Review of Economic Studies (2020) , 1–0 doi:10.1093/restud/rdaa
© The Author(s) 2020. Published by Oxford University Press on behalf of The Review of Economic Studies Limited.

Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring

FELIPE VARAS
Duke University, Fuqua School of Business
IVÁN MARINOVIC
Stanford University, GSB
ANDRZEJ SKRZYPACZ
Stanford University, GSB

We study the design of monitoring in dynamic settings with moral hazard. An agent (e.g. a firm) benefits from reputation for quality, and a principal (e.g. a regulator) can learn the agent's quality via costly inspections. Monitoring plays two roles: an incentive role, because outcomes of inspections affect agent's reputation, and an informational role because the principal directly values the information. We characterize the optimal monitoring policy inducing full effort. When information is the principal's main concern, optimal monitoring is deterministic with periodic reviews. When incentive provision is the main concern, optimal monitoring is random with a constant hazard rate.

JEL Codes: C73, D82, D83, D84.

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 are to be tested for compliance? What about the timing of internal audits to measure divisional performance to allocate capital within organizations?

Monitoring is fundamental for the implementation of any regulation. It is essential for enforcement and, ultimately, for resource allocation. However, monitoring is costly in practice, and according to the OECD (2014), "regulators in many countries are increasingly under pressure to do 'more with less.' A 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 many cases, monitoring outcomes are public and can thus have a significant impact on a firm's reputation. Regulators can exploit this reputational concern when designing their monitoring strategies to strengthen firms' incentives to provide quality. In essence, monitoring is a form of information acquisition. The information that a regulator collects via monitoring serves multiple purposes: not only it provides valuable information that can help the regulator improve the allocation of resources in the economy, but it also is an important incentive device when agents are concerned about their reputation. As such, monitoring is often a substitute to monetary rewards.[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 inspections 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.

A central prediction of our analysis is that periodic reviews and random Inspections are optimal in different environments since they serve distinct purposes.

Random inspections are efficient in providing incentives, and hence are optimal in circumstances where moral hazard considerations are important. The severity of moral hazard depends on the difficulty of providing quality and its persistence. These factors affect the prevalence of random inspections in practice. For example, restaurant hygiene monitoring programs typically rely on random inspections[2]

Deterministic reviews, by contrast, are the most efficient when an essential reason for inspections is learning to improve decisions and incentives to provide quality are relatively less important. This seems to be a good description in case of safety inspections where learning about a safety hazard is crucial for the monitor to prevent a disaster.

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 trade-offs in such markets.

2. Or else a restaurant could improve its hygiene conditions merely to pass a pre-announced inspection but pay less attention to hygiene when inspections are not expected. Jin and Leslie (2003) provides evidence of the positive impact of random inspection on restaurant hygiene.

Deterministic reviews are thus common for safety monitoring systems such as those managed by the Federal Aviation Administration[3] Our results indicate that deterministic reviews are particularly useful when the monitored agents have a relatively mild moral hazard problem (for example, because they care directly about maintaining quality) and when the direct value of information is high.

A different example of the variety of timing of tests is familiar from the educational institutions, where professors use both random quizzes and pre-announced tests. Consistent with our model, tests that could be easy to prepare for on a short notice (but with such preparation being not very productive) are often performed as "random quizzes" or "cold calls." On the other hand, tests that require a deeper understanding of the material and hence more sustained studying effort but are important for evaluating students and for giving them guidance about what other choices to make, are often scheduled.

In this paper, we study a model with investment in quality and costly inspections. The objective is to identify the 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 combines the previous two features that we commonly observe in practice. It relies on deterministic reviews to periodically acquire information and combine it with random inspections to provide incentives at the lowest possible cost. In principle, we would expect the optimal policy to be complex, fine-tuning the probability of monitoring over time. However, we show that the optimal policy is simple, and can be easily 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.

A policy with a constant hazard rate minimizes the cost of inspections subject to the incentive compatibility constraints. However, when learning is important, random policies are inefficient because inspections might be performed when there is little uncertainty about the agent's type, in which case learning is not as valuable yet. Then, the benefit of delaying inspections until the value of learning is large enough is greater than the associated increase in cost caused by the departure from the constant hazard rate policy. However, over time, in the absence of inspections, the value of learning stabilizes, and the trade-off is dominated by cost minimization. This explains why the optimal policy eventually shifts towards a constant hazard rate.

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 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.

3. The Federal Aviation Administration regulates all aspects of civil aviation in that nation. The FAA mandates periodic aircraft inspections. For example, 'A checks' happen every 400-600 flight hours (see FAR 91.409b).

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 the 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 by considering the case in which 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.

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. However, our general intuition is that these additional considerations (such as dynamic punishments, or direct monetary incentives) that make the moral hazard less severe or the direct value of information higher, should lead the optimal policy to favor deterministic monitoring over randomization.

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 the 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.[4]

Our paper is closely related to previous work by Lazear (2006) and Eeckhout et al. (2010), who study the optimal allocation of monitoring resources in static settings and without reputation concerns. In particular, 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.

4. 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 the 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 the properties of good monitoring programs.

Another related literature, initiated by Becker (1968), looks at the deterrence effect of policing and enforcement and the optimal monitoring policy to deter criminal behavior in static settings.[5] In a dynamic context, 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. Our work is also related to the literature looking at inspections games with exogenous fines, in particular, the work by Solan and Zhao (2019), who look at a repeated inspection game in which there is a capacity constraint on the number of firms that the regulator can inspect in a given period, and firms are fined if they are found in violation of the rules. In their setting, it is not possible to provide full compliance, and the objective of the regulator is to maximize compliance over time.

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 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.[6] At 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 the design of optimal monitoring when some information comes exogenously.

Finally, our paper is somewhat related to the literature that has explored the 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

We start by describing the general setting. Then, we provide a discussion of potential applications and some specific examples of how the model can be micro-founded to study them.

Agents, Technology, and Effort: There are two players: a principal and a firm/agent. 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

5. See for example , Polinsky and Shavell (1984), Reinganum and Wilde (1985), Mookherjee and Png (1989), Bassetto and Phelan (2008), Bond and Hagerty (2010).

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

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.)

A monitoring policy specifies an increasing sequence of inspections $(T_n)_{n \geq 1}$ times.[7] 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 $\mathbb{F}^P = (\mathcal{F}_t^P)_{t \geq 0}$ 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.[8] 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 $x_t \equiv E^{\tilde{a}}(\theta_t | \mathcal{F}_t^P)$ where the expectation is taken with respect to the measure induced by the conjectured effort, $\tilde{a}$. In words, reputation is the market's belief about the firm's current quality. It evolves based on the market's conjecture about 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.[9] The marginal cost of effort is k, hence the firm's expected payoff at time t is

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

7. 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.

8. Notice that the principal filtration includes the complete history of inspection outcomes and dates.

9. 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.

In the absence of asymmetric information, the maximal 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 twice continuously differentiable 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

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

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.

Incentive Compatibility and Optimal Policies. We seek to characterize monitoring policies that maximize the principal's payoff among those inducing full effort.[10] 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 2.1. *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:*[11]

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.

10. 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 at the end of the paper.

11. 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.

Definition 2.2. *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.

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.[12] 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 3.2).

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 the 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.[13] 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 a manager's reputation has an important role in 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 a 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

12. If the principal payoff is $\tilde{u}(\theta_t,x_t)$ then the expected payoff is $u(x_t)=x_t\tilde{u}(H,x_t)+(1-x_t)\tilde{u}(L,x_t)$.
13. See Motta (2003) for a capital budgeting model driven by career concerns along these lines.

inspecting a certain fraction of them in every period to approximate the optimal random policy we describe. The non-commitment case is beyond the scope of this paper.[14]

2.1. *Examples*

To further motivate the model, we turn to three applications. They illustrate how the firm and principal payoffs can be micro-founded.

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

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

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

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

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:

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

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

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, his 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 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.[15] The principal's preferences are linear if $q_1 = q$

15. 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.

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.[16] 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.[17] 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 headquarter 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 $\iota_t = x_t$.[18] 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 about the current productivity once the cash-flows arrive. We study the possibility of exogenous news arrivals in the appendix.

16. 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.

17. 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.

18. 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.

3. INCENTIVE COMPATIBLE POLICIES

In the next section, we derive optimal monitoring policies. To that end, in this section, we characterize necessary and sufficient conditions for a monitoring policy to be incentive compatible.

Consider the firm's continuation payoff under full effort at time T_{n+1}:

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

This expression represents the expected present value of the firm's future revenues net of effort costs. A key observation is that the law of iterated expectations and the Markov nature of the quality process imply that $E^{\bar{a}}[x_t|\mathcal{F}_{T_{n+1}}] = E^{\bar{a}}[\theta_t|\theta_{T_{n+1}}]$, and moreover:

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

Therefore, under any incentive-compatible monitoring policy, if the firm is inspected at T_{n+1} and has quality $\theta_{T_{n+1}} = \theta$, then its continuation payoff is:

$$\Pi(\theta) \equiv \frac{\bar{a}}{r} + \frac{\theta - \bar{a}}{r+\lambda} - \frac{\bar{a}k}{r}. \tag{3.1}$$

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. 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.

Consider the firm's effort incentives. Effort may affect the firm's payoff by changing the outcome of future inspections. The expected marginal benefit of exerting effort over an interval of size $\mathrm{d}t$ is

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

where the expectation is over the next inspection time, T_{n+1}. 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 today 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 and our derivation of $\Pi(H) - \Pi(L)$, we can express the necessary and sufficient condition for full effort to be incentive compatible as follows.

Proposition 3.1. *Full effort is incentive compatible if and only if for all $n \geq 0$,*

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

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

This condition states that for a monitoring policy to be incentive compatible next expected discounted inspection date $E\left[e^{-(r+\lambda)(T_{n+1}-t)}\right]$ has to be sufficiently high. Future monitoring affects incentives today because effort has a persistent effect on quality, so shirking today can lead to a persistent drop in quality that can be detected by the principal in the near future. Therefore, what matters for incentives at a given point in time is not just the monitoring intensity at that point but the cumulative discounted likelihood of monitoring in the near future. Future inspections are discounted both by r and the switching intensity λ because effort today matters insofar as quality is persistent.

Finally, 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. Therefore, the optimal monitoring policy is the same whether 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 marginal cost of effort are independent of θ_t.

