathcal{M}(\mathbf{U}, x_\tau^\theta) \right) - \tilde{\nu}_\tau (1 - q_\tau), \tag{35b}$$

where S corresponds to the function Q defined in (Arutyunov et al., 2005, p. 1816). It follows from the system of Equations (4.1) in Arutyunov et al. (2005) that

$$\tilde{\zeta}_\tau = \tilde{\zeta}_0 - \int_0^\tau r\tilde{\zeta}_s \mathrm{d}s - \int_0^\tau \tilde{\zeta}_s \mathrm{d}M_s^c - \sum_k (1 - e^{-\Delta M_{\tau_k}^d})\tilde{\zeta}_{\tau_k -} \tag{36a}$$

$$\tilde{\nu}_\tau = \tilde{\nu}_0 - \int_0^\tau (r+\lambda)\tilde{\nu}_s \mathrm{d}s - \int_0^\tau \tilde{\nu}_s \mathrm{d}M_s^c - \tilde{\Psi}_\tau - \sum_k (1 - e^{-\Delta M_{\tau_k}^d})\tilde{\nu}_{\tau_k -} \tag{36b}$$

$$\tilde{H}(\tau) = \tilde{H}(0) - \int_0^\tau \tilde{\zeta}_s u'(x_s^\theta)\dot{x}_s^\theta \mathrm{d}s - \int_0^\tau \tilde{\zeta}_s \dot{x}_s^\theta (U_H - U_L)\mathrm{d}M_s^c \\ - \sum_k (1 - e^{-\Delta M_{\tau_k}^d})\tilde{\zeta}_{\tau_k -}\dot{x}_{\tau_k}^\theta (U_H - U_L); \tag{36c}$$

where the Lagrange muliplier Ψ_τ is a positive nondecreasing function satisfying

$$\tilde{\Psi}_\tau = \int_0^\tau \mathbf{1}_{\{q_u = \underline{q}\}} \mathrm{d}\tilde{\Psi}_u;$$

and that the adjoint variables satisfy the transversality conditions

$$\tilde{\zeta}_0 = -1$$
$$\tilde{\nu}_0 \leq 0$$
$$\tilde{\nu}_0(q_0 - \underline{q}) = 0$$

In addition, it follows from equation (4.2) that the following optimality and complementary slackness conditions must be satisfied:

$$\tilde{S}(\tau) \leq 0 \tag{37a}$$

$$M_\tau = \int_0^\tau \mathbf{1}_{\{S(u)=0\}} \mathrm{d}M_u. \tag{37b}$$

Noting that the adjoint variables in the optimal control formulation correspond to the derivative of the value function, we verify that (37) correspond to the optimality conditions (16) and (18). In addition, (37) also coincide with the first order conditions from the Hamiltonian maximization in (Seierstad and Sydsaeter, 1986, Theorem 2, p. 332).

As it is common in the analysis of discounted optimal control problems, it is convenient to express all the co-state variables in term of their current value counterparts. We define the following current value multipliers: $\zeta_\tau \equiv e^{r\tau + M_\tau}\tilde{\zeta}_\tau$, $\nu_\tau \equiv e^{r\tau + M_\tau}\tilde{\nu}_\tau$, $H(\tau) \equiv e^{r\tau + M_\tau}\tilde{H}(\tau)$, and $S(\tau) \equiv e^{r\tau + M_\tau}\tilde{S}(\tau)$. It follows that we can write the current value versions of $\tilde{H}$ and $\tilde{S}$ as

$$H(\tau) = \zeta_\tau \left(rU_\tau - u(x_\tau^\theta)\right) + \nu_\tau (r+\lambda) q_\tau \tag{38a}$$

$$S(\tau) = \zeta_\tau \left(U_\tau - \mathcal{M}(\mathbf{U}, x_\tau^\theta)\right) - \nu_\tau (1 - q_\tau) \tag{38b}$$

Next, we define the current value Lagrange multiplier as

$$\Psi_\tau = \tilde{\Psi}_0 + \int_0^\tau e^{rs + M_s} \mathrm{d}\tilde{\Psi}_s.$$

We can invert the previous equation and write

$$\tilde{\Psi}_\tau = \Psi_0 + \int_0^\tau e^{-rs-M_s}\mathrm{d}\Psi_s,$$

and replacing in equations (36a)-(36b) we get

$$\begin{aligned}
e^{-r\tau-M_\tau}\zeta_\tau &= \zeta_0 - \int_0^\tau re^{-rs-M_s}\zeta_s\mathrm{d}s - \int_0^\tau e^{-rs-M_s}\zeta_s\mathrm{d}M_s^c - \sum_k(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\zeta_{\tau_k-}\\
e^{-r\tau-M_\tau}\nu_\tau &= \nu_0 - \int_0^\tau (r+\lambda)e^{-rs-M_s}\nu_s\mathrm{d}s - \int_0^\tau e^{-rs-M_s}\nu_s\mathrm{d}M_s^c - \Psi_0 - \int_0^\tau e^{-rs-M_s}\mathrm{d}\Psi_s\\
&\quad - \sum_k(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\nu_{\tau_k-}.
\end{aligned}$$

It can be readily verified then that $\zeta_\tau = \zeta_0 = -1$. It can also be verified that

$$\nu_\tau = \nu_0 - \int_0^\tau \lambda\nu_s\mathrm{d}s - \Psi_\tau \tag{39}$$

Equation (39) corresponds to the integral representation of the traditional differential equation for the co-state variable (Seierstad and Sydsaeter, 1986, Equation (91) in Theorem 2, p. 332). Notice that the current value adjoint variable ν_τ is continuous at the jump time τ_k. Equation (36c) implies that at any jump time τ_k we have

$$\tilde{H}(\tau_k) - \tilde{H}(\tau_k-) = -(1-e^{-\Delta M^d_{\tau_k}})\tilde{\zeta}_{\tau_k-}\dot{x}^\theta_{\tau_k}(U_H - U_L)$$

This correspond to the same optimality condition as the one in (Seierstad and Sydsaeter, 1986, Note 7, p. 197). By definition of the Hamiltonian $H(\tau)$, we have

$$\tilde{H}(\tau_k) - \tilde{H}(\tau_k-) = \tilde{\zeta}_{\tau_k}\left(rU_{\tau_k} - u(x^\theta_{\tau_k})\right) + \tilde{\nu}_{\tau_k}(r+\lambda)q_{\tau_k} - \tilde{\zeta}_{\tau_k-}\left(rU_{\tau_k-} - u(x^\theta_{\tau_k-})\right) - \tilde{\nu}_{\tau_k-}(r+\lambda)q_{\tau_k-}$$

By defintion, $\tilde{\zeta}_\tau = -e^{-r\tau-M_\tau}$ and $\tilde{\nu}_{\tau_k} = e^{-r\tau_k-M_{\tau_k}}\nu_{\tau_k} = e^{-r\tau_k-M_{\tau_k}}\nu_{\tau_k-} = e^{-\Delta M^d_{\tau_k}}\tilde{\nu}_{\tau_k-}$, so we have that

$$\begin{aligned}
(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\dot{x}^\theta_{\tau_k}(U_H-U_L) &= -e^{-r\tau_k-M_{\tau_k}}\left(rU_{\tau_k} - u(x^\theta_{\tau_k})\right) + e^{-\Delta M^d_{\tau_k}}\tilde{\nu}_{\tau_k-}(r+\lambda)q_{\tau_k}\\
&\quad + e^{-r\tau_k-M_{\tau_k-}}\left(rU_{\tau_k-} - u(x^\theta_{\tau_k-})\right) - \tilde{\nu}_{\tau_k-}(r+\lambda)q_{\tau_k-}\\
(1-e^{-\Delta M^d_{\tau_k}})e^{\Delta M^d_{\tau_k}}\dot{x}^\theta_{\tau_k}(U_H-U_L) &= -rU_{\tau_k} + u(x^\theta_{\tau_k}) + e^{r\tau_k+M_{\tau_k-}}\tilde{\nu}_{\tau_k-}(r+\lambda)q_{\tau_k}\\
&\quad + e^{\Delta M^d_{\tau_k}}rU_{\tau_k-} - e^{\Delta M^d_{\tau_k}}u(x^\theta_{\tau_k-}) - e^{r\tau_k+M_{\tau_k}}\tilde{\nu}_{\tau_k-}(r+\lambda)q_{\tau_k-}\\
(e^{\Delta M^d_{\tau_k}}-1)\dot{x}^\theta_{\tau_k}(U_H-U_L) &= r(e^{\Delta M^d_{\tau_k}}U_{\tau_k-} - U_{\tau_k}) - u(x^\theta_{\tau_k})\left(e^{\Delta M^d_{\tau_k}}-1\right)\\
&\quad - \nu_{\tau_k-}(r+\lambda)\left(e^{\Delta M^d_{\tau_k}}q_{\tau_k-} - q_{\tau_k}\right).
\end{aligned}$$

Replacing the expressions for the jump in U_τ and q_τ given by

$$\begin{aligned}
e^{\Delta M^d_\tau}U_{\tau-} - U_\tau &= (e^{\Delta M^d_\tau}-1)\mathcal{M}(\mathbf{U}, x^\theta_\tau)\\
e^{\Delta M^d_\tau}q_{\tau-} - q_\tau &= e^{\Delta M^d_\tau} - 1
\end{aligned}$$

we find that

$$(e^{\Delta M^d_{\tau_k}}-1)\dot{x}^\theta_{\tau_k}(U_H-U_L)=r(e^{\Delta M^d_{\tau_k}}-1)\mathcal{M}(\mathbf{U},x^\theta_{\tau_k})-u(x^\theta_{\tau_k})\left(e^{\Delta M^d_{\tau_k}}-1\right)$$
$$-\nu_{\tau_k-}(r+\lambda)\left(e^{\Delta M^d_{\tau_k}}-1\right).$$

Simplifying, the previous condition reduces to the following optimality condition at the atom:

$$r\mathcal{M}(\mathbf{U},x^\theta_{\tau_k})=u(x^\theta_{\tau_k})+\dot{x}^\theta_{\tau_k}(U_H-U_L)+(r+\lambda)\nu_{\tau_k}. \tag{40}$$

This also corresponds to the transversality condition at free time $\bar{\tau}$ in Equation (4.4) in Arutyunov et al. (2005). Moreover, this condition also coincides with the optimality condition (21) in our heuristic derivation using dynamic programming, and also coincides with the condition for free final time problems in (Seierstad and Sydsaeter, 1986, Equation (152) in Theorem 16, p. 398).