Remark 3.2. *Proposition 3.1 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*

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

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

4. OPTIMAL MONITORING POLICY

We now describe an optimal monitoring policy that induces full effort. To do so, we first optimize over the distribution of the first inspection time, taking as given some continuation payoffs for the principal. We then discuss how one can obtain via a recursive computation the actual continuation payoffs under the optimal policy. As we show, the qualitative features of the optimal policy do not depend on the continuation payoffs. The optimal policy belongs to a two-dimensional class (with one parameter for each possible outcome of the last inspection), and that greatly simplifies the computation of continuation payoffs under the optimal policy. The next theorem provides a general characterization of the optimal monitoring policy, and it is the main result of this paper.

Theorem *4.1 (Optimal Monitoring Policy)* *Let τ^{bind} be the largest time such that deterministic monitoring at that time is incentive compatible, which is given by $e^{-(r+\lambda)\tau^{bind}} = \underline{q}$, and let $F_\theta^* : \mathbb{R}_+ \times \{L, H\} \to [0,1]$ be an optimal policy following and inspection in which $\theta_{T_n} = \theta$. An optimal policy 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 deterministic 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

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

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 4.1 states that the optimal policy belongs to the following simple family of monitoring policies. For a given outcome in the last inspection, there is a time $\hat{\tau}_\theta^*$ 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 $\hat{\tau}_\theta^*$: the timing of the next inspection is deterministic, and the incentive constraints bind at most right after an inspection (so that $\tau_\theta^* \le \tau^{\text{bind}}$). There is a special case in which $\hat{\tau}_\theta^* = 0$ so the policy is fully random and requires monitoring at a constant hazard rate. In general, the optimal random policy has an atom at $\hat{\tau}_\theta^*$ such that the incentive constraints hold exactly at $\tau = 0$ (and then are slack till $\hat{\tau}_\theta^*$ 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.

Real-world inspection schemes share several qualitative aspects of the optimal policy. In many applications, monitoring systems feature random inspections. For example, restaurant hygiene inspections in the U.S are random and have a reputational impact because the outcome of the inspections is disclosed to the public (see Jin and Leslie (2003)). In the U.S., firms are inspected for health and safety by the Occupational Safety and Health Administration. Safety inspections also happen randomly every year. As Levine et al. (2012) demonstrate empirically, random inspections seem to play an important incentive role.[19] Also, random quality inspections are widely used in Europe to evaluate schools.

However, not all monitoring systems feature purely random inspections. Many monitoring applications exhibit regular periodic inspections; for example, the Public Company Accounting Oversight Board (PCAOB) monitors audit firms annually or triennially, depending on their size. The PCAOB inspections are noisy, as it evaluates a random sample of the auditor's past audit engagements. In the U.K., schools are supposed to be inspected by Office for Standards in Education, Children Services and Skills (Ofsted) at least once over a four-year cycle, so monitoring also exhibits a deterministic component.

We show that a policy with a constant hazard rate minimizes the cost of inspections subject to the incentive compatibility constraints. However, random policies are inefficient because inspections might not be performed when information is most valued. 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 the departure from the constant hazard rate policy. However, over time, the value of learning grows slower and slower. For example, as beliefs get closer to the steady-state, they change very slowly, and the trade-off is dominated by cost minimization. This 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.

In the case of our previous applications, our model predicts that when learning is particularly valuable to the regulator (for example, when health and safety is a concern, as in the OSHA application), then optimal monitoring systems feature frequent deterministic inspections. That is, even if incentive issues are not particularly relevant (e.g., effort is costless), learning about hazards to prevent accidents is relevant to the regulator. Similarly, in the case of school quality inspections, learning is important as it allows the regulator to steer children to the best schools. The importance of learning may explain perhaps why in the U.K. all schools incorporate a periodic, deterministic component, to its random inspection scheme.

Some of our applications feature inspection frequencies that depend on the outcome of the last inspection (i.e., past performance). For instance, in the U.K. since 2009, the frequency of school inspections varies according to each school's past performance, whereby inadequate schools are inspected every two years, and outstanding schools are inspected every five years. Our model can generate such asymmetric frequency if u is relatively convex at the bottom and linear at the top of the distribution. In other words, if the information is particularly valuable when schools perform badly, then the optimal monitoring system will feature a higher frequency when the outcome of the inspection is bad.

4.1. *Analysis Principal Problem*

Having discussed the shape of the optimal policy and its implementations, we proceed to analyze the principal problem and to provide a sketch of the derivation of the policy in Theorem 4.1. The full verification arguments are provided in the appendix. The first

19. In a natural field experiment, they found that companies subject to random OSHA inspections showed a 9.4 percent decrease in injury rates compared with uninspected firms.

step in the analysis is to express the principal's problem as a linear program. According to Proposition 3.1, incentive compatibility constraint at time t depends only on the distribution of the time to the next inspection, $\tau_{n+1} \equiv T_{n+1} - t$ 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 principal's 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.[20] Let $V_\theta(\tau|\mathbf{U})$ be the principal payoff under the full effort by the firm, conditional on monitoring at time τ. It depends on the last inspection result θ and the continuation payoffs. It is given by

$$V_\theta(\tau|\mathbf{U}) = \int_0^\tau e^{-rs} u(x_s^\theta) \mathrm{d}s + e^{-r\tau} \mathcal{M}(\mathbf{U}, x_\tau^\theta), \tag{4.2}$$

where $x_\tau^\theta \equiv \theta e^{-\lambda\tau} + \bar{a}\left(1 - e^{-\lambda\tau}\right)$ is the expected quality at τ given starting quality θ. From here, we can write the principal problem as choosing the distribution of the next inspection time, F, subject to the incentive compatibility constraints:

$$\begin{cases} \mathscr{G}^\theta(\mathbf{U}) = \max_F \int_0^\infty V_\theta(\tau|\mathbf{U}) \mathrm{d}F(\tau) \\ \text{subject to} \\ \int_\tau^\infty e^{-(r+\lambda)(s-\tau)} \frac{\mathrm{d}F(s)}{1-F(\tau-)} \geq \underline{q} \quad \forall \tau \geq 0, \end{cases} \tag{4.3}$$

The principal payoff under the optimal policy is then given by the fixed point $\mathscr{G}^\theta(\mathbf{U}) = \mathbf{U}$, $\theta \in \{L, H\}$.

From now on, we omit the dependence of $V_\theta(\tau|\mathbf{U})$ on θ and $\mathbf{U}$ to simplify the notation and we just write $V(\tau)$, understanding that it represents the principal payoff for a given state θ and continuation payoff $\mathbf{U}$. In order to simplify the principal problem in (4.3), we can replace the incentive compatibility constraint in (4.3) by

$$\int_\tau^\infty e^{-(r+\lambda)(s-\tau)} \mathrm{d}F(s) \geq \underline{q}(1 - F(\tau-)), \forall \tau \geq 0.$$

Notice that we have added extra constraints for some values of τ for which $F(\tau) = 1$; however, we can include them without loss of generality as they are trivially satisfied by any feasible policy. We can now write the principal problem in (4.3) as

$$\begin{cases} \max_F \int_0^\infty V(\tau) \mathrm{d}F(\tau) \\ \text{subject to} \\ \int_\tau^\infty \left(e^{-(r+\lambda)(s-\tau)} - \underline{q}\right) \mathrm{d}F(s) \geq 0, \forall \tau \geq 0 \\ \int_0^\infty \mathrm{d}F(\tau) = 1. \end{cases} \tag{4.4}$$

The advantage of the formulation in (4.4) over the one in (4.3) is that the former is a linear programming problem.

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

To develop some intuition of the shape of the optimal policy, consider the problem of the principal when only the incentive compatibility constraint at time 0 is relevant. Ignoring the second constraint in (4.4) we get the Lagrangian

$$\mathcal{L}=\int_0^\infty V(\tau)\mathrm{d}F(\tau)+\mu\left(\int_0^\infty e^{-(r+\lambda)\tau}\mathrm{d}F(\tau)-\underline{q}\right),$$

where $\mu\geq 0$ is the Lagrange multiplier of the time zero incentive compatibility constraint. For some arbitrary time τ, consider a policy that satisfies both constraints, and a perturbation to this policy that increases $\mathrm{d}F(\tau+\mathrm{d}\tau)$ and reduces $\mathrm{d}F(\tau)$ by the same amount, so that the total probability stays constant (so the second constraint is satisfied). The marginal effect of this perturbation on the Lagrangian is

$$V(\tau+\mathrm{d}\tau)-V(\tau)+\mu\left(e^{-(r+\lambda)(\tau+\mathrm{d}\tau)}-e^{-(r+\lambda)\tau}\right)$$

Dividing by $\mathrm{d}t$, taking the limit, and multiplying by $e^{(r+\lambda)\tau}$ we get

$$\underbrace{e^{(r+\lambda)\tau}V'(\tau)}_{\equiv h(\tau)}-\mu(r+\lambda) \tag{4.5}$$

At the time $\hat{\tau}$ of an atom this derivative is zero since we do not want to postpone the atom anymore, which we could do by reducing the atom at $\hat{\tau}$ and increasing the probability of monitoring $\mathrm{d}\tau$ later. We show in the formal proof in the appendix that the function $h(\tau)$ is quasi-convex and decreasing to the left of $\hat{\tau}$ (which means that (4.5) is positive for all $\tau<\hat{\tau}$, so it is optimal to not have any probability of monitoring before $\hat{\tau}$. Just after time $\hat{\tau}$, (4.5) is negative so we want to front-load monitoring, which means that monitoring occurs at a constant hazard rate of monitoring that makes the incentive compatibility constraint binding (unless the atom at $\hat{\tau}$ entails monitoring with probability one).[21]

Building upon the previous analysis of the benefits of front-loading monitoring when the shadow cost of the incentive compatibility constraint is high, we can proceed with the proof of Theorem 4.1. The proof relies on the theory of weak duality for infinite-dimensional linear programming problems (Anderson and Nash, 1987, Theorem 2.1). In particular, the proof consists of constructing multipliers for the dual problem such that the value of the dual is the same as the expected payoff of the policy in Theorem 4.1. By weak duality, any feasible solution for the dual problem provides an upper bound for the value of the primal problem. Thus, if we can find feasible multipliers such that the value of the dual is equal to the expected payoff of the policy in Theorem 4.1, then this policy maximizes the principal's expected payoff.

21. A complication arises because $h(\tau)$ is quasi-convex rather than just decreasing. As a result, the reasoning so far could imply a second time $\tilde{\tau}$ at which $h(\tau)$ crosses $\mu(r+\lambda)$ from below (and remains above thereafter). This would suggest the possibility that having a second atom is optimal. The full analysis of the problem, once we incorporate all the remaining incentive compatibility constraints, verifies that adding an extra atom is suboptimal.