B.4 Proof of Theorem 1

Proof of Lemma 3

Proof. At any point of continuity we have that

$$\mathrm{d}\nu_\tau=-\lambda\nu_\tau\mathrm{d}\tau-\mathrm{d}\Psi_\tau. \tag{41}$$

We also have the optimality conditions

$$S(\tau)\le 0 \tag{42a}$$
$$M_\tau=\int_0^\tau \mathbf{1}_{\{S(u)=0\}}\mathrm{d}M_u. \tag{42b}$$

Condition (42a) corresponds to

$$S(\tau)=\mathcal{M}(\mathbf{U},x^\theta_\tau)-U_\tau-(1-q_\tau)\nu_\tau\le 0.$$

Differentiation $S(\tau)$ we find

$$\begin{aligned}\mathrm{d}S(\tau)=&\dot{x}^\theta_\tau(U_H-U_L)\mathrm{d}\tau-\mathrm{d}U_\tau+\nu_\tau\mathrm{d}q_\tau-(1-q_\tau)\mathrm{d}\nu_\tau\\ =&\dot{x}^\theta_\tau(U_H-U_L)\mathrm{d}\tau-\left(rU_\tau-u(x^\theta_\tau)\right)\mathrm{d}\tau-\left(U_\tau-\mathcal{M}(\mathbf{U},x^\theta_\tau)\right)\mathrm{d}M^c_\tau\\ &+\nu_\tau\left((r+\lambda)q_\tau\mathrm{d}t-(1-q_\tau)\mathrm{d}M^c(\tau)\right)+(1-q_\tau)\left(\lambda\nu_\tau\mathrm{d}\tau+\mathrm{d}\Psi_\tau\right)\\ =&\left(\dot{x}^\theta_\tau(U_H-U_L)+u(x^\theta_\tau)-rU_\tau+\nu_\tau(rq_\tau+\lambda)\right)\mathrm{d}t+(1-q_\tau)\mathrm{d}\Psi_\tau+S(\tau)\mathrm{d}M^c_\tau\end{aligned}$$

The optimality condition (42b) implies that $S(\tau)\mathrm{d}M_\tau=0$. Thus we can write the evolution of $S(\tau)$ as

$$\mathrm{d}S(\tau)=\left(\dot{x}^\theta_\tau(U_H-U_L)+u(x^\theta_\tau)-rU_\tau+\nu_\tau(rq_\tau+\lambda)\right)\mathrm{d}t+(1-q_\tau)\mathrm{d}\Psi_\tau. \tag{43}$$

Whenever $q_\tau>\underline{q}$ we have that $\mathrm{d}\Psi_\tau=0$, which means that $S(\tau)$ is absolutely continuous in any interval (τ',τ'') with $q_\tau>\underline{q}$ (notice that q_τ is continuous between jumps so wlog we can assume that if $q_{\tilde{\tau}}>\underline{q}$ at some time $\tilde{\tau}$ between jumps then there is neighborhood of $\tilde{\tau}$ such that $q_\tau>\underline{q}$) . Note as well that $S(\tau)\mathrm{d}M^c_\tau=0$

implies that we can write

$$\mathrm{d}U_\tau = \left(rU_\tau - u(x^\theta_\tau)\right)\mathrm{d}\tau - (1-q_\tau)\nu_\tau \mathrm{d}M^c_\tau \tag{44}$$

Let $\dot{S}(\tau)$ denote the drift of $S(\tau)$, which is given by

$$\dot{S}(\tau) \equiv \dot{x}^\theta_\tau(U_H - U_L) + u(x^\theta_\tau) - rU_\tau + \nu_\tau(rq_\tau + \lambda). \tag{45}$$

Differentiating $\dot{S}(\tau)$ we find

$$\mathrm{d}\dot{S}(\tau) = \left(\ddot{x}^\theta_\tau(U_H - U_L) + u'(x^\theta_\tau)\dot{x}^\theta_\tau\right)\mathrm{d}\tau - r\mathrm{d}U_\tau + (rq_\tau + \lambda)\mathrm{d}\nu_\tau + r\nu_\tau \mathrm{d}q_\tau \tag{46}$$

Replacing equations (41) and (44), and the equation for $\mathrm{d}q_\tau$ in (46) we find that

$$\mathrm{d}\dot{S}(\tau) = \left(r^{-1}\ddot{x}^\theta_\tau(U_H - U_L) + r^{-1}u'(x^\theta_\tau)\dot{x}^\theta_\tau - rU_\tau + u(x^\theta_\tau) + r^{-1}(r^2 q_\tau - \lambda^2)\nu_\tau\right)\mathrm{d}\tau. \tag{47}$$

The support of M^c_τ is $A \equiv \{\tau : S(\tau) = 0\}$, which correspond to the set of maximizers of $S(\tau)$. Accordingly, for any time $\tau \in A$, we have that $S(\tau) = \dot{S}(\tau) = 0$ and $\ddot{S}(\tau) \le 0$. Suppose that there is $\tilde{\tau}$ such that $S(\tau) = 0$, $\dot{S}(\tilde{\tau}) = 0$ and $\ddot{S}(\tilde{\tau}) = 0$, and replacing $S(\tilde{\tau}) = 0$ and $\ddot{x}^\theta_\tau = -\lambda\dot{x}^\theta_\tau$ in (47), then we get that it must be the case that

$$\nu_{\tilde{\tau}} = \frac{\dot{x}^\theta_{\tilde{\tau}}}{\lambda(r+\lambda)}\left(u'(x^\theta_{\tilde{\tau}}) - (r+\lambda)(U_H - U_L)\right) \tag{48}$$

Let's define

$$z_\tau \equiv \frac{\dot{x}^\theta_\tau}{\lambda(r+\lambda)}\left(u'(x^\theta_\tau) - (r+\lambda)(U_H - U_L)\right).$$

Differentiating z_τ we get

$$\begin{aligned}
\mathrm{d}z_\tau &= \left(\frac{\ddot{x}^\theta_\tau}{\lambda(r+\lambda)}\left(u'(x^\theta_\tau) - (r+\lambda)(U_H - U_L)\right) + \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau)\right)\mathrm{d}t \\
&= \left(\frac{\ddot{x}^\theta_\tau}{\dot{x}^\theta_\tau}z_\tau + \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau)\right)\mathrm{d}\tau \\
&= \left(-\lambda z_\tau + \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau)\right)\mathrm{d}\tau
\end{aligned}$$

On the other hand, whenever $q_\tau > \underline{q}$ we have that $\mathrm{d}\Psi_\tau = 0$ so

$$\mathrm{d}\nu_\tau = -\lambda\nu_\tau \mathrm{d}\tau.$$

Accordingly

$$\mathrm{d}(\nu_\tau - z_\tau) = -\lambda(\nu_\tau - z_\tau)\mathrm{d}\tau - \frac{(\dot{x}^\theta_\tau)^2}{\lambda(r+\lambda)}u''(x^\theta_\tau)\mathrm{d}\tau,$$

so for any $\tau > \tilde{\tau}$

$$\nu_\tau - z_\tau = \int_{\tilde{\tau}}^{\tau} e^{-\lambda(\tau - s)}\frac{(\dot{x}^\theta_s)^2}{\lambda(r+\lambda)}u''(x^\theta_s)\mathrm{d}s > 0.$$

This means that there is at most one $\tilde{\tau} \in A$ satisfying equation (48), which means that there is at most one $\tilde{\tau} \in A$ such that $\ddot{S}(\tilde{\tau}) = 0$, and any other $\tau \in A$ satisfies $\ddot{S}(\tau) < 0$. This means that all, but at most one, $\tau \in A$, are isolated points. And, by Theorem 7.14.23 in (Bogachev, 2007), the only atomless measure in A is the trivial zero measure, which means that $M^c_\tau - M^c_{\tau'} = 0$ for all $\tau \in [\tau', \tau'')$

□

Proof of Lemma 4

Proof. The fist step is to verify that $S(\tau)$ is continuous at any atom τ_k. We have that

$$S(\tau_k-) = \mathcal{M}(\mathbf{U}, x^\theta_{\tau_k}) - U_{\tau_k-} - \nu_{\tau_k-}(1 - q_{\tau_k-})$$

Using the fact that ν_τ is continuous at a jump τ_k, we find that

$$\begin{aligned} S(\tau_k) &= \mathcal{M}(\mathbf{U}, x^\theta_{\tau_k}) - U_{\tau_k} - \nu_{\tau_k-}(1 - q_{\tau_k}) \\ &= e^{\Delta M^d_{\tau_k}} \left(\mathcal{M}(\mathbf{U}, x^\theta_{\tau_k}) - U_{\tau_k-} - \nu_{\tau_k-}(1 - q_{\tau_k-})\right) \\ &= e^{\Delta M^d_{\tau_k}} S(\tau_k-) \end{aligned}$$

Hence, $S(\tau_k-) = S(\tau_k) = 0$. At any jump time, the following necessary condition must hold

$$r\mathcal{M}(\mathbf{U}, x^\theta_{\tau_k}) = u(x^\theta_{\tau_k}) + \dot{x}^\theta_{\tau_k}(U_H - U_L) + (r + \lambda)\nu_{\tau_k} \tag{49}$$

The objective now is to show that equation (49) cannot be satisfied at more than one point. Let's define