The first step is to derive the dual optimization problem. The Lagrangian for the principal problem is

$$\mathcal{L}(F,\Psi,\eta)=\int_0^\infty V(\tau)\mathrm{d}F(\tau)+\int_0^\infty\int_\tau^\infty\left(e^{-(r+\lambda)(s-\tau)}-\underline{q}\right)\mathrm{d}F(s)\mathrm{d}\Psi(\tau)$$
$$+\eta\left(1-\int_0^\infty \mathrm{d}F(\tau)\right),$$

where $\Psi(\tau)$ is the cumulative Lagrange multiplier (an integral of the individual Lagrange multipliers on the continuum of incentive compatibility constraints). The dual problem can be derived starting from the Lagrangian above. By changing the order of integration, the Lagrangian can be written as

$$\mathcal{L}(F,\Psi,\eta)=\eta+\int_0^\infty\left(V(\tau)-\eta+\int_0^\tau\left(e^{-(r+\lambda)(\tau-s)}-\underline{q}\right)\mathrm{d}\Psi(s)\right)\mathrm{d}F(\tau).$$

The optimization of the Lagrangian is finite only if the following inequality is satisfied for all $\tau\geq 0$:

$$V(\tau)-\eta+\int_0^\tau\left(e^{-(r+\lambda)(\tau-s)}-\underline{q}\right)\mathrm{d}\Psi(s)\leq 0.$$

It follows that the dual of the maximization problem (4.4) is given by

$$\begin{cases}\min_{\eta,\Psi}\ \eta\\ \text{subject to}\\ V(\tau)-\eta+\int_0^\tau\left(e^{-(r+\lambda)(\tau-s)}-\underline{q}\right)\mathrm{d}\Psi(s)\leq 0,\forall\tau\geq 0\\ \Psi(0)\geq 0\\ \Psi(\tau)\text{ is nondecreasing}\end{cases}\tag{4.6}$$

Remark 4.2. *Our strategy for the proof has been to conjecture the shape of the optimal policy and then verify its optimality by looking at the dual problem. This approach works in hindsight, once we know the general shape of the solution. However, it is difficult to guess the solution directly from the principal problem* (4.4). *An alternative approach consists of rewriting the principal problem* (4.3) *as a dynamic optimization problem, which can then be analyzed using tools from optimal control. This was the original approach we followed for the analysis. For the interested Reader, we present such formulation in the appendix.*

The first step in the analysis of the dual problem, is to consider the best monitoring policy within the class of random policies described in Theorem 4.1. This comes down to solving the following maximization problem

$$\max_{\hat{\tau}\in[0,\tau^{\text{bind}}]}\left(\frac{e^{(r+\lambda)\hat{\tau}}-1}{1-\underline{q}}\right)\underline{q}V(\hat{\tau})+\left(\frac{1-e^{(r+\lambda)\hat{\tau}}\underline{q}}{1-\underline{q}}\right)\int_{\hat{\tau}}^\infty m^*e^{-m^*(\tau-\hat{\tau})}V(\tau)\mathrm{d}\tau.\tag{4.7}$$

The first order condition of the optimization problem (4.7) can be written in terms of the function $h(\tau)=e^{(r+\lambda)\tau}$ as follows:

$$h(\hat{\tau}^*)=\int_{\hat{\tau}^*}^\infty \rho e^{-\rho(s-\hat{\tau}^*)}h(s)\mathrm{d}s,\tag{4.8}$$

where $\rho \equiv (r+\lambda+m^*)$.[22] The objective function in problem (4.7) is quasi-concave, so it can be shown that $\hat{\tau}^* < \tau^{\text{bind}}$ only if

$$h(\tau^{\text{bind}}) < \int_{\tau^{\text{bind}}}^{\infty} \rho e^{-\rho(s-\tau^{\text{bind}})} h(s)\mathrm{d}s, \tag{4.10}$$

and that $\hat{\tau}^* > 0$ only if [23]

$$h(0) > \int_0^{\infty} \rho e^{-\rho s} h(s)\mathrm{d}s. \tag{4.11}$$

Notice that the optimality conditions can be fully specified in terms of the h function identified in our perturbation analysis in equation (4.5). Equations (4.9) through (4.11) are instrumental in the construction in the multipliers for the dual problem, which provide a verification argument for the optimality of our conjectured policy. If both inequalities, (4.10) and (4.11), hold, then there is a unique $\hat{\tau}^* < \tau^{\text{bind}}$ that satisfies the first order condition in equation (4.8). It follows from here that, for an arbitrary continuation value of $\mathbf{U} = (U_L, U_H)$, the optimal policy can be fully characterized in terms of the function $h(\tau)$. The next proposition provides this characterization.

Proposition 4.3. *Let* $\tau^*_\theta \equiv \operatorname{argmax}_{\tau \in [0,\tau^{bind}]} V_\theta(\tau|\mathbf{U})$, *and* $h_\theta(\tau) \equiv e^{(r+\lambda)\tau} V'_\theta(\tau|\mathbf{U})$. *The optimal policy is the following:*

1. *If* $\tau^*_\theta < \tau^{bind}$, *then the optimal policy is deterministic monitoring at time* τ^*.
2. *If* $\tau^*_\theta = \tau^{bind}$ *and*

$$h_\theta(\tau^{bind}) \geq \int_{\tau^{bind}}^{\infty} \rho e^{-\rho(s-\tau)} h_\theta(s)\mathrm{d}s,$$

then the optimal policy is deterministic monitoring at time τ^{bind}.

3. *If* $\tau^*_\theta = \tau^{bind}$ *and*

$$h_\theta(\tau^{bind}) < \int_{\tau^{bind}}^{\infty} \rho e^{-\rho(s-\tau)} h_\theta(s)\mathrm{d}s,$$

then the optimal policy is random with a distribution given by Theorem 4.1, where $\hat{\tau}^*_\theta$ *is given by*

$$\hat{\tau}^*_\theta = \inf\left\{ \tau \in [0,\tau^{bind}] : h_\theta(\tau) \leq \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h_\theta(s)\mathrm{d}s \right\}.$$

Proposition 4.3 establishes Theorem 4.1. However, this characterization of the optimal policy is for an arbitrary value of the principal's continuation value $\mathbf{U}$. To fully solve the principal problem then we also have to solve for the continuation value consistent with the optimal continuation policy. We turn to that problem next.

22. After some straightforward manipulations of the first order condition, we can write it as

$$\frac{V'(\hat{\tau}^*)}{r+\lambda+m^*} = E[V(\tau)|\tau > \hat{\tau}^*] - V(\hat{\tau}^*). \tag{4.9}$$

We arrive to equation (4.8) using integration by parts.

23. The inequalities in (4.10) and (4.11) follow from a single crossing argument provided in the appendix in Lemma B.4.

4.2. *Solving the Bellman Equation and Finding the Optimal Policy*

In this section we provide a simple characterization of a Bellman equation that allows us to find the optimal policy and corresponding continuation payoffs. Theorem 4.1 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. If we ignore the incentive compatibility constraint, the optimal policy is deterministic and is given by the maximizer of $V(\tau)$. However, such a policy might entail infrequent monitoring and violate the incentive compatibility constraint. In such a case, we need to consider policies that might entail some randomization. By Theorem 4.1, the optimal random policy is fully described by the monitoring rate m^* and the length of the quiet period, captured by $\hat{\tau}^*_\theta$, which pins down the size of the atom that initializes the random monitoring phase. Given the simple form of an optimal policy, we can reduce the optimization problem for the fixed point to the analysis of a simple one-dimensional maximization problem. 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 solution to the principal's problem is thus given by the fixed point:

$$U_L = \max\{\mathcal{G}^L_{\text{det}}(U_L, U_H), \mathcal{G}^L_{\text{rand}}(U_L, U_H)\} \tag{4.12a}$$

$$U_H = \max\{\mathcal{G}^H_{\text{det}}(U_L, U_H), \mathcal{G}^H_{\text{rand}}(U_L, U_H)\} \tag{4.12b}$$

The operator in (4.12) is a contraction so a unique fixed point exists (see proof of Lemma D.1 in the appendix).

To build additional economic intuition, it is useful to analyze two polar cases. In the next subsection, we analyze a relaxed problem in which we ignore the incentive compatibility constraint. In this case, the principal monitors to maximize the benefit of learning net of the monitoring cost. We show that the optimal monitoring policy is deterministic with periodic reviews. Next, we look at the optimal policy when learning is not valuable (i.e., $u(x)$ is linear). In this case, the principal looks for the incentive compatible policy that minimizes the cost of inspections. We show that the cost-minimizing policy entails random monitoring with a constant hazard rate. In the general case, the trade-off echoes these two benchmarks. As in the linear case, to minimize costs subject to satisfying incentive constraints, it is optimal to front-load incentives and hence to monitor with a constant hazard rate. However, when $u(x)$ is convex, as reputation moves from one of the extremes towards the steady-state, inspections generate additional value from learning. The value of learning is zero at the extreme reputations and grows originally fast because beliefs move fast after inspections.

4.3. *Optimal Policy without Moral Hazard*

Without moral hazard, it is optimal to concentrate all the monitoring probability on the time $V(\tau)$ reaches its maximum. The optimal policy is deterministic with an inspection date $\tau^* = \text{argmax}_{\tau \geq 0} V(\tau)$. If the solution is interior, then τ^* satisfies the first order condition $V'(\tau^*) = 0$.[24] The following proposition characterizes the optimal policy in this case.

Proposition 4.4. *In absence of moral hazard, the optimal policy is deterministic with an inspection date*

$$\tau^*_\theta = \underset{\tau \geq 0}{\text{argmax}} \int_0^\tau e^{-rs} u(x^\theta_s) \mathrm{d}s + e^{-r\tau} \mathcal{M}(\mathbf{U}, x^\theta_\tau),$$

where $\mathbf{U} = (U_L, U_H)$ *is the unique solution to the fixed point problem*

$$U_L = \max_{\tau \geq 0} \int_0^\tau e^{-rs} u(x^L_s) \mathrm{d}s + e^{-r\tau} \mathcal{M}(\mathbf{U}, x^L_\tau)$$
$$U_H = \max_{\tau \geq 0} \int_0^\tau e^{-rs} u(x^H_s) \mathrm{d}s + e^{-r\tau} \mathcal{M}(\mathbf{U}, x^H_\tau).$$

To develop some intuition, we consider the first order condition $V'(\tau^*) = 0$, which amounts to

$$u(x^\theta_{\tau^*}) + \dot{x}_{\tau^*}(U_H - U_L) = rM(\mathbf{U}, x^\theta_{\tau^*}). \tag{4.13}$$

The left hand side is the flow payoff that the principal gets in absence of monitoring while the right hand side is the (normalized) payoff of monitoring immediately. At the optimal inspection date τ^*, the principal is indifferent between inspecting now or later. Additional intuition can be obtained by thinking about the continuation value at time t as a function of the belief at time x_t, which we denote by $U(x_t)$. If the principal does not monitor at time t, the continuation value satisfies the standard HJB equation

$$rU(x_t) = u(x_t) + \lambda(\bar{a} - x_t) U'(x_t). \tag{4.14}$$

If the principal inspects at time τ^*, the value function satisfies the value matching condition $U(x_{\tau^*}) = M(\mathbf{U}, x^\theta_{\tau^*})$. Moreover, at the optimal inspection time, the value function satisfies the smooth pasting condition $U'(x_{\tau^*}) = U_H - U_L$. Substituting both conditions in the HJB equation (4.14) we get the first order condition (4.13). Figure 1, illustrates the optimal policy. The time of inspection τ^*_θ is such $x^\theta_{\tau^*_\theta}$ equals the threshold belief $x^*(\theta)$. The thresholds $x^*(\theta)$ 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 $\bar{a}$. In the online appendix, we provide a full analysis of the Principal problem using dynamic programming.

24. It follows from the quasi-convexity of $h(\tau)$ that the first order condition together with the second order condition are sufficient.

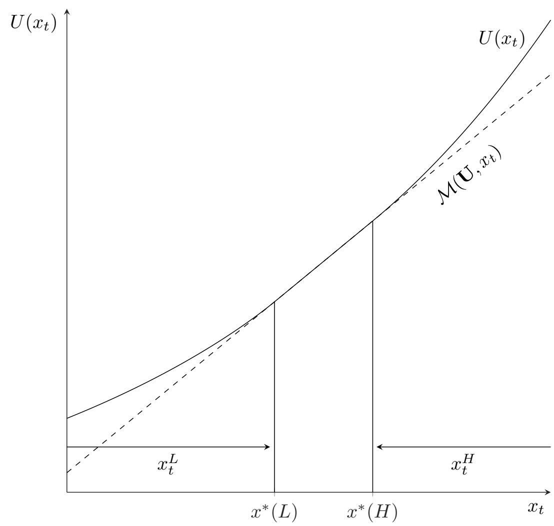


FIGURE 1

Value Function. The optimal policy requires to monitor whenever $x_t^L = x^*(L)$ and $x_t^H = x^*(H)$.

4.4. *Linear Payoffs: Information without Direct Social Value*

Next, 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 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 principal. Hence, the principal's problem boils down to minimizing the expected monitoring costs, subject to the incentive compatibility constraints. Accordingly, using Proposition 3.1, we can reduce the principal's problem to the following cost minimization problem:

$$\begin{cases} C_0 = \inf_{(T_n)_{n\geq 1}} E\left[\sum_{n\geq 1} e^{-rT_n} c \middle| \mathcal{F}_0^P\right] \\ \text{subject to:} \\ \frac{k}{\lambda} \leq \frac{1}{r+\lambda} E\left[e^{-(r+\lambda)(T_{n+1}-t)} | \mathcal{F}_t\right] \forall t \in [T_n, T_{n+1}). \end{cases} \tag{4.15}$$

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 4.5. *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 4.5 follows from the fact that in (4.15) 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 a rate $m^*=(r+\lambda)\underline{q}/(1-\underline{q})$, 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.

As mentioned above, random monitoring is prevalent in the real world: restaurant hygiene is inspected randomly in the U.S.; school quality is inspected randomly in the U.K.; firms are inspected randomly for safety and health hazards in the U.S. However, most monitoring systems feature hazard rates that evolve over time, suggesting that the linear model is not a good description of these real-world applications. In Section **??**, we discuss other potential explanations of why hazard rates are not constant in the real world.

In the online appendix, we provide an alternative proof based on a perturbation argument. However, we can immediately verify that the policy in Proposition 4.5 is optimal applying Proposition 4.3 to the particular case in which $u(x)$ is linear. If $u(x)=x$ and the principal uses a policy which is random with constant arrival rate m^*, then a simple calculation yields

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

Thus, we get that

$$h(\tau)=e^{\lambda\tau}(r+m^*)c,$$

which means that

$$h(0)=(r+m^*)c<\int_0^\infty \rho e^{-\rho s}h(s)\mathrm{d}s=(r+m^*)c\left(1+\frac{\lambda}{r+m^*}\right).$$

Thus, by Proposition 4.5, the policy with constant hazard rate m^* is optimal.

Remark 4.6. *It follows from the alternative proof of Proposition 4.5 found in the appendix 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 r_P, then Proposition 4.5 still holds as long as $r_P<r+\lambda$. If the principal is sufficiently impatient, that is if $r_P>r+\lambda$, then the optimal policy in the linear case involves purely deterministic monitoring.*

4.5. *Comparative Statics*

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 ***4.7 (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.
2. There is $k^{\dagger}<\lambda/(r+\lambda)$ such that for any $k>k^{\dagger}$ the optimal policy is random.
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 $\theta_{T_{n-1}}=L$ is random.
4. Consider the limit when $\lambda\to\infty$, $c\to 0$, and $\lambda c\in(0,\infty)$. In this limit, the optimal policy is random with constant monitoring rate m^*.

Figure 2a 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. In contrast, when the cost of monitoring is high, the optimal policy specifies constant random monitoring starting at time zero. Similarly, Figure 2b shows 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 intermediate. We provide a more detailed analysis of the comparative statics in the next section in the context of a model with linear quadratic preferences and quality driven by Brownian motion.

Cost is a key dimension that determines the optimal design of a monitoring system. Coming back to our school monitoring example, a well-documented report estimates that a school quality review system in the United States analogous to the British system –which targets the entire universe of schools within a three years cycle— would cost between \$635 million and \$1.1 billion annually, depending on the methodology.[25] Another report estimates that a system that reviews every school every three years would cost approximately \$2.5 billion a year (see "On Her Majesty's School Inspection Service" by Craig D. Jerald). Currently, school quality systems around the world target the entire universe of schools. In this case, our results are suggestive that, given relatively high monitoring costs, a random inspection scheme could be an efficient way to lower costs while still ensuring incentives are in place.[26]

5. 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. The policy in the linear case remains optimal for a general class of quality processes. In this section, we analyze the optimal policy when information is valuable, and quality follows the Ornstein-Uhlenbeck process

$$d\theta_t=\lambda(a_t-\theta_t)dt+\sigma dB_t, \tag{5.16}$$

25. For example, Rhode Island, after 12 years decided to eliminate its school quality inspection system due to budget cuts.

26. Of course, the previous recommendation should be taken with a grain of salt as we cannot look at the nominal cost of monitoring in isolation, but we need to look at the cost of monitoring relative to the value of information captured by $u(x)$.

where B_t is a Brownian motion.

The incentive compatibility in Proposition 3.1 holds for any process for quality satisfying the stochastic differential equation

$$d\theta_t = \lambda(a_t - \theta_t)dt + dZ_t,$$

where Z_t is a martingale. In particular, it holds when Z_t is a Brownian motion B_t, so quality follows the Ornstein-Uhlenbeck process in equation (5.16).

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 $\Sigma_t \equiv \mathrm{Var}(\theta_t|\mathcal{F}_t^M)$. 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 (5.16), the distribution of θ_t is Gaussian with moments

$$x_t = \theta_0 e^{-\lambda t} + \bar{a}\left(1 - e^{-\lambda t}\right) \tag{5.17}$$

$$\Sigma_t = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right). \tag{5.18}$$

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 the optimal cost of inspection $\mathcal{C}$ is given by the solution to the fixed point problem

$$\mathcal{C} = \min\left\{\int_0^\infty C(\tau)dF(\tau) : \int_\tau^\infty \left(e^{-(r+\lambda)(s-\tau)} - \underline{q}\right)dF(s) \geq 0, \quad \forall \tau \geq 0\right\},$$
$$C(\tau) \equiv \int_0^\tau \gamma\Sigma_s ds + e^{-r\tau}(c + \mathcal{C}).$$

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 distribution over the monitoring date $F(\tau)$. 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[27]

$$\mathcal{C}=\min\left\{\min_{\bar{\tau}\in[0,\tau^{\text{bind}}]}\frac{\int_0^{\bar{\tau}}e^{-r\tau}\gamma\Sigma_\tau \mathrm{d}\tau+e^{-r\bar{\tau}}c}{1-e^{-r\bar{\tau}}},\right.$$
$$\left.\min_{\hat{\tau}\in[0,\tau^{\text{bind}}]}\frac{\int_0^{\hat{\tau}}e^{-r\tau}\gamma\Sigma_\tau \mathrm{d}\tau+e^{-r\hat{\tau}}\left(\frac{1-e^{(r+\lambda)\hat{\tau}}\underline{q}}{1-\underline{q}}\right)\int_{\hat{\tau}}^{\infty}e^{-(r+m^*)(\tau-\hat{\tau})}\gamma\Sigma_\tau \mathrm{d}\tau+\delta(\hat{\tau})c}{1-\delta(\hat{\tau})}\right\}, \tag{5.19}$$

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 5.1. *Suppose that θ_t follows the Ornstein-Uhlenbeck process in* (5.16)*, 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 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 4.7: 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 (5.18) and (5.19) we see that the

27. In the whole paper, we focus on policies that induce full effort from the agent. In this quadratic specification, it is possible to verify that if the cost of monitoring is not to high, the optimal policy we described is better than any other stationary policy, that is deterministic or monitors at a rate m^*, even when we consider policies that do not induce effort at all times. For example, such is the case if $\gamma=1$, $r=0.1$, $k=0.5$, $c=0.25$, $\bar{a}=0.5$, $\lambda=1$ and $\sigma^2=1$. In this numerical example, the optimal policy is random with an atom, and it is given by $\hat{\tau}=0.16<\tau^{\text{bind}}=0.54$.

monitoring policy only depends on the cost of monitoring per unit or risk, $c/\gamma\sigma^2$, so increasing γ/σ^2 is equivalent to reducing the cost of monitoring. We have the following proposition characterizing the comparative statics in the linear-quadratic case.