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

And notice that

$$\dot{S}(\tau) = rS(\tau) + G(\tau) = u(x^\theta_\tau) + \dot{x}^\theta_\tau(U_H - U_L) - rU_\tau + \nu_\tau(rq_\tau + \lambda). \tag{50}$$

Accordingly, for any atom τ_k, the following conditions must be satisfied

$$\begin{aligned} S(\tau_k-) &= 0 \\ G(\tau_k-) &= 0 \\ \dot{S}(\tau_k-) &= 0. \end{aligned}$$

We have from equation (43) that

$$\mathrm{d}S(\tau) = \dot{S}(\tau)\mathrm{d}\tau + (1 - q_\tau)\mathrm{d}\Psi_\tau$$

Notice that, because both $G(\tau)$ and $S(\tau)$ are continuous at the atom τ_k, so it is $\dot{S}(\tau)$. Moreover, because τ_k is a local maximum of $S(\tau)$ and $S(\tau_k-) = S(\tau_k) = \dot{S}(\tau_k-) = \dot{S}(\tau_k) = 0$, it must be the case that $\mathrm{d}\Psi_{\tau_k} = 0$. It follows that $S(\tau)$ is differentiable at τ_k and that $\ddot{S}(\tau_k-) \leq 0$, and equation (50) then implies that $\dot{G}(\tau_k-) \leq 0$. Differentiating $G(\tau)$, we find that

$$\mathrm{d}G(\tau) = \left(u'(x^\theta_\tau)\dot{x}^\theta_\tau - (r + \lambda)\left(\dot{x}^\theta_\tau(U_H - U_L) + \lambda\nu_\tau\right)\right)\mathrm{d}\tau - (r + \lambda)\mathrm{d}\Psi_\tau$$

Let's $J(\tau)$ be given by

$$J(\tau) \equiv u'(x^\theta_\tau)\dot{x}^\theta_\tau - (r + \lambda)\dot{x}^\theta_\tau(U_H - U_L) - (r + \lambda)\lambda\nu_\tau. \tag{51}$$

Notice that whenever the IC constraint is slack we have $\dot{G}(\tau) = J(\tau)$, so in particular $\dot{G}(\tau_k-) = J(\tau_k-)$ for any atom τ_k. Next, if we differentiate equation (51) we get

$$\begin{aligned}\mathrm{d}J(\tau) &= \left(u''(x_\tau^\theta)(\dot{x}_\tau^\theta)^2 - \lambda u'(x_\tau^\theta)\dot{x}_\tau^\theta - (r+\lambda)(-\lambda\dot{x}_\tau^\theta)(U_H - U_L)\right)\mathrm{d}\tau - \lambda(r+\lambda)\mathrm{d}\nu_\tau \\ &= \left(u''(x_\tau^\theta)(\dot{x}_\tau^\theta)^2 - \lambda u'(x_\tau^\theta)\dot{x}_\tau^\theta + (r+\lambda)\lambda\dot{x}_\tau^\theta(U_H - U_L)\right)\mathrm{d}\tau + \lambda(r+\lambda)(\lambda\nu_\tau\mathrm{d}\tau + \mathrm{d}\Psi_\tau) \\ &= \left(u''(x_\tau^\theta)(\dot{x}_\tau^\theta)^2 - \lambda u'(x_\tau^\theta)\dot{x}_\tau^\theta + (r+\lambda)\lambda(\dot{x}_\tau^\theta(U_H - U_L) + \lambda\nu_\tau)\right)\mathrm{d}\tau + \lambda(r+\lambda)\mathrm{d}\Psi_\tau,\end{aligned}$$

which can be rewritten as

$$\mathrm{d}J(\tau) = \left(u''(x_\tau^\theta)(\dot{x}_\tau^\theta)^2\right)\mathrm{d}\tau - \lambda J(\tau)\mathrm{d}\tau + \lambda(r+\lambda)d\Psi_\tau.$$

Thus, for any $\tau \in [\tau_k, \tau_{k+1})$ we have

$$\begin{aligned}J(\tau) &= -\int_\tau^{\tau_{k+1}} e^{\lambda(s-\tau)}\left((\dot{x}_s^\theta)^2 u''(x_s^\theta)\mathrm{d}s + \lambda(r+\lambda)\mathrm{d}\Psi_s\right) + e^{\lambda(\tau_{k+1}-\tau)}J(\tau_{k+1}-) \\ &= -\int_\tau^{\tau_{k+1}} e^{\lambda(s-\tau)}\left((\dot{x}_s^\theta)^2 u''(x_s^\theta)\mathrm{d}s + \lambda(r+\lambda)\mathrm{d}\Psi_s\right) + e^{\lambda(\tau_{k+1}-\tau)}\dot{G}(\tau_{k+1}-) < 0,\end{aligned}$$

where we have used the fact that $J(\tau_{k+1}-) = \dot{G}(\tau_{k+1}-) \leq 0$. But then,

$$\mathrm{d}G(\tau) = J(\tau)\mathrm{d}\tau - (r+\lambda)\mathrm{d}\Psi_\tau < 0$$

for all $\tau \in (\tau_k, \tau_{k+1})$ which contradicts the requirement that $G(\tau_{k+1}-) = 0$.

□

Proof of Theorem 1

Proof. Lemma 3 implies that, in the absence of an atom, q_τ is increasing if $q_\tau > \underline{q}$ because q_τ increases whenever $\mathrm{d}M_\tau^{c*} = 0$. Hence, because there is at most one atom, this means that either there is monitoring with probability one at the atom, or the incentive compatibility constraint is binding thereafter. If this were not the case, q_τ would eventually reach one, which would require a second atom and contradict lemma 4. Thus lemmas 3 and 4 imply that the optimal monitoring policy takes the following form:

1. There is $\tilde{\tau}$ such that for any $\tau \in [0, \tilde{\tau})$ we have $q_\tau = \underline{q}$.
2. There is $\hat{\tau}$ such that for any $\tau \in [\tilde{\tau}, \hat{\tau})$ there is no monitoring and $q_\tau > \underline{q}$.
3. There is an atom at time $\hat{\tau}$. If the probability of monitoring at the atom is less than one, then there is a constant rate of monitoring after $\hat{\tau}$.

Thus, the problem of solving for the optimal policy is reduced to finding $\tilde{\tau}$ and $\hat{\tau}$. The last step of the proof shows that $\tilde{\tau}$ is either zero or infinity. The intuition is the following. Analogous to standard contracting models, equation (10) works as a promise keeping constraint. Equation (11) implies that the largest possible atom consistent with $q_{\tau-}$ is $(q_{\tau-} - \underline{q})/(1-\underline{q})$, which corresponds to the atom in Theorem 1. On the other hand, once the incentive compatibility constraint is binding, equation (10) implies that the largest monitoring rate consistent with the promise keeping and the incentive compatibility constraint is m^*. Thus, because the benefit of monitoring is increasing over time, the optimal policy requires to perform as much monitoring as possible once it becomes profitable to do so. Hence, the support of the monitoring distribution is either a singleton (deterministic monitoring) or an interval $[\hat{\tau}, \infty]$.

First, notice that any atom has to be of size

$$\Delta M_\tau^d = \log\left(\frac{1-\underline{q}}{1-q_{\tau-}}\right),$$

and that the continuation payoff at the atom date satisfies

$$U_{\tau-} = \left(\frac{1-q_{\tau-}}{1-\underline{q}}\right) U_\tau + \left(\frac{q_{\tau-}-\underline{q}}{1-\underline{q}}\right) \mathcal{M}(\mathbf{U}, x_\tau^\theta)$$

Whenever the IC constraint is binding on an interval of time, the monitoring rate is given by

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

The payoff at time zero of a policy with monitoring at a rate m in $[0,\tilde{\tau})$ and an atom at time $\hat{\tau} = \tilde{\tau}+\delta$ is

$$\begin{aligned}\mathcal{U}(\tilde{\tau},\delta) = \int_0^{\tilde{\tau}} e^{-(r+m)\tau}\left(u(x_\tau^\theta) + m\mathcal{M}(\mathbf{U}, x_\tau^\theta)\right)\mathrm{d}\tau + \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r\tau - m\hat{\tau}} u(x_\tau^\theta)\mathrm{d}\tau \\ + e^{-r(\tilde{\tau}+\delta)-m\tilde{\tau}}\left[\left(\frac{1-q_{\tilde{\tau}+\delta-}}{1-\underline{q}}\right)U_{\tilde{\tau}+\delta} + \left(\frac{q_{\tilde{\tau}+\delta-}-\underline{q}}{1-\underline{q}}\right)\mathcal{M}(\mathbf{U}, x_{\tilde{\tau}+\delta}^\theta)\right] \quad (52)\end{aligned}$$

where

$$U_{\tilde{\tau}+\delta} = \int_{\tilde{\tau}+\delta}^{\infty} e^{-(r+m)(\tau-\tilde{\tau}-\delta)}\left(u(x_\tau^\theta) + m\mathcal{M}(\mathbf{U}, x_\tau^\theta)\right)\mathrm{d}\tau$$

Suppose that the IC constraint is binding at time 0, that is assume that $q_0 = \underline{q}$, then we have that

$$q_{\tilde{\tau}+\delta-} = e^{(r+\lambda)\delta}\underline{q},$$

which means that δ must satisfy

$$\delta \leq \frac{1}{r+\lambda}\log\frac{1}{\underline{q}}.$$

Replacing $q_{\tilde{\tau}+\delta-}$ in (52) we get

$$\begin{aligned}\mathcal{U}(\tilde{\tau},\delta) = \int_0^{\tilde{\tau}} e^{-(r+m)\tau}\left(u(x_\tau^\theta) + m\mathcal{M}(\mathbf{U}, x_\tau^\theta)\right)\mathrm{d}\tau + \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r\tau - m\tilde{\tau}} u(x_\tau^\theta)\mathrm{d}\tau \\ + e^{-r(\tilde{\tau}+\delta)-m\tilde{\tau}}\left[\left(\frac{1-e^{(r+\lambda)\delta}\underline{q}}{1-\underline{q}}\right)U_{\tilde{\tau}+\delta} + \left(\frac{e^{(r+\lambda)\delta}-1}{1-\underline{q}}\right)\underline{q}\mathcal{M}(\mathbf{U}, x_{\tilde{\tau}+\delta}^\theta)\right] \quad (53)\end{aligned}$$

Next, we show that for any given δ we have that $\partial\mathcal{U}(\tilde{\tau},\delta)/\partial\tilde{\tau} = 0 \Rightarrow \partial^2\mathcal{U}(\tilde{\tau},\delta)/\partial\tilde{\tau}^2 > 0$. This means that the maximum cannot have an interior value for $\tilde{\tau}$.