Proposition ***5.2 (Comparative Statics)*** *Suppose that θ_t follows the Ornstein-Uhlenbeck process in* (5.16)*, 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 $\lambda_n\to\infty$, $\sigma_n=\sigma\sqrt{\lambda_n}$, $\tilde{c}_n=c/\gamma\sigma_n^2$. In this limit, the optimal policy is random with constant monitoring rate m^*.

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 the i.i.d. limit, where quality shocks are highly transitory, shows how the possibility of window dressing moves the optimal policy towards random monitoring.

6. FINAL REMARKS

The dissemination of information about quality incentivizes firms to invest in quality. Monitoring systems are key sources of information in a wide spectrum of applications, ranging from school reviews and product safety to bank solvency and audit quality. Since monitoring systems are usually costly to implement, it is important for governments and regulatory agencies to design them efficiently. However, little is known about optimal monitoring schemes, particularly in situations in which monetary transfers —such as fines— are too small to be relevant, and public information is scarce.

Since Becker (1968), the literature has largely focused on monitoring as a punishment device to deter misbehavior. In this paper, we develop a reputational theory of monitoring that emphasizes its informational role. Our theory is built on the premise that the reputational impact of inspections is often significantly more relevant than the limited pecuniary punishments that our legal system permits. It also emphasizes the notion that monitoring is, in essence, a form of costly information acquisition. As such, our theory is most relevant in settings where the external flow of information is insufficient.

Specifically, we study the optimal monitoring policy in a principal-agent setting, in which the agent is driven by reputation concerns, and fines are infeasible. The agent exerts hidden effort to affect product quality, and his payoff depends on the product's perceived quality. The principal's monitoring policy plays a dual role that is present in most monitoring systems: *i*) a learning role, as monitoring provides valuable information to the principal even in the absence of incentive issues and *ii*) an incentive role, as monitoring outcomes publicly reveal the agent's quality and affect his demand. These two aspects are not only natural but critical in shaping the structure of the optimal monitoring policy. While learning favors the use of deterministic inspections, incentive provision favors the use of random inspections.

One might expect that the combination of both ingredients in a dynamic game would lead to complex, time-varying, monitoring policies; however, the optimal policy is surprisingly simple and easy to implement in practice. Depending on the outcome of the last inspection, the optimal policy is a mixture of a deterministic periodic review and random inspections with a constant hazard rate and fixed delay. The optimal policy is consistent with some features observed in the real world. For example, many monitoring systems —such as school reviews in the European Union or safety inspections in the U.S.— incorporate a combination of periodic, deterministic components, and random inspection schemes.

Our model is stylized and ignores several aspects that may be important in practice. We conclude with a discussion on the impact that these aspects on the optimal monitoring policy, and how our model can be extended to incorporate them.

First, to isolate the effect of reputation concerns, we have considered settings without fines, or any sort of monetary transfer, where the agent's incentives are driven purely by reputation/career concerns. We believe that this assumption is natural in many applications in which fines play a secondary role or are outright forbidden. For example, in the case of public schools, there is limited scope to fine schools that exhibit poor performance or to provide high power incentives to school principals. More generally, as thoroughly discussed in Dewatripont et al. (1999), there is a limited role for financial incentives in bureaucracies and many governance agencies, where most incentives are provided by career concerns. In the case of auditing firms, we see that auditors rarely pay fines to the PCAOB, and, if any, these fines are often perceived as immaterial by the leading audit companies.

That being said, in many other situations, fines do play an important role, which may affect the design of monitoring systems. At the extreme, with arbitrary fines, monitoring design becomes trivial since a very small intensity of monitoring combined with a large fine would implement first-best (Lazear (2006)). Our analysis suggests that under limited fines the optimal policy shifts 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.[28] Relatedly, we have considered cases in which monitoring is costly to the monitor but (relatively) costless to the agent. This assumption is broadly consistent with some of our applications, such as restaurant hygiene inspections, health and safety reviews, or auditor inspections. However, in many other situations, regular inspections place a burden on the inspected firms. When inspections are costly to the agent, frequent inspections serve a similar role as money burning technologies considered in the literature on optimal delegation (see Amador and Bagwell (2013)), and the optimal design is likely to change.

28. 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.

In this case, the regulator may use the frequency of inspection as an incentive tool, by rewarding good performance with a lower frequency of inspections and punishing bad schools with a higher frequency of inspections. This might be one of the reasons why in the U.K., good performing schools are inspected every 5 years, while bad performing schools are inspected every 2 years (see Jerald (2012)), or why in some countries (for example, the Netherlands), complying firms are provided inspection holidays (OECD, 2014). In our model, more frequent inspections after bad outcomes arise if the value of information is higher for low performing schools (formally, if the payoff is more convex for low reputation).

Our baseline model ignores external news such as customer reviews and complaints, newspaper articles, and accidents. For example, restaurant hygiene inspections are often triggered by a customer complaint (see, e.g., Jones et al. (2004)). OSHA safety inspections respond to the arrival of news about a firm's safety hazards (Levine et al. (2012)). The optimal monitoring system is sensitive to the presence of such news. The impact of news crucially depends on the specific details of the news process. For example, in the online appendix, we show that the qualitative aspects of the model remain unchanged if the news process reveals the firm's current type at a Poisson arrival rate that is independent of the type. In other instances, the principal might be more likely to learn if the quality of the product is either bad or low (that is, the arrival rate of news depends on the quality of the product or service). In the online appendix we study the case in which the arrival rate depends on quality when the principal's preferences are linear. We show that the arrival of bad news often leads to more frequent inspections and stronger enforcement, which is consistent with how health and safety inspections are performed. We also find that in the bad news case, the rate of monitoring decreases in reputation. The economic mechanism behind these results follows the insight in Board and Meyer-ter-Vehn (2013) that, in the bad news case, the agent's incentives increase in reputation.

We have not considered the possibility that the firm can communicate with the principal. This assumption is natural when the agent does not observe his type, as it might be the case in our application to school quality (the incentive compatibility constraint does not change if we assume that the firm does not observe its type, so the optimal policy remains the same). However, in some applications, firms observe their types and are supposed to self-report any problems.[29] In New York, restaurants that fail a hygiene inspection can secure a quick re-inspection. Such self-reporting could improve the performance of the optimal monitoring policy by avoiding unnecessary inspections. For example, if we allow firms that failed the last inspection to self-report improvements, then the firm's reputation remains at zero until the firm requested re-certification. This improves the principal's payoff due to the role of learning, and the possible lower certification costs (the second effect is ambiguous because allowing re-certification might reduce incentives). While we do not provide a characterization of the optimal policy with self-reporting, we expect that the trade-offs between random and deterministic inspections that we stress in this paper will remain relevant in such a model, while new insights are likely to emerge (for example, a characterization of the timing at which firms are allowed to re-certify upon request).

Another issue is whether the productivity of effort is the same in the good and the bad state. In our model, we assume that the productivity of effort is the same across

29. 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.

states. This assumption simplifies the incentive constraints because the marginal return to effort is the same across states; if the productivity of effort differs across states, then our analysis holds as long as we want to maintain full effort in both states. In this case, the relevant incentive compatibility constraint is the one for the state with the lowest productivity of effort.[30]

Finally, two assumptions have a crucial role in our analysis. First, we have only considered policies that induce full effort after all histories; however, for some parameters, the optimal policy will likely prescribe no effort at all, after some bad histories, as a punishment. Even if the full effort is optimal in the first-best, this prescription can be optimal if conditional on the full effort, the probability of maintaining high quality is very high. The intuition for this conjecture comes from the case with $\bar{a}=1$, in which the firm can maintain quality forever as long as it works. Because an inspection revealing low quality only arises off-the-equilibrium path, we can relax the incentive compatibility constraints (at no cost) using the worst possible punishment for the firm after that outcome, which amounts to stop monitoring the firm altogether. This punishment leads to no effort off path, and it is akin to revoking a firm's license, which leads to the lowest possible payoff for the firm. By continuity, we expect that using such strong punishments with some probability remains optimal if $\bar{a}$ is very close to 1 (see Marinovic, Skrzypacz, and Varas (2018) for analysis along these lines in the context of voluntary certification). However, if the cost of inspections is not too high and $\bar{a}$ is sufficiently smaller than one, then we expect that 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 prevent the principal from implementing zero effort as a punishment. For example, in the case of 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. The optimal policy might also fail to implement the same effort at all times if the cost of effort is strictly convex, and the optimal policy implements interior effort. In this case, any policy that entails a deterministic review necessarily implements a time-varying effort. When the cost of effort is convex, our analysis holds if the marginal cost of effort evaluated at $\bar{a}$ is low enough so the optimal policy implements the maximum level of effort.[31]

Second, the simplicity of the incentive compatibility constraint rests on the assumption that the payoff of the firm is linear in reputation, which means that the firm only cares about its average reputation. However, in some markets, the firm's payoffs are likely non-linear in the firm's reputation, as is likely the case for restaurants, where consumers only go to restaurants that have a sufficiently high hygiene reputation, making

30. If the agent does not observe the current state, then 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. Our intuition about the optimal policy, in that case, is that, because after a deviation to lower effort the agent assigns a lower probability to the high state than the principal, so the incentive compatibility constraint in the high state becomes slack if productivity is higher in the low state, we can still characterize the optimal policy using our current methods. However, if the productivity in the low state is lower, then the analysis gets more complicated because we may need to keep the incentive compatibility constraint to prevent "double deviations".

31. Also, if the payoff function is linear, so there is no value of information acquisition, a constant hazard rate with constant effort might still be optimal as this allows to smooth the cost of effort over time. Of course, the same caveat about the role of low effort punishment still applies in this case.

the restaurant's payoffs convex in reputation. In these cases, monitoring could have an additional effect of providing direct value to the firm. If the firm's payoff is convex, then the firm can be rewarded with frequent inspections, as this increases the volatility of reputation. This leads to the seemingly counterfactual implication that firms with high performance are inspected more often. We believe that this possibility is likely to push the optimal monitoring policy towards deterministic reviews. Analyzing the optimal policy in this case is more difficult than in our model because information has direct value to the firm, so inspections provide additional incentives, and the frequency of inspections can be used by the regulator to reward or punish the firm. The full analysis of these kinds of punishments, as well as the analysis of time-varying effort, requires the use of a different set of techniques, similar to the ones in Fernandes and Phelan (2000), and we believe this is an important direction for future research.