Differentiating (53) we get

$$\frac{\partial}{\partial\tilde{\tau}}\mathcal{U}(\tilde{\tau},\delta) = e^{-(r+m)\tilde{\tau}}\left(u(x^\theta_{\tilde{\tau}}) + m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}})\right) + e^{-r(\tilde{\tau}+\delta)-m\tilde{\tau}}u(x^\theta_{\tilde{\tau}+\delta}) - e^{-(r+m)\tilde{\tau}}u(x^\theta_{\tilde{\tau}})$$
$$- m\int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r\tau-m\tilde{\tau}}u(x^\theta_\tau)\mathrm{d}\tau - (r+m)e^{-r(\tilde{\tau}+\delta)-m\tilde{\tau}}\left[\left(\frac{1-e^{(r+\lambda)\delta}\underline{q}}{1-\underline{q}}\right)U_{\tilde{\tau}+\delta} + \left(\frac{e^{(r+\lambda)\delta}-1}{1-\underline{q}}\right)\underline{q}\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta})\right]$$
$$+ e^{-r(\tilde{\tau}+\delta)-m\tilde{\tau}}\left[\left(\frac{1-e^{(r+\lambda)\delta}\underline{q}}{1-\underline{q}}\right)\frac{\partial}{\partial\tilde{\tau}}U_{\tilde{\tau}+\delta} + \left(\frac{e^{(r+\lambda)\delta}-1}{1-\underline{q}}\right)\underline{q}\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H - U_L)\right]$$

where

$$\frac{\partial}{\partial\tilde{\tau}}U_{\tilde{\tau}+\delta} = -u(x^\theta_{\tilde{\tau}+\delta}) - m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) + (r+m)U_{\tilde{\tau}+\delta}$$

Rearranging terms we get

$$\begin{aligned}
\frac{\partial}{\partial\tilde{\tau}}\mathcal{U}(\tilde{\tau},\delta) &= e^{-(r+m)\tilde{\tau}}\Big[m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}}) + e^{-r\delta}u(x^\theta_{\tilde{\tau}+\delta}) \\
&- m\int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})}u(x^\theta_\tau)\mathrm{d}\tau - (r+m)\left(\frac{e^{\lambda\delta}-e^{-r\delta}}{1-\underline{q}}\right)\underline{q}\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) \\
&- \left(\frac{e^{-r\delta}-e^{\lambda\delta}\underline{q}}{1-\underline{q}}\right)\left(u(x^\theta_{\tilde{\tau}+\delta}) + m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta})\right) + \left(\frac{e^{\lambda\delta}-e^{-r\delta}}{1-\underline{q}}\right)\underline{q}\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H-U_L)\Big] \\
&= e^{-(r+m)\tilde{\tau}}\Big[m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}}) + (e^{\lambda\delta}-e^{-r\delta})\left(\frac{\underline{q}}{1-\underline{q}}\right)u(x^\theta_{\tilde{\tau}+\delta}) \\
&- m\int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})}u(x^\theta_\tau)\mathrm{d}\tau - (r+m)(e^{\lambda\delta}-e^{-r\delta})\left(\frac{\underline{q}}{1-\underline{q}}\right)\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) \\
&- \left(\frac{e^{-r\delta}-e^{\lambda\delta}\underline{q}}{1-\underline{q}}\right)m\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) + (e^{\lambda\delta}-e^{-r\delta})\left(\frac{\underline{q}}{1-\underline{q}}\right)\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H-U_L)\Big] \\
&= e^{-(r+m)\tilde{\tau}}m\Big[\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}u(x^\theta_{\tilde{\tau}+\delta}) \\
&- \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})}u(x^\theta_\tau)\mathrm{d}\tau - (r+m)\frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) \\
&- \left(\frac{e^{-r\delta}-e^{\lambda\delta}\underline{q}}{1-\underline{q}}\right)\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H-U_L)\Big]
\end{aligned}$$

So, finally, we can write

$$\frac{\partial}{\partial\tilde{\tau}}\mathcal{U}(\tilde{\tau},\delta) = e^{-(r+m)\tilde{\tau}}m\Big[\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}u(x^\theta_{\tilde{\tau}+\delta})$$
$$- \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})}u(x^\theta_\tau)\mathrm{d}\tau - \left(\frac{r}{r+\lambda}e^{\lambda\delta} + \frac{\lambda}{r+\lambda}e^{-r\delta}\right)\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H-U_L)\Big]$$

Let's define

$$G(\tilde{\tau}) = \mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}u(x^\theta_{\tilde{\tau}+\delta})$$
$$- \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})}u(x^\theta_\tau)\mathrm{d}\tau - \left(\frac{r}{r+\lambda}e^{\lambda\delta} + \frac{\lambda}{r+\lambda}e^{-r\delta}\right)\mathcal{M}(\mathbf{U},x^\theta_{\tilde{\tau}+\delta}) + \frac{(e^{\lambda\delta}-e^{-r\delta})}{r+\lambda}\dot{x}^\theta_{\tilde{\tau}+\delta}(U_H-U_L)$$

So

$$\frac{\partial}{\partial \tilde{\tau}} \mathcal{U}(\tilde{\tau}, \delta) = e^{-(r+m)\tilde{\tau}} m G(\tilde{\tau})$$

Clearly, the first order condition is satisfied only if $G(\tilde{\tau}) = 0$. Moreover, $G(\tilde{\tau}) = 0$ implies that $\frac{\partial^2}{\partial \tilde{\tau}^2} \mathcal{U}(\tilde{\tau}, \delta) = G'(\tilde{\tau})$. Differentiating $G(\tilde{\tau})$ we get

$$\begin{aligned}
G'(\tilde{\tau}) &= \dot{x}^{\theta}_{\tilde{\tau}}(U_H - U_L) + \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) + u(x^{\theta}_{\tilde{\tau}}) \\
&\quad - r \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u(x^{\theta}_{\tau}) \mathrm{d}\tau - \left(\frac{r}{r+\lambda} e^{\lambda\delta} + \frac{\lambda}{r+\lambda} e^{-r\delta} \right) \dot{x}^{\theta}_{\tilde{\tau}+\delta}(U_H - U_L) + \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} \ddot{x}^{\theta}_{\tilde{\tau}+\delta}(U_H - U_L) \\
&= \dot{x}^{\theta}_{\tilde{\tau}}(U_H - U_L) + \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) + u(x^{\theta}_{\tilde{\tau}}) \\
&\quad - r \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u(x^{\theta}_{\tau}) \mathrm{d}\tau - \left(\frac{r}{r+\lambda} e^{\lambda\delta} + \frac{\lambda}{r+\lambda} e^{-r\delta} \right) \dot{x}^{\theta}_{\tilde{\tau}+\delta}(U_H - U_L) - \lambda \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} \dot{x}^{\theta}_{\tilde{\tau}+\delta}(U_H - U_L) \\
&= \left(\dot{x}^{\theta}_{\tilde{\tau}} - e^{\lambda\delta} \dot{x}^{\theta}_{\tilde{\tau}+\delta} \right)(U_H - U_L) + \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) + u(x^{\theta}_{\tilde{\tau}}) \\
&\quad - r \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u(x^{\theta}_{\tau}) \mathrm{d}\tau
\end{aligned}$$

Noting that

$$\frac{\partial}{\partial \delta} e^{\lambda\delta} \dot{x}^{\theta}_{\tilde{\tau}+\delta} = \lambda e^{\lambda\delta} \dot{x}^{\theta}_{\tilde{\tau}+\delta} + e^{\lambda\delta} \ddot{x}^{\theta}_{\tilde{\tau}+\delta} = \lambda e^{\lambda\delta} \dot{x}^{\theta}_{\tilde{\tau}+\delta} - \lambda e^{\lambda\delta} \dot{x}^{\theta}_{\tilde{\tau}+\delta} = 0$$

we conclude that

$$G'(\tilde{\tau}) = \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) + u(x^{\theta}_{\tilde{\tau}}) - r \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u(x^{\theta}_{\tau}) \mathrm{d}\tau \tag{54}$$

Using integration by parts we find that

$$-r \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u(x^{\theta}_{\tau}) \mathrm{d}\tau = e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) - u(x^{\theta}_{\tilde{\tau}}) - \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u'(x^{\theta}_{\tau}) \dot{x}^{\theta}_{\tau} \mathrm{d}\tau$$

Replacing in equation (54) we arrive to

$$\begin{aligned}
G'(\tilde{\tau}) &= \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) + u(x^{\theta}_{\tilde{\tau}}) + e^{-r\delta} u(x^{\theta}_{\tilde{\tau}+\delta}) - u(x^{\theta}_{\tilde{\tau}}) - \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u'(x^{\theta}_{\tau}) \dot{x}^{\theta}_{\tau} \mathrm{d}\tau \\
&= \frac{(e^{\lambda\delta} - e^{-r\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) \dot{x}^{\theta}_{\tilde{\tau}+\delta} - \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-r(\tau-\tilde{\tau})} u'(x^{\theta}_{\tau}) \dot{x}^{\theta}_{\tau} \mathrm{d}\tau \qquad (55)
\end{aligned}$$

Replacing $\dot{x}^{\theta}_{\tau} = \lambda(\bar{a} - \theta) e^{-\lambda\tau}$ in equation (55) we get

$$G'(\tilde{\tau}) = \lambda(\bar{a} - \theta) e^{-\lambda\tilde{\tau}} \left[\frac{(1 - e^{-(r+\lambda)\delta})}{r+\lambda} u'(x^{\theta}_{\tilde{\tau}+\delta}) - \int_{\tilde{\tau}}^{\tilde{\tau}+\delta} e^{-(r+\lambda)(\tau-\tilde{\tau})} u'(x^{\theta}_{\tau}) \mathrm{d}\tau \right] \tag{56}$$

On the one hand, if $\theta = 0$, then we have that $u'(x^\theta_{\tilde{\tau}+\delta}) > u'(x^\theta_\tau)$ for all $\tilde{\tau} + \delta > \tau$, which means that

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

On the other hand, if $\theta = 1$, then we have that $u'(x^\theta_{\tilde{\tau}+\delta}) < u'(x^\theta_\tau)$ for all $\tilde{\tau} + \delta > \tau$, which means that

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

This means that, for any $\delta \geq 0$, we have $\frac{\partial}{\partial\tilde{\tau}} \mathcal{U}(\tilde{\tau}, \delta) = 0$ implies $\frac{\partial^2}{\partial\tilde{\tau}^2} \mathcal{U}(\tilde{\tau}, \delta) > 0$ which means that the optimal monitoring policy can not have an interior $\tilde{\tau}$, that is $\tilde{\tau}^* \in \{0, \infty\}$.

□

Proof of Proposition 3

Comparative static c: Let G_{det} and G_{rand} be the maximization problems in the operators above so we write the optimization in the fixed point problem as

$$\max_{\alpha \in [0,1]} \alpha G_{\text{rand}} + (1 - \alpha) G_{\text{det}}$$

We can fix the continuation values and show that we have single crossing in (c, U_H, U_L). In the previous expressions, we have that

$$\begin{aligned}
\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial(-c)} &= e^{-r\hat{\tau}} \left(\frac{r}{r + \lambda\underline{q}} e^{(r+\lambda)\hat{\tau}} \underline{q} + \frac{\lambda\underline{q}}{r + \lambda\underline{q}} \right) - e^{-r\bar{\tau}} \\
&= e^{\lambda\hat{\tau}} \underline{q} \left(\frac{r}{r + \lambda\underline{q}} + e^{-(r+\lambda)\hat{\tau}} \frac{\lambda}{r + \lambda\underline{q}} \right) - e^{-(r+\lambda)\bar{\tau}} e^{\lambda\bar{\tau}} \\
&\leq e^{\lambda\hat{\tau}} \underline{q} - e^{-(r+\lambda)\tau_{\text{bind}}} e^{\lambda\bar{\tau}} \\
&= \left(e^{\lambda\hat{\tau}} - e^{\lambda\bar{\tau}} \right) \underline{q}
\end{aligned}$$

which is negative if $\hat{\tau} < \bar{\tau}$. Next, we have that

$$\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial U_H} = e^{-r\hat{\tau}} \left[\left(\frac{e^{(r+\lambda)\hat{\tau}} - 1}{1 - \underline{q}} \right) \underline{q} x^{\theta}_{\hat{\tau}} + \left(\frac{1 - e^{(r+\lambda)\hat{\tau}}\underline{q}}{1 - \underline{q}} \right) \int_{\hat{\tau}}^{\infty} e^{-(r+m)(\tau - \hat{\tau})} m x^{\theta}_{\tau} \mathrm{d}\tau \right] - e^{-r\bar{\tau}} x^{\theta}_{\bar{\tau}}$$

If we replace

$$\int_{\hat{\tau}}^{\infty} e^{-(r+m)(\tau-\hat{\tau})} x^{\theta}_{\tau} \mathrm{d}\tau = \frac{\bar{a}}{r+m} + \frac{x^{\theta}_{\hat{\tau}} - \bar{a}}{r + \lambda + m}$$

$$\int_{\hat{\tau}}^{\infty} e^{-(r+m)(\tau-\hat{\tau})} (1 - x^{\theta}_{\tau}) \mathrm{d}\tau = \frac{1 - \bar{a}}{r+m} - \frac{x^{\theta}_{\hat{\tau}} - \bar{a}}{r + \lambda + m},$$

and after some tedious simplifications we obtain

$$\begin{aligned}
\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial U_H} &= e^{-r\hat{\tau}} \left[(1 - \underline{q}) e^{(r+\lambda)\hat{\tau}} \frac{m}{r+\lambda} x^{\theta}_{\hat{\tau}} + \left(1 - e^{(r+\lambda)\hat{\tau}}\underline{q}\right) \frac{\lambda(1-\underline{q})}{(r+\lambda\underline{q})(r+\lambda)} m\bar{a} \right] - e^{-r\bar{\tau}} x^{\theta}_{\bar{\tau}} \\
&= e^{\lambda\hat{\tau}} \underline{q} x^{\theta}_{\hat{\tau}} + \left(e^{-r\hat{\tau}} - e^{\lambda\hat{\tau}}\underline{q}\right) \frac{\lambda\underline{q}}{r+\lambda\underline{q}} \bar{a} - e^{-r\bar{\tau}} x^{\theta}_{\bar{\tau}}
\end{aligned}$$

Noticing that

$$e^{\lambda\tau} x^{\theta}_{\tau} = \theta + \bar{a}(e^{\lambda\tau} - 1),$$

we obtain

$$\begin{aligned}
\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial U_H} &= e^{\lambda\hat{\tau}} \underline{q} x^{\theta}_{\hat{\tau}} + \left(e^{-r\hat{\tau}} - e^{\lambda\hat{\tau}}\underline{q}\right) \frac{\lambda\underline{q}}{r+\lambda\underline{q}} \bar{a} - e^{-r\bar{\tau}} x^{\theta}_{\bar{\tau}} \\
&= \underline{q}\,(\theta - \bar{a}) + \left[e^{\lambda\hat{\tau}} \frac{r}{r+\lambda\underline{q}} + e^{-r\hat{\tau}} \frac{\lambda}{r+\lambda\underline{q}} \right] \underline{q}\bar{a} - e^{-(r+\lambda)\bar{\tau}} \left(\theta + \bar{a}(e^{\lambda\bar{\tau}} - 1)\right) \\
&\leq \underline{q}\,(\theta - \bar{a}) + \left[e^{\lambda\hat{\tau}} \frac{r}{r+\lambda\underline{q}} + e^{-r\hat{\tau}} \frac{\lambda}{r+\lambda\underline{q}} \right] \underline{q}\bar{a} - \underline{q}(\theta - \bar{a}) - e^{-r\bar{\tau}} \bar{a} \\
&= \left[e^{\lambda\hat{\tau}} \frac{r}{r+\lambda\underline{q}} + e^{-r\hat{\tau}} \frac{\lambda}{r+\lambda\underline{q}} \right] \underline{q}\bar{a} - e^{-r\bar{\tau}} \bar{a}
\end{aligned}$$

The last expression is increasing in $\hat{\tau}$, which means that if $\hat{\tau} \leq \bar{\tau}$ then

$$\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial U_H} \leq -e^{\lambda\bar{\tau}} \left(e^{-(r+\lambda)\bar{\tau}} - \underline{q}\right) \frac{r\bar{a}}{r+\lambda\underline{q}} \leq 0,$$

where the last inequality follows from the IC constraint. We can repeat the same calculations for U_L.

$$\begin{aligned}
\frac{\partial(G_{\text{rand}} - G_{\text{det}})}{\partial U_L} &= e^{-r\hat{\tau}} \left[\left(\frac{e^{(r+\lambda)\hat{\tau}} - 1}{1 - \underline{q}} \right) \underline{q}(1 - x^{\theta}_{\hat{\tau}}) + \left(\frac{1 - e^{(r+\lambda)\hat{\tau}}\underline{q}}{1 - \underline{q}} \right) \int_{\hat{\tau}}^{\infty} e^{-(r+m)(\tau - \hat{\tau})} m(1 - x^{\theta}_{\tau}) \mathrm{d}\tau \right] \\
&\qquad - e^{-r\bar{\tau}}(1 - x^{\theta}_{\bar{\tau}}) \\
&= e^{\lambda\hat{\tau}} \underline{q}(1 - x^{\theta}_{\hat{\tau}}) + \left(e^{-r\hat{\tau}} - e^{\lambda\hat{\tau}}\underline{q}\right) \frac{\lambda\underline{q}}{r+\lambda\underline{q}} (1 - \bar{a}) - e^{-r\bar{\tau}}(1 - x^{\theta}_{\bar{\tau}})
\end{aligned}$$