Acknowledgments. 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 Columbia, Duke, FTG (Princeton), MIT, Stanford, Toulouse, UCLA, Warwick, and the Columbia/Duke/MIT/Northwestern IO Theory Conference for helpful comments.

REFERENCES

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 K. Bagwell (2013). The theory of optimal delegation with an application to tariff caps. *Econometrica 81*(4), 1541–1599.

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

Anderson, E. J. and P. Nash (1987). *Linear Programming in Infinite Dimensional Spaces: Theory and Applications*. New York: Wiley.

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.

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.

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.

Jerald, C. D. (2012). On her majesty's school inspection service. *Washington, DC: Educator Sector. Retrieved on August 8*, 2014.

Jin, G. Z. and P. Leslie (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics 118*(2), 409–451.

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

Jones, T. F., B. I. Pavlin, B. J. LaFleur, L. A. Ingram, and W. Schaffner (2004). Restaurant inspection scores and foodborne disease. *Emerging Infectious Diseases 10*(4), 688.

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.

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

Levine, D. I., M. W. Toffel, and M. S. Johnson (2012). Randomized government safety inspections reduce worker injuries with no detectable job loss. *Science 336*(6083), 907–911.

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 10*(2), 58–82.

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.

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.

Solan, E. and C. Zhao (2019). Optimal dynamic inspections. Working Paper.

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

A. INCENTIVE COMPATIBILITY: PROOF PROPOSITION 3.1

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

$$Z_t^{LH} = N_t^{LH} - \int_0^t (1-\theta_s)\lambda a_s \mathrm{d}s$$

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

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

$$\mathrm{d}\theta_t = \lambda(a_t - \theta_t)\mathrm{d}t + \mathrm{d}Z_t, \tag{A.1}$$

which leads to equation (A.1). Full effort is incentive compatible if and only if for any deviation $\hat{a}_t$ (with an associated process for quality $\hat{\theta}_t$)

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

Letting $\Delta \equiv \Pi(H) - \Pi(L)$ and replacing the solution for θ_t in (A.1), we can write the incentive compatibility condition as

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

For any deviation we have that

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

So, we can write the incentive compatibility condition as

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

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

B. ANALYSIS PRINCIPAL PROBLEM: PROOF OF THEOREM 4.1

In order to analyze the dual problem (4.6), we first derive an ordinary differential equation, which will be essential for the construction of the multipliers, that must be satisfied whenever the inequality constraint in the dual problem (4.6) binds. Denoting the first date at which there is monitoring with positive probability by $\hat{\tau}$, we look to construct multipliers when the incentive compatibility constraint is binding after $\hat{\tau}$. Instead of directly working with the multiplier $\Psi(\tau)$, it is convenient to work with the discounted version of the multiplier, $\tilde{\Psi}(\tau)$, that is defined as

$$\tilde{\Psi}(\tau) = \Psi(0) + \int_0^{\tau} e^{(r+\lambda)s}\mathrm{d}\Psi(s).$$

Clearly, the multiplier $\Psi(\tau)$ is nondecreasing if and only if $\tilde{\Psi}(\tau)$ is non-decreasing. In Lemma B.1 we show that in any interval over which the constraint in the dual problem binds, the Lagrange multipliers solve the following differential equation

$$\tilde{\Psi}'(\tau) = \frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}(\tau) - h(\tau)\Big).$$

In the analysis of the principal's problem, we need to distinguish between several cases. First, we consider the most interesting case in which the incentive compatibility constraint is binding and the optimal policy is random. Second, we consider the case in which $V(\tau)$ attains an interior maximum at $\hat{\tau}^* < \tau^{\text{bind}}$. so the incentive compatibility constraint is slack. Finally, we consider the corner case in which the incentive compatibility constraint is binding and the optimal policy is deterministic with $\hat{\tau}^* = \tau^{\text{bind}}$.

Binding IC constraint with Random Inspections. Suppose that under the optimal policy, incentive compatibility constraint binds at time 0 and at all times $\tau > \hat{\tau}^*$; and it is slack in between. Since $F(\tau)$ is constant on $[0,\hat{\tau}^*)$, the inequality constraint in the dual problem (4.6) does not need to bind, and we only need to verify that the first constraint in the dual problem is not violated on $[0,\hat{\tau}^*)$. On the other hand, because $F(\tau)$ is strictly increasing on $[\hat{\tau}^*,\infty)$, the first inequality constraint in (4.6) must bind on $[\hat{\tau}^*,\infty)$. In particular, if we set $\tilde{\Psi}(\tau)$ to be constant on $[0,\hat{\tau}^*)$, then we find that at the date of the atom

$$e^{(r+\lambda)\hat{\tau}^*}V(\hat{\tau}^*) - e^{(r+\lambda)\hat{\tau}^*}\eta + \tilde{\Psi}(\hat{\tau}^*) - \tilde{\Psi}(0) - \underline{q}\int_0^{\hat{\tau}^*} e^{(r+\lambda)(\hat{\tau}-s)}\mathrm{d}\tilde{\Psi}(s) =$$
$$e^{(r+\lambda)\hat{\tau}^*}V(\hat{\tau}) - e^{(r+\lambda)\hat{\tau}^*}\eta + (1-\underline{q})(\tilde{\Psi}(\hat{\tau}^*) - \tilde{\Psi}(0)) = 0,$$

and from this equation we can conclude that

$$\eta = e^{-(r+\lambda)\hat{\tau}^*}(1-\underline{q})\big(\tilde{\Psi}(\hat{\tau}) - \tilde{\Psi}(0)\big) + V(\hat{\tau}^*). \tag{B.1}$$

The verification argument requires that the value of the dual, which is given by the multiplier η, is equal to the value of the primal given the policy $F(\tau)$ in Theorem 4.1. Thus, the multipliers η must be equal to

$$\eta = \left(\frac{e^{(r+\lambda)\hat{\tau}^*} - 1}{1-\underline{q}}\right)\underline{q}V(\hat{\tau}^*) + \left(\frac{1 - e^{(r+\lambda)\hat{\tau}^*}\underline{q}}{1-\underline{q}}\right)\int_{\hat{\tau}^*}^{\infty} m^* e^{-m^*(\tau-\hat{\tau}^*)}V(\tau)\mathrm{d}\tau. \tag{B.2}$$

If we substitute the expression for η in equation (B.1), which comes from feasibility conditions for the dual, into the value of the dual problem in equation (B.2), we get an expression for the change of the multiplier $\tilde{\Psi}(\tau)$ at time $\hat{\tau}^*$ given by $\tilde{\Psi}(\hat{\tau}^*) - \tilde{\Psi}(0)$:

$$\tilde{\Psi}(\hat{\tau}^*) - \tilde{\Psi}(0) = \frac{e^{(r+\lambda)\hat{\tau}^*}}{1-\underline{q}}\left(1 - e^{(r+\lambda)\hat{\tau}^*}\underline{q}\right)e^{-(r+\lambda)\hat{\tau}^*}\frac{h(\hat{\tau}^*)}{r+\lambda}, \tag{B.3}$$

where we have substituted the first order condition for $\hat{\tau}^*$ in equation (4.8). Notice that the cumulative multiplier $\tilde{\Psi}(\tau)$ jumps at the time $\hat{\tau}^*$ when the incentive compatibility starts to bind. For $\tau \geq \hat{\tau}^*$, we construct the multipliers using the solution to the ODE in equation (B.13) with the transversality condition $\lim_{\tau\to\infty} e^{-\rho\tau}\tilde{\Psi}(\tau) = 0$. From here we get that the Lagrange multiplier $\tilde{\Psi}(\tau)$ is characterized by the solution to the ordinary differential equation on $(\hat{\tau}^*,\infty)$ together with the jump in the multiplier at time $\hat{\tau}^*$ in equation (B.3), so the Lagrange multiplier is given by

$$\tilde{\Psi}(\tau) = \begin{cases} \left(e^{(r+\lambda)\hat{\tau}^*} - 1\right)\frac{\underline{q}}{1-\underline{q}}\frac{h(\hat{\tau}^*)}{r+\lambda} & \text{if } \tau \in [0,\hat{\tau}^*) \\ \frac{1}{1-\underline{q}}\int_{\tau}^{\infty} e^{-\rho(s-\tau)}h(s)\mathrm{d}s & \text{if } \tau \in [\hat{\tau}^*,\infty). \end{cases} \tag{B.4}$$

The only remaining step in the construction of the multipliers is to verify that $\tilde{\Psi}(\tau)$ in equation (B.4) is nondecreasing. It can be easily verified that $\tilde{\Psi}'(\hat{\tau}^*) = 0$, which means that $\Psi'(\tau)$ satisfies the following ODE

$$\tilde{\Psi}''(\tau) = \frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}'(\tau) - h'(\tau)\Big), \tilde{\Psi}'(\hat{\tau}^*) = 0, \tag{B.5}$$

where $h'(\hat{\tau}^*) < 0$.[32] We show in the appendix (Lemma B.3) that if the first order condition in equation (4.8) is satisfied, then the solution to the ordinary differential equation (B.5) is non-negative. Moreover, the jump in $\tilde{\Psi}(\tau)$ at time $\hat{\tau}^*$ given in equation (B.3) is positive. Thus, we conclude the multiplier $\tilde{\Psi}(\tau)$ in equation (B.4) is non-decreasing.

32. By $\tilde{\Psi}'(\hat{\tau}^*)$ we mean the right derivative of $\tilde{\Psi}(\tau)$ at $\hat{\tau}^*$ as $\tilde{\Psi}(\hat{\tau}^*)$ is discontinuous at this point.

In sum, the multipliers $(\eta,\tilde{\Psi})$ described by equations (B.2) and (B.4), are dual feasible with a value of the dual problem that equals the expected payoff of the policy in Theorem 4.1. Thus, $F^*_\theta(\tau)$ is optimal by weak duality.

The verification argument in the case in which $\hat{\tau}^*=0$ is very similar. If

$$h(0)\leq\int_0^\infty \rho e^{-\rho s}h(s)\mathrm{d}s, \tag{B.6}$$

then, by Lemma B.4, we have that that for all $\tau\geq 0$

$$h(\tau)<\int_\tau^\infty \rho e^{-\rho(s-\tau)}h(s)\mathrm{d}s, \tag{B.7}$$

and we can construct the multipliers in the same way as we did before on the interval $[\hat{\tau}^*,\infty)$. In particular, the multiplier $\tilde{\Psi}(\tau)$ is given by[33]

$$\tilde{\Psi}(\tau)=\frac{1}{1-\underline{q}}\int_\tau^\infty e^{-\rho(s-\tau)}h(s)\mathrm{d}s \tag{B.8}$$

$$\eta=\int_0^\infty m^* e^{-m^*\tau}V(\tau)\mathrm{d}\tau,$$

so the value of the dual corresponds to the expected payoff of the monitoring policy $F^*_\theta(\tau)=1-e^{-m^*\tau}$.