Replacing

$$1 - x^{\theta}_{\tau} = e^{-\lambda\tau}(1 - \theta) + (1 - e^{-\lambda\tau})(1 - \bar{a})$$

we get that

$$\begin{aligned}\frac{\partial(G_{\text{rand}}-G_{\text{det}})}{\partial U_L} &= \underline{q}(\bar{a}-\theta)+\left(e^{\lambda\hat{\tau}}\underline{q}\frac{r}{r+\lambda\underline{q}}+e^{-r\hat{\tau}}\frac{\lambda\underline{q}}{r+\lambda\underline{q}}\right)(1-\bar{a})-e^{-r\bar{\tau}}(1-x^{\theta}_{\bar{\tau}}) \\ &= \underline{q}(\bar{a}-\theta)+\left(e^{\lambda\hat{\tau}}\underline{q}\frac{r}{r+\lambda\underline{q}}+e^{-r\hat{\tau}}\frac{\lambda\underline{q}}{r+\lambda\underline{q}}\right)(1-\bar{a})-e^{-(r+\lambda)\bar{\tau}}\left(\bar{a}-\theta+e^{\lambda\bar{\tau}}(1-\bar{a})\right) \\ &\leq \left[e^{\lambda\hat{\tau}}\frac{r}{r+\lambda\underline{q}}+e^{-r\hat{\tau}}\frac{\lambda}{r+\lambda\underline{q}}\right]\underline{q}(1-\bar{a})-e^{-r\bar{\tau}}(1-\bar{a}) \\ &\leq 0\end{aligned}$$

where the last inequality follows if $\hat{\tau}\leq\bar{\tau}$ by the same reason as in the case of U_H. Hence, in order to verify single crossing in $(-c,U_L,U_H)$ it is enough to show that $\hat{\tau}\leq\bar{\tau}$. Notice that, for a given continuation value (U_L,U_H), the solution to the deterministic problem, $\bar{\tau}$, is increasing in c, and that whenever $\bar{\tau}<\tau^{\text{bind}}$ (so the IC constraint is slack), the solution to the optimal control problem must be $\bar{\tau}$. Let $c^{\dagger}=\sup\{c\geq 0:\bar{\tau}<\tau^{\text{bind}}\}$, so for any $c<c^{\dagger}$ the solution for a given continuation value (U_L,U_H) is $\bar{\tau}$. On the other hand, for any $c\geq c^{\dagger}$ we have that $\bar{\tau}=\tau^{\text{bind}}\geq\hat{\tau}$, which means that $G_{\text{rand}}-G_{\text{det}}$ satisfies single crossing in $(U_L,U_H,-c)$ which means that $\alpha(U_H,U_L,c)$ is decreasing in U_H,U_L and increasing in c. Moreover, as U_L and U_H are both decreasing in c we can conclude that $\alpha(U_H(c),U_L(c),c)$ is increasing in c, which means that there is $\tilde{c}$ such that for any $c\leq\tilde{c}$ the solution has deterministic monitoring while for any $c>\tilde{c}$ the solution has random monitoring.

Next, we prove that random monitoring dominates deterministic monitoring when k is large enough and when $\bar{a}$ is high or low enough. For this, it is enough to establish that full random monitoring (that is $\hat{\tau}=0$) dominates fully deterministic as this guarantees that some randomization is going to be used in the optimal policy. Before proving the statements in the proposition, we start proving the following Lemma

Lemma 5. *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 derivate 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 $q\in(0,1)$. ☐

Optimality of random monitoring for large k: We compare the payoff of deterministic monitoring with the payoff of full random monitoring (that is $\hat{\tau}=0$) when k converges to its upper bound, $\lambda/(r+\lambda)$ and show that 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

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

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^{det}_{θ}, solves the system of equations

$$\begin{aligned}
B^{\text{det}}_L &= \int_0^{\tau} e^{-rt}u(x^L_t)\mathrm{d}t + e^{-r\tau}(x^L_{\tau}B^{\text{det}}_H + (1-x^L_{\tau})B^{\text{det}}_L) \\
B^{\text{det}}_H &= \int_0^{\tau} e^{-rt}u(x^H_t)\mathrm{d}t + e^{-r\tau}(x^H_{\tau}B^{\text{det}}_H + (1-x^H_{\tau})B^{\text{det}}_L).
\end{aligned}$$

Solving this system we get that the payoff is given by

$$\begin{aligned}
B^{\text{det}}_L &= \frac{\int_0^{\tau} e^{-rt}u(x^L_t)\mathrm{d}t}{1-e^{-r\tau}} + \frac{e^{-r\tau}x^L_{\tau}}{1-e^{-r\tau}(x^H_{\tau}-x^L_{\tau})}\frac{\int_0^{\tau} e^{-rt}(u(x^H_t)-u(x^L_t))\mathrm{d}t}{1-e^{-r\tau}} \\
B^{\text{det}}_H &= \frac{\int_0^{\tau} e^{-rt}u(x^H_t)\mathrm{d}t}{1-e^{-r\tau}} - \frac{e^{-r\tau}(1-x^H_{\tau})}{1-e^{-r\tau}(x^H_{\tau}-x^L_{\tau})}\frac{\int_0^{\tau} e^{-rt}(u(x^H_t)-u(x^L_t))\mathrm{d}t}{1-e^{-r\tau}},
\end{aligned}$$

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

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

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

$$B_L^{\text{rand}} = \int_0^\infty e^{-(r+m^*)t}(u(x_t^L) + m^*(x_t^L B_H^{\text{rand}} + (1 - x_t^L)B_L^{\text{rand}}))\text{d}t$$
$$B_H^{\text{rand}} = \int_0^\infty e^{-(r+m^*)t}(u(x_t^H) + m^*(x_t^H B_H^{\text{rand}} + (1 - x_t^H)B_L^{\text{rand}}))\text{d}t,$$

where

$$B_L^{\text{rand}} = \int_0^\infty e^{-(r+m^*)t}u(x_t^L)\text{d}t + \frac{m^*}{r+m^*}B_L^{\text{rand}} + \frac{m^*\lambda\bar{a}}{(r+m^*)(r+\lambda+m^*)}(B_H^{\text{rand}} - B_L^{\text{rand}})$$
$$B_H^{\text{rand}} = \int_0^\infty e^{-(r+m^*)t}u(x_t^H)\text{d}t + \frac{m^*}{r+m^*}B_L^{\text{rand}} + \left[\frac{m^*}{r+\lambda+m^*} + \frac{m^*\lambda\bar{a}}{(r+m^*)(r+\lambda+m^*)}\right](B_H^{\text{rand}} - B_L^{\text{rand}})$$

From here we get

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

So, replacing in the previous equations

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

We can also write

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

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.

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

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

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

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

Finally, we show that if $\bar{a}$ is large enough, then the upper bound for $\mathcal{G}^{\theta}_{\text{det}}(U)$ is below the lower bound for $\mathcal{G}^{\theta}_{\text{rand}}$. 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 5, we let $\beta \equiv \frac{r}{r+\lambda}$ so we can write

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

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

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

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

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

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

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

and by dominated convergence

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

From Lemma 5 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

Proof of Proposition 4

Proof. We show that the objective function in the model with linear quadratic preferences and brownian shocks can be reduced to the objective function in the model with binary quality and linear quadratic $u(x)$. The objective function in the case linear quadratic case is $u(x_t, \Sigma_t) = x_t - \gamma\Sigma_t$ where

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

On the other hand, if we set $\bar{a} = 1/2$ in the binary case we get that

$$x_t^2 = x_t - \frac{1}{4}(1 - e^{-2\lambda t}).$$

It follows that we can normalize the cost of monitoring and reduce the optimization problem in the linear quadratic case with Brownian quality shocks to the same optimization problem as the one in the binary case with $\bar{a} = 1/2$ and linear quadratic utility function. □

Proof of Proposition 5

Proof. For the first part, notice that the existence of $\tilde{c}^\dagger$ follows directly from Proposition 4a. Next, let's define

$$H^{\text{det}}(\bar{\tau}) \equiv \frac{\int_0^{\bar{\tau}} e^{-r\tau}\gamma\Sigma_\tau \mathrm{d}\tau + e^{-r\bar{\tau}}c}{1 - e^{-r\bar{\tau}}}$$

$$H^{\text{rand}}(\hat{\tau}) \equiv \frac{\int_0^{\hat{\tau}} e^{-r\tau}\gamma\Sigma_\tau \mathrm{d}\tau + e^{-r\hat{\tau}}\left(\frac{1-e^{(r+\lambda)\hat{\tau}}\underline{q}}{1-\underline{q}}\right)\int_{\hat{\tau}}^{\infty} e^{-(r+m^*)(\tau-\hat{\tau})}\gamma\Sigma_\tau \mathrm{d}\tau + \delta(\hat{\tau})c}{1 - \delta(\hat{\tau})}$$

First, we have that

$$H^{\text{det}}_c = \frac{e^{-r\bar{\tau}}}{1 - e^{-r\bar{\tau}}},$$

which means that

$$H^{\text{det}}_{c\bar{\tau}} = -\frac{re^{-r\bar{\tau}}}{(1 - e^{-r\bar{\tau}})^2} < 0$$

which means that $\bar{\tau}$ is increasing in c. Secon, we have that

$$H^{\text{rand}}_c = \frac{\delta(\hat{\tau})}{1 - \delta(\hat{\tau})}$$

so

$$H^{\text{rand}}_{c\hat{\tau}} = \frac{\delta'(\hat{\tau})}{(1-\delta(\hat{\tau}))^2}$$

where

$$\begin{aligned}
\delta'(\hat{\tau}) &= \frac{\underline{q}e^{\lambda\hat{\tau}}}{1-\underline{q}}\left[\lambda\frac{r}{r+m^*} + re^{-(r+\lambda)\hat{\tau}}\left(1-\frac{m^*}{(r+m^*)\underline{q}}\right)\right] \\
&= \frac{r\lambda\underline{q}e^{\lambda\hat{\tau}}}{1-\underline{q}}\left[\frac{1}{r+m^*} - e^{-(r+\lambda)\hat{\tau}}\frac{1-\underline{q}}{r+\lambda\underline{q}}\right] \\
&= re^{\lambda\hat{\tau}}\frac{\lambda\underline{q}}{r+\lambda\underline{q}}\left[1-e^{-(r+\lambda)\hat{\tau}}\right] > 0.
\end{aligned}$$

This means that $H^{\text{rand}}_{c\hat{\tau}} > 0$ so $\hat{\tau}^*$ is decreasing in c and p^* is increasing.