Slack IC constraint. We consider the case in which $V(\tau)$ has a maximum on $\hat{\tau}\in[0,\tau^{\text{bind}})$. In this case, monitoring is deterministic and $\hat{\tau}<\tau^{\text{bind}}$. The IC constraint is slack so we can set $\Psi(0)=0$ and $\eta=V(\hat{\tau})$ we get that the value of the objective function in the dual problem equals $V(\hat{\tau})$. Because $V(\hat{\tau})>V(\tau)$ for all $\tau<\hat{\tau}$, all the constraint in the dual problem are satisfied for $\tau<\hat{\tau}$. For $\tau>\hat{\tau}$ we need to consider two cases: (1) $\hat{\tau}$ it is strictly positive and (2) $\hat{\tau}=0$.

Case 1. If $\hat{\tau}$ is strictly positive, then $V'(\hat{\tau})=0$ and $V''(\hat{\tau})<0$, which means that $h(\hat{\tau})=0$ and $h'(\hat{\tau})<0$. Because $h(\tau)$ is quasi-convex, there is at most one local maximum. If $\tau''>0$ is another local maximum, then there must be a local minimum $\hat{\tau}<\tau''<\tau'$. Hence, at τ' we have $h(\tau'')=0$ and $h'(\tau'')>0$. However, Lemma B.2 implies that $h'(\tau')>0$, which means that τ' cannot be another local maximum.

By the previous argument, there cannot be an interior minimum before $\hat{\tau}$. This means that $V(\tau)$ must be increasing for all $\tau<\hat{\tau}$. Given that there is at most one interior local maximum, the global maximum of $V(\tau)$ must belong to $\{\hat{\tau},\infty\}$.

1. If $\lim_{\tau\to\infty}V(\tau)\leq V(\hat{\tau})$, we can set $\tilde{\Psi}(\tau)=0$ and all the constrains are satisfied as $V(\tau)-V(\hat{\tau})\leq 0$ for all $\tau\geq 0$.

33. Because in this case we are assuming that $V(\tau)$ has no maximum on $[0,\tau^{\text{bind}})$, it must be that $V'(0)=h(0)\geq 0$. Hence, inequality (B.6) implies

$$\frac{1}{1-\underline{q}}\int_0^\infty e^{-\rho s}h(s)\mathrm{d}s\geq 0.$$

Differentiating equation (B.8) and using inequality (B.7) we get

$$\tilde{\Psi}'(\tau)=\rho\tilde{\Psi}(\tau)-\frac{h(\tau)}{1-\underline{q}}=\frac{1}{1-\underline{q}}\int_\tau^\infty \rho e^{-\rho(s-\tau)}h(s)\mathrm{d}s-\frac{h(\tau)}{1-\underline{q}}>0,$$

which means that $\tilde{\Psi}(\tau)$ is nondecreasing, and because $\tilde{\Psi}(0)\geq 0$ we get that $\tilde{\Psi}(\tau)$ is nonnegative.

2. If $\lim_{\tau\to\infty} V(\tau) > V(\hat{\tau})$. Then the global maximum is infinity. In this case, we set $\tilde{\Psi}(\tau)=0$ on $[0,\hat{\tau}]$ and use equation (B.13) to construct the multipliers for $\tau \geq \hat{\tau}$, in particular

$$\tilde{\Psi}(\tau) = e^{\rho(\tau-\hat{\tau})}\tilde{\Psi}(\hat{\tau}) - \frac{1}{1-\underline{q}}\int_{\hat{\tau}}^{\tau} e^{\rho(\tau-s)}h(s)ds$$
$$= -\frac{1}{1-\underline{q}}\int_{\hat{\tau}}^{\tau} e^{\rho(\tau-s)}h(s)ds.$$

Because $\hat{\tau}>0$ we have that $h(\hat{\tau})=0$ so $\tilde{\Psi}'(\hat{\tau})=0$ and $\tilde{\Psi}''(\hat{\tau}) = -h'(\hat{\tau})/(1-\underline{q})>0$, so $\tilde{\Psi}'(\tau)>0$ on $(\hat{\tau},\hat{\tau}+\epsilon)$. If $\hat{\tau}=0$, then $\tilde{\Psi}'(\hat{\tau}) = \tilde{\Psi}'(0) = -h(0)/(1-\underline{q})>0$.

We need to verify that $\tilde{\Psi}(\tau)$ is non-decreasing. Looking for a contradiction, suppose that $\tilde{\Psi}(\tau)$ is decreasing at some point and let $\tau^{\dagger} = \inf\{\tau>\hat{\tau} : \tilde{\Psi}'(\tau)<0\}$. Because $\tilde{\Psi}(\hat{\tau})=0$ and $\tilde{\Psi}'(\tau)$ is nondecreasing on $(\hat{\tau},\tau^{\dagger})$ and strictly increasing in some subset, we have that $\tilde{\Psi}(\tau^{\dagger})>0$. The derivative of $\tilde{\Psi}(\tau)$ is

$$\tilde{\Psi}'(\tau) = -\frac{1}{1-\underline{q}}\left(\int_{\hat{\tau}}^{\tau} \rho e^{\rho(\tau-s)}h(s)ds + h(\tau)\right),$$

which means that at time $\tau^{\dagger}$

$$\int_{\hat{\tau}}^{\tau^{\dagger}} \rho e^{\rho(\tau^{\dagger}-s)}h(s)ds + h(\tau^{\dagger}) = 0$$

If $\hat{\tau}>0$, then $h(\hat{\tau})=0$, so, because $h(\tau)$ is quasi-convex, we have

$$0 = \int_{\hat{\tau}}^{\tau^{\dagger}} \rho e^{\rho(\tau^{\dagger}-s)}h(s)ds + h(\tau^{\dagger}) \geq h(\tau^{\dagger}) + \max\{h(\hat{\tau}),h(\tau^{\dagger})\}\int_{\hat{\tau}}^{\tau^{\dagger}} \rho e^{\rho(\tau^{\dagger}-s)}ds$$
$$= h(\tau^{\dagger}) + \max\{h(\hat{\tau}),h(\tau^{\dagger})\} = h(\tau^{\dagger}) + \max\{0,h(\tau^{\dagger})\}.$$

However, $\tilde{\Psi}'(\tau^{\dagger})=0 \Rightarrow h(\tau^{\dagger}) = (r+\lambda)\Psi(\tau^{\dagger})>0$, which yields a contradiction.

Case 2. If $\hat{\tau}=0$, and $h(\tau)\leq 0$ for all $\tau\geq 0$ then $\hat{\tau}=0$ is a global maximum and there is nothing to prove. So, suppose that $h(\tau)>0$ for some τ. Let

$$\Psi(0) = \max\left\{0, \frac{1}{1-\underline{q}}\int_0^{\infty} e^{-\rho s}h(s)ds\right\},$$

and construct the multipliers using (B.13) with initial the condition $\Psi(0)$ defined above. By the comparison principle for ODEs the solution to (B.13) is

$$\tilde{\Psi}(\tau) = \max\left\{-\frac{1}{1-\underline{q}}\int_{\tau'}^{\tau} e^{\rho(\tau-s)}h(s)ds, \frac{1}{1-\underline{q}}\int_{\tau}^{\infty} e^{-\rho(s-\tau)}h(s)ds\right\}.$$

Because $\tilde{\Psi}'(\tau)$ is increasing in $\tilde{\Psi}(\tau)$, we also have by the comparison principle applied to (B.5) that

$$\tilde{\Psi}'(\tau) = \max\left\{-\frac{\rho}{1-\underline{q}}e^{\rho\tau}h(0) - \frac{1}{1-\underline{q}}\left(\int_0^{\tau} \rho e^{\rho(\tau-s)}h'(s)ds + h(\tau)\right), \frac{1}{1-\underline{q}}\int_{\tau}^{\infty} e^{-\rho(s-\tau)}h'(s)ds\right\}$$

Let $\tau' = \inf\{\tau\geq 0 : h(\tau)>0\}$. At time τ', $h(\tau)$ is crossing zero from below, which means that $h'(\tau')>0$, so Lemma B.2 implies that $h'(\tau)>0$ and $h(\tau)>0$ for all $\tau>\tau'$, which means that

$$\tilde{\Psi}'(\tau) \geq \frac{1}{1-\underline{q}}\int_{\tau}^{\infty} e^{-\rho(s-\tau)}h'(s)ds \geq 0.$$

On the other hand, for all $\tau<\tau'$ we have that $h(\tau)<0$ and

$$\tilde{\Psi}(\tau)\geq -\frac{1}{1-\underline{q}}\int_{\tau'}^{\tau} e^{\rho(\tau-s)}h(s)ds\geq 0,$$

so

$$\tilde{\Psi}'(\tau)=\frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}(\tau)-h(\tau)\Big)>0.$$

We conclude that $\tilde{\Psi}(\tau)$ is nonnegative and nondecreasing.