Next, let's consider the comparative statics with respect to k. First, notice that $H^{\text{det}}(\bar{\tau}, c)$ is independent of k and that the cost of effort becomes relevant only once the incentive compatibility constraint is binding. Next, we consider the maximization of $H^{\text{rand}}(\hat{\tau})$. Because k enters into the maximization problem only through $\underline{q}$ it is enough to show that $\hat{\tau}$ is decreasing in $\underline{q}$. After some lengthy computations, we have that $H_{\hat{\tau}} = 0$ if and only if

$$g(\hat{\tau},\underline{q}) \equiv r(r+\lambda(2-\underline{q}))\left(2\tilde{c}\lambda(r+2\lambda)-1\right) + 2\underline{q}r(r+2\lambda)e^{-\lambda\hat{\tau}} - (r+2\lambda)(r+\lambda\underline{q})e^{-2\lambda\hat{\tau}} + 2\lambda^2\underline{q}e^{-(r+2\lambda)\hat{\tau}} = 0$$

Let $z \equiv e^{-(r+2\lambda)\hat{\tau}}$ and write

$$g(z,\underline{q}) \equiv r(r+\lambda(2-\underline{q}))\left(2\tilde{c}\lambda(r+2\lambda)-1\right) + 2\underline{q}r(r+2\lambda)z^{\frac{\lambda}{r+2\lambda}} - (r+2\lambda)(r+\lambda\underline{q})z^{\frac{2\lambda}{r+2\lambda}} + 2\lambda^2\underline{q}z = 0. \quad (57)$$

The incentive compatibility constraint requires that $\hat{\tau} \leq \tau^{\text{bind}}$, which means that

$$z \geq \underline{q}^{\frac{r+2\lambda}{r+\lambda}}.$$

Hence, we get that

$$\begin{aligned}
g_z(z,\underline{q}) &= 2\underline{q}r\lambda z^{-\frac{r+\lambda}{r+2\lambda}} - 2\lambda(r+\lambda\underline{q})z^{-\frac{r}{r+2\lambda}} + 2\lambda^2\underline{q} \\
&\leq 2\lambda\left(r+\lambda\underline{q}\right)\left(1-z^{-\frac{r}{r+2\lambda}}\right) \\
&\leq 0.
\end{aligned}$$

Next, we verify that $g_{\underline{q}}(z,\underline{q}) > 0$. Notice that we can write

$$g(z,\underline{q}) = g_0(z) + g_1(z)\underline{q}.$$

Hence, if $g(\hat{z},\underline{q}) = 0$, then it must be the case that

$$g_1(\hat{z})\underline{q} = -g_0(\hat{z})$$

which means that it is enough to show that $g_0(\hat{z}) < 0$ evaluated at the solution. From equation (57) we have that

$$g_0(z) = r(r+2\lambda)\left(2\tilde{c}\lambda(r+2\lambda)-1\right) - (r+2\lambda)rz^{\frac{2\lambda}{r+2\lambda}},$$

where

$$r(r+2\lambda)\left(2\tilde{c}\lambda(r+2\lambda)-1\right) - (r+2\lambda)rz^{\frac{2\lambda}{r+2\lambda}} \leq r(r+2\lambda)\left(2\tilde{c}\lambda(r+2\lambda)-1\right)$$

so the inequality follows from the sufficient condition

$$\tilde{c} \leq \frac{1}{2\lambda(r+2\lambda)}.$$

Finally, we verify that the optimal policy in the i.i.d limit is random. The optimal policy is random if $g(1,\underline{q}) > 0$, which means that if

$$\lambda\tilde{c} \geq \frac{1-\underline{q}^{\frac{\lambda}{r+\lambda}}}{2(r+2\lambda)} + \frac{1-\underline{q}^{\frac{r+2\lambda}{r+\lambda}}}{2r(2-\underline{q})} - \frac{2\underline{q}^{\frac{\lambda}{r+\lambda}} + 2\underline{q}^2 - 4\underline{q} - 1}{2(\underline{q}-2)(r+\lambda(2-\underline{q}))}$$

then we have constant monitoring starting at time zero.

On the other hand, there is random monitoring only if

$$\lambda\tilde{c} > \frac{1-\underline{q}^{\frac{\lambda}{r+\lambda}}}{2(r+2\lambda)} + \frac{\underline{q}^{\frac{2\lambda}{r+\lambda}} - \underline{q}^{\frac{r+2\lambda}{r+\lambda}}}{2r(2-\underline{q})} + \frac{\underline{q}^{\frac{2\lambda}{r+\lambda}} - 2(1-\underline{q})^2\underline{q}^{\frac{\lambda}{r+\lambda}}}{2(\underline{q}-2)(r+\lambda(2-\underline{q}))}$$

In order to keep the steady state distribution constant as we take λ to infinity, we consider the case in which σ^2/λ is constant. As shocks are more transitory, monitoring becomes less informative so we consider the limit when $\tilde{c}_\lambda \equiv \lambda\tilde{c} > 0$ is constant to adjust for this lower informativeness. Thus we can write

$$\tilde{c}_\lambda > \frac{1-\underline{q}^{\frac{\lambda}{r+\lambda}}}{2(r+2\lambda)} + \frac{\underline{q}^{\frac{2\lambda}{r+\lambda}} - \underline{q}^{\frac{r+2\lambda}{r+\lambda}}}{2r(2-\underline{q})} + \frac{\underline{q}^{\frac{2\lambda}{r+\lambda}} - 2(1-\underline{q})^2\underline{q}^{\frac{\lambda}{r+\lambda}}}{2(\underline{q}-2)(r+\lambda(2-\underline{q}))}.$$

In the limit, $\underline{q} \to k$, which means that the right hand side above converges to zero, so random monitoring is optimal. □

D Exogenous News

In this appendix, we consider the model with exogenous news and $\mu_L \neq \mu_H$. We characterize how the relation between monitoring and market beliefs depends on the nature of the news process. We find that when the news process conveys negative states faster than positive states, then monitoring tends to intensify when the firm's reputation is low because then the moral hazard issue is more severe. By contrast, when the news process conveys negative states faster than positive states, then monitoring intensifies in good times since effort incentives are weaker when the firm's reputation is high.

D.1 Incentive Compatibility and the Principal's Problem with News

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

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

where we use the fact that if $a_t = \bar{a}$ for any $t \geq T_n$ then, given $\theta_{T_n} = \theta$, the continuation payoff is $\Pi_0^\theta = \Pi(\theta)$ (recall that $\Pi(\theta)$ is given by (5)). From here it follows that full effort $a_\tau = \bar{a}$ is incentive compatible if and only if:[26]

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

The evolution of D_τ can be derived (analogously to what we have done before) to be

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

with a boundary condition $D_{\bar{\tau}} = \Delta \equiv \Pi(H) - \Pi(L) = 1/(r+\lambda)$.

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 facing the principal is the same going forward. Hence, we can write the problem recursively using as state variables both 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

$$\begin{aligned} \mathscr{G}^\theta(\mathbf{U}) = &\sup_{\bar{\tau}, m_\tau, \Pi_0^{\neg\theta}} \int_0^{\bar{\tau}} e^{-r\tau - M_{\tau^-}} \left(x_\tau^\theta + \mu_H x_\tau^\theta U_H + \mu_L (1 - x_\tau^\theta) U_L + m_\tau \mathcal{M}(\mathbf{U}, x_\tau) \right) \mathrm{d}\tau \\ &+ e^{-r\bar{\tau} - M_{\bar{\tau}}} \mathcal{M}(\mathbf{U}, x_{\bar{\tau}}) \end{aligned}$$

subject to

$$\begin{aligned} \dot{\Pi}_\tau^H &= (r + \mu_H + m_\tau)\Pi_\tau^H - x_\tau + k\bar{a} + \lambda(1-\bar{a})(\Pi_\tau^H - \Pi_\tau^L) - (\mu_H + m_\tau)\Pi(H), \ \Pi_{\bar{\tau}}^H = \Pi(H) \\ \dot{\Pi}_\tau^L &= (r + \mu_L + m_\tau)\Pi_\tau^L - x_\tau + k\bar{a} - \lambda\bar{a}(\Pi_\tau^H - \Pi_\tau^L) - (\mu_L + m_\tau)\Pi(L), \ \Pi_{\bar{\tau}}^L = \Pi(L) \\ \Pi_0^\theta &= \Pi(\theta) \\ \frac{k}{\lambda} &\leq \Pi_\tau^H - \Pi_\tau^L, \ \forall \tau \in [0, \bar{\tau}] \\ 0 &\leq m_\tau. \end{aligned}$$

Note that in the previous formulation, the continuation payoff given the counterfactual type $\neg\theta$ (if $\theta = H$ then $\neg\theta = L$ and vice versa), which we denote by $\Pi_0^{\neg\theta}$, is not given by $\Pi(\neg\theta)$. The solution of this problem critically depends on the intensity of bad versus good news arrivals. We first consider the symmetric case.

We consider the asymmetric case, $\mu_H \neq \mu_L$, so that the intensity of news arrival depends on firm's quality. Such asymmetry seems natural: in some industries and under some market conditions, good news tend to be revealed faster than bad news, among other things because firms themselves may delay the release

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

of bad news. Sometimes, bad news tend to be revealed faster than good news, perhaps because news agencies and TV broadcasts face stronger demand for bad news stories.

The main question we address here is how monitoring rates are affected by reputation when exogenous news are asymmetric. We do not solve the full problem here, and instead we focus on the case in which the principal's preferences are linear. Based on our previous analysis, it is natural to conjecture that the optimal policy has 1) no atoms in the distribution of monitoring (in particular, $\bar{\tau} = \infty$), and 2) the monitoring rate is positive (i.e., $m_\tau > 0$) only if the incentive compatibility constraint is binding, that is if $\Pi^H_\tau - \Pi^L_\tau = k/\lambda$. We can use the maximum principle to verify if our conjectured policy is optimal. We relegate a detailed discussion of the optimality conditions to the appendix.

Given this monitoring policy, we can follow the same steps as before, and derive the monitoring rate using the incentive compatibility constraint: $(\dot{\Pi}^H_\tau - \dot{\Pi}^L_\tau) = 0$ and $\Pi^H_\tau - \Pi^L_\tau = k/\lambda$. These conditions are necessary for the incentive compatibility constraints to bind at all times. They imply:

$$m_\tau = \alpha + \beta \Pi^L_\tau, \tag{58}$$

where

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

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 focus on the simplest case with parameters such that the optimal policy has $m_\tau > 0$ for all $\tau \geq 0$; this case illustrates the effect of introducing exogenous news on the optimal monitoring policy at the lowest cost of technical complications.[27] Using the relation $\Pi^H_\tau = \Pi^L_\tau + D_\tau = \Pi^L_\tau + k/\lambda$ and the monitoring rate (58) we write the evolution of the low quality firm continuation value as

$$\dot{\Pi}^L_\tau = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi^L_\tau + \beta(\Pi^L_\tau)^2 - x_\tau. \tag{59}$$

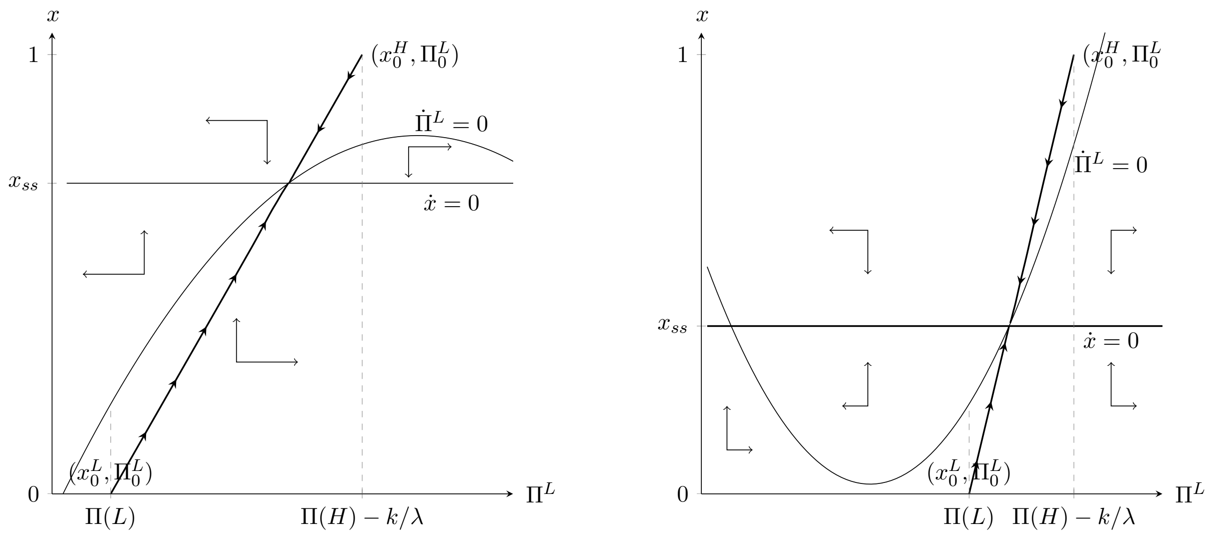
If $\theta_0 = L$ then the initial condition is $\Pi^L_0 = \Pi(L)$. If $\theta_0 = H$ (and the incentive compatibility is binding) the initial condition is $\Pi^L_0 = \Pi(H) - k/\lambda$.[28] We can analyze the evolution of monitoring by studying the phase diagram in the space (x_τ, Π^L_τ) in Figure 7.

Using the ODE for Π^L_τ in equation (59) we get a quadratic equation for the steady state:

$$0 = -(\mu_L + \alpha)\Pi(L) + (r + \mu_L + \alpha - \beta\Pi(L))\Pi^L + \beta(\Pi^L)^2 - x. \tag{60}$$

[27]Such policy is optimal when the rates of exogenous news arrivals are low. When those rates 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.

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



(a) Case with bad news

(b) Case with good news

Figure 7: Phase diagram. The (x_τ, Π_τ^L) system has two steady states. In each case, one of the steady states is a saddle point. If the optimal solution is such that $m_\tau > 0$ 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_τ^L must be monotone between news arrivals. This immediately implies that the evolution of monitoring between news is monotone as well.

This quadratic equation has two solutions. We show that in the good news case only the largest solution is consistent with a positive monitoring rate, while in the bad news only the smallest one is consistent with a positive monitoring rate. So if the solution has positive monitoring rate at all times, then the solution must correspond to the saddle point trajectory in the phase diagram in Figure 7.

From inspection of the phase diagram it is clear that Π_τ^L is monotone: it starts decreasing after good news and starts increasing after bad news. This implies the dynamics of optimal monitoring that are described in Figure 6. In the bad news case, monitoring increases after (bad) news. The opposite is optimal in the good news case. As previously mentioned, this is driven by the dynamics of reputational incentives. In the bad news case, incentives weaken as reputation goes down. As Board and Meyer-ter-Vehn (2013) point out, a high reputation firm has more to lose from a collapse in its reputation following a breakdown than a low reputation firm. Hence, inspections are most needed 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, inspections are thus most needed when reputation is high. Accordingly monitoring complements exogenous news, being used when exogenous news are ineffective at providing incentives. We still need to verify that: (1) the optimal monitoring policy is optimal, and (2) show that the dynamics of the firm's continuation value satisfy the monotonicity properties in Figure 7. We consider the optimality conditions for the optimal policy in Section D.2 and study the steady states of the firm's continuation payoffs in Section D.3.

D.2 Necessary Conditions with Asymmetric News

The Hamiltonian for the optimal control problem is

$$\begin{aligned}\mathcal{H}(\Pi^L_\tau, \Pi^H_\tau, \zeta_\tau, \nu^L_\tau, \nu^H_\tau, \psi_\tau, m_\tau, \tau) = {} & \zeta_\tau((r+m_\tau)U^\theta_\tau - x^\theta_\tau - \mu_H x^\theta_t U_H - \mu_L(1-x^\theta_\tau)U_L - m_\tau \mathcal{M}(\mathbf{U}, x_\tau) \\ & + \psi_\tau(\Pi^H_\tau - \Pi^L_\tau - k/\lambda) + \nu^H_\tau\big((r+\mu_H+m_\tau)\Pi^H_\tau - x_\tau + k\bar{a} + \lambda(1-\bar{a})(\Pi^H_\tau - \Pi^L_\tau) \\ & - (\mu_H + m_\tau)\Pi(H)\big) + \nu^L_\tau\big((r+\mu_L+m_\tau)\Pi^L_\tau - x_\tau + k\bar{a} - \lambda\bar{a}(\Pi^H_\tau - \Pi^L_\tau) \\ & - (\mu_L + m_\tau)\Pi(L)\big)\end{aligned}$$

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

$$\begin{aligned}\dot{\nu}^H_\tau &= -(\mu_H + \lambda(1-\bar{a}))\nu^H_\tau - \psi_\tau + \lambda\bar{a}\nu^L_\tau \\ \dot{\nu}^L_\tau &= -(\mu_L + \lambda\bar{a})\nu^L_\tau + \psi_\tau + \lambda(1-\bar{a})\nu^H_\tau.\end{aligned}$$

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

$$S(\tau) = \mathcal{M}(\mathbf{U}, x_\tau) + \nu^H_\tau(\Pi^H_\tau - \Pi(H)) + \nu^L_\tau(\Pi^L_\tau - \Pi(L)) - U^\theta_\tau.$$

We pin-down the boundary condition for the co-state variables ν^θ_τ 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

$$0 = \mathcal{M}(\mathbf{U}, \theta) - U_\theta + \nu^H_0(\Pi^H_0 - \Pi(H)) + \nu^L_0(\Pi^L_0 - \Pi(L)).$$

If the incentive compatibility constraint is binding at time zero, so $\Pi^H_0 - \Pi^L_0 = k/\lambda$, then when $\theta_0 = L$ and $m_0 > 0$ the initial value of the co-state variable ν^H_0 is

$$c = -\nu^H_0\left(\frac{1}{r+\lambda} - \frac{k}{\lambda}\right).$$

The initial value of the co-state variable ν^L_0 is determined by the transversality condition $\lim_{\tau\to\infty}\nu^L_\tau = \nu^L_{ss}$. If the incentive compatibility constraint at time zero were slack (that is $m_0 = 0$) then the initial value would be $\nu^H_0 = 0$. The determination of ν^L_0 is more complicated in this latter case as ν^L_τ can jump at the junction time τ^m in which the IC constraint becomes binding. Similarly, if $\theta = H$ then we have that ν^L_0 is given by

$$c = \nu^L_0\left(\frac{1}{r+\lambda} - \frac{k}{\lambda}\right)$$

while ν^H_0 is determined by the transversality condition $\lim_{\tau\to\infty}\nu^H_\tau = \nu^H_{ss}$. 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

$$\begin{aligned}\psi_\tau\big((\Pi^H_\tau - \Pi^L_\tau) - (\Pi(H) - \Pi(L))\big) = {} & \dot{x}^\theta_\tau(U_H - U_L) - \dot{U}^\theta_\tau + (-(\mu_H + \lambda(1-\bar{a}))\nu^H_\tau + \lambda\bar{a}\nu^L_\tau)(\Pi^H_\tau - \Pi(H)) + \nu^H_\tau\dot{\Pi}^H_\tau \\ & + (-(\mu_L + \lambda\bar{a})\nu^L_\tau + \lambda(1-\bar{a})\nu^H_\tau)(\Pi^L_\tau - \Pi(L)) + \nu^L_\tau\dot{\Pi}^L_\tau\end{aligned}$$

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

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

A necessary 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 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. The Hamiltonian in our problem is not concave, so traditional theorems on the sufficiency of the maximum principle do not apply. However, our problem is a special case of the generalized linear control processes considered by Lansdowne (1970), for which he proves the sufficiency of the maximum principle. The results in Lansdowne (1970) do not apply directly to our problem due to the presence of a state constraint; however, because the state constraint in our problem is linear, his sufficiency result can be extended to our setting.

D.3 Monotonicity of Monitoring Policy with Asymmetric News

Proof. Looking at the phase diagram in Figure 7 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. Hence, we only need to rule out that in the optimal policy the continuation values converge to the stable steady state. We show this by verifying 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 (60), respectively. We show next that only one of these roots is consistent with $m_\tau \geq 0$.

Claim 2 (Bad News)**.** *If $\mu_L > \mu_H$ then*

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

Given that we are in the bad news case, $m_\tau > 0$ 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 3 (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

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

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 $\Pi_-^L > 0$. □