Binding IC constraint with Deterministic Inspection. From now on, we can focus on the more interesting case in which $V(\tau)$ does not have a maximum on $[0,\tau^{\text{bind}})$ so the IC constraint is binding. We need to consider two cases depending if $h(\tau^{\text{bind}})-\int_{\tau^{\text{bind}}}^{\infty}\rho e^{-\rho(s-\tau)}h(s)\text{d}s$ is negative or not. We already presented in the main text the case in which

$$h(\tau^{\text{bind}})<\int_{\tau^{\text{bind}}}^{\infty}\rho e^{-\rho(s-\tau^{\text{bind}})}h(s)\text{d}s, \tag{B.9}$$

so it is only left to consider the case in which the previous condition is not satisfied so there is an atom at τ^{bind}. Thus, consider the case in which

$$h(\tau^{\text{bind}})\geq\int_{\tau^{\text{bind}}}^{\infty}\rho e^{-\rho(s-\tau^{\text{bind}})}h(s)\text{d}s \tag{B.10}$$

Suppose that the constraint $\Gamma(\tau)\leq 0$ in the dual problem is binding at τ^{bind}, then we have that

$$e^{(r+\lambda)\tau^{\text{bind}}}V(\tau^{\text{bind}})-e^{(r+\lambda)\tau^{\text{bind}}}\eta=0,$$

so we immediately get that

$$\eta=V(\tau^{\text{bind}}). \tag{B.11}$$

The constraint $\Gamma(\tau)\leq 0$ has to be satisfy for all $\tau<\tau^{\text{bind}}$. Setting $\Psi(\tau)$ to be constant on $[0,\tau^{\text{bind}}]$, the previous condition reduces to $\Gamma(\tau)=e^{(r+\lambda)\tau}\big(V(\tau)-V(\tau^{\text{bind}})\big)\leq 0$, which holds because we are considering the case in which $V(\tau)$ does not have an interior maximum on $[0,\tau^{\text{bind}})$. The next step is to construct the multipliers for $\tau>\tau^{\text{bind}}$. We set $\tilde{\Psi}(\tau)=\frac{h(\tau^{\text{bind}})}{r+\lambda}$ for all $\tau\in[0,\tau^{\text{bind}}]$ and construct the multipliers for $\tau>\tau^{\text{bind}}$ using the ODE in equation (B.13) with the appropriate initial condition:

$$\tilde{\Psi}'(\tau)=\frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}(\tau)-h(\tau)\Big),\tilde{\Psi}(\tau^{\text{bind}})=\frac{h(\tau^{\text{bind}})}{r+\lambda}. \tag{B.12}$$

Hence, we have that

$$\begin{aligned}
\tilde{\Psi}(\tau)&=\frac{1}{r+\lambda}e^{\rho(\tau-\tau^{\text{bind}})}h(\tau^{\text{bind}})-\frac{1}{1-\underline{q}}\int_{\tau^{\text{bind}}}^{\tau}e^{\rho(\tau-s)}h(s)ds\\
&=\frac{1}{r+\lambda}e^{\rho(\tau-\tau^{\text{bind}})}\left[h(\tau^{\text{bind}})-\frac{r+\lambda}{1-\underline{q}}e^{-\rho(\tau-\tau^{\text{bind}})}\int_{\tau^{\text{bind}}}^{\tau}e^{\rho(\tau-s)}h(s)ds\right]\\
&=\frac{1}{r+\lambda}e^{\rho(\tau-\tau^{\text{bind}})}\left[h(\tau^{\text{bind}})-\int_{\tau^{\text{bind}}}^{\tau}\rho e^{-\rho(s-\tau^{\text{bind}})}h(s)ds\right],
\end{aligned}$$

where $\rho \equiv (r+\lambda)/(1-\underline{q})$. The second derivative is given by

$$\begin{aligned}\tilde{\Psi}''(\tau) &= \frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}'(\tau)-h'(\tau)\Big)\\ &= \frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}'(\tau)-h'(\tau)\Big).\end{aligned}$$

If inequality (B.10) is satisfied, then $\lim_{\tau\to\infty}\tilde{\Psi}(\tau)\geq 0$. Substituting $\tilde{\Psi}(\tau^{\text{bind}})$ in (B.12) we get that $\tilde{\Psi}'(\tau^{\text{bind}})=0$ so the second derivative satisfies

$$\tilde{\Psi}''(\tau) = \frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}'(\tau)-h'(\tau)\Big), \tilde{\Psi}'(\tau^{\text{bind}})=0,$$

where $h'(\tau^{\text{bind}})<0$, so it follows from Lemma B.3 that $\tilde{\Psi}'(\tau)\geq 0$.

B.1. *Proof of Theorem 4.1: Technical Lemmas*

In this section, we prove a series of lemmas required for the construction of the Lagrange multipliers in the analysis of the dual problem.

Lemma B.1. *Let $h(\tau)\equiv e^{(r+\lambda)\tau}V'(\tau)$. Suppose that there is $\hat{\tau}\geq 0$ such that*

$$h(\tau)-e^{(r+\lambda)\tau}\eta+\int_0^\tau\Big(1-\underline{q}e^{(r+\lambda)(\tau-s)}\Big)\mathrm{d}\tilde{\Psi}(s)=0, \forall\tau\geq\hat{\tau},$$

then on $(\hat{\tau},\infty)$ the Lagrange multiplier $\tilde{\Psi}(\tau)$ must satisfy the following differential equation

$$\tilde{\Psi}'(\tau)=\frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}(\tau)-h(\tau)\Big). \tag{B.13}$$

Proof. Let's define

$$\Gamma(\tau)\equiv e^{(r+\lambda)\tau}V(\tau)-e^{(r+\lambda)\tau}\eta+\int_0^\tau\Big(e^{(r+\lambda)s}-\underline{q}e^{(r+\lambda)\tau}\Big)\mathrm{d}\Psi(s),$$

which corresponds to the left hand side of the constraint in the dual problem multiplied by $e^{(r+\lambda)\tau}$. By definition, the multipliers (η,Ψ) are dual feasible if and only if $\Gamma(\tau)\leq 0$. If we write $\Gamma(\tau)$ in terms of the multiplier $\tilde{\Psi}(\tau)$ we get

$$\Gamma(\tau)=e^{(r+\lambda)\tau}V(\tau)-e^{(r+\lambda)\tau}\eta+\int_0^\tau\Big(1-\underline{q}e^{(r+\lambda)(\tau-s)}\Big)\mathrm{d}\tilde{\Psi}(s).$$

At any point in which $\tilde{\Psi}(\tau)$ is differentiable, we have

$$\Gamma'(\tau)=e^{(r+\lambda)\tau}V'(\tau)+\big(1-\underline{q}\big)\tilde{\Psi}'(\tau)+(r+\lambda)\left(e^{(r+\lambda)\tau}(V(\tau)-\eta)-\underline{q}\int_0^\tau e^{(r+\lambda)(\tau-s)}\mathrm{d}\tilde{\Psi}(s)\right),$$

which can be written as

$$\Gamma'(\tau)=e^{(r+\lambda)\tau}V'(\tau)+\big(1-\underline{q}\big)\tilde{\Psi}'(\tau)+(r+\lambda)\Big(\Gamma(\tau)-\tilde{\Psi}(\tau)\Big). \tag{B.14}$$

From here we get that if there is $\hat{\tau}$ such that $\Gamma(\tau)=0$ for all $\tau>\hat{\tau}$, then it must be the case that $\tilde{\Psi}(\tau)$ satisfies the following differential equation

$$\tilde{\Psi}'(\tau)=\frac{1}{1-\underline{q}}\Big((r+\lambda)\tilde{\Psi}(\tau)-h(\tau)\Big). \tag{B.15}$$

where $h(\tau)$ is defined as

$$h(\tau)\equiv e^{(r+\lambda)\tau}V'(\tau), \tag{B.16}$$

||

Lemma B.2. *$h(\tau)$ is quasi-convex, and strictly convex when increasing.*

Proof. Recall

$$h(\tau)\equiv e^{(r+\lambda)\tau}V'(\tau)=e^{\lambda\tau}\Big(u(x^\theta_\tau)-r\mathcal{M}(\mathbf{U},x^\theta_\tau)+\dot{x}^\theta_\tau(U_H-U_L)\Big).$$

We have:

$$\begin{aligned}
h'(\tau)&=\lambda h(\tau)+e^{\lambda\tau}\Big(\dot{x}^\theta_\tau\Big[u'(x^\theta_\tau)-(r+\lambda)(U_H-U_L)\Big]\Big)\\
h''(\tau)&=\lambda h'(\tau)+\lambda e^{\lambda\tau}\Big(\dot{x}^\theta_\tau\Big[u'(x^\theta_\tau)-(r+\lambda)(U_H-U_L)\Big]\Big)\\
&\quad+e^{\lambda\tau}\Big(-\lambda\dot{x}^\theta_\tau\Big[u'(x^\theta_\tau)-(r+\lambda)(U_H-U_L)\Big]\Big)+e^{\lambda\tau}\Big(\dot{x}^\theta_\tau\Big)^2u''(x^\theta_\tau)\\
&=\lambda h'(\tau)+e^{\lambda\tau}\Big(\dot{x}^\theta_\tau\Big)^2u''(x^\theta_\tau)
\end{aligned}$$

So when $h'(\tau)=0$ then $h''(\tau)>0$. So whenever the derivative of h becomes zero, h is strictly convex locally. So h is quasi-convex: $h'(\tau)$ changes sign only at most once from negative to positive. ||

Lemma B.3. *Suppose that $h'(\hat{\tau})\le 0$, then the solution to the ODE*

$$f'(\tau)=\frac{1}{1-\underline{q}}\big((r+\lambda)f(\tau)-h'(\tau)\big)$$

with initial condition $f(\hat{\tau})=0$ is non-negative on $(\hat{\tau},\infty)$ if and only if

$$h(\hat{\tau})\ge\int_{\hat{\tau}}^{\infty}\rho e^{-\rho(s-\hat{\tau})}h(s)ds.$$

Proof. The solution to the ODE is

$$f(\tau)=-\frac{1}{1-\underline{q}}\int_{\hat{\tau}}^{\tau}e^{\frac{r+\lambda}{1-\underline{q}}(\tau-s)}h'(s)\mathrm{d}s.$$

We want to show that for $\tau>\hat{\tau}$

$$\int_{\hat{\tau}}^{\tau}e^{-\rho s}h'(s)ds\le 0.$$

for which, given the quasi-convexity of h, it is necessary and sufficient to show that

$$Z\equiv\int_{\hat{\tau}}^{\infty}e^{-\rho s}h'(s)ds\le 0.$$

Integrating by parts we get

$$\int_{\hat{\tau}}^{\infty} e^{-\rho s} h'(s) ds = \lim_{\tau\to\infty} h(\tau) e^{-\rho\tau} - e^{-\rho\hat{\tau}} h(\hat{\tau}) + \int_{\hat{\tau}}^{\infty} \rho e^{-\rho s} h(s) ds$$

Since

$$h(\tau)e^{-\rho\tau} = e^{(r+\lambda-\rho)\tau} V'(\tau) = e^{(\lambda-\rho)\tau}\left(u(x_\tau^\theta) - r\mathcal{M}(\mathbf{U}, x_\tau^\theta) + \dot{x}_\tau^\theta (U_H - U_L)\right)$$

we have that

$$\lim_{\tau\to\infty} h(\tau)e^{-\rho\tau} = 0$$

Therefore

$$\begin{aligned} Z &= -e^{-\rho\hat{\tau}} h(\hat{\tau}) + \int_{\hat{\tau}}^{\infty} \rho e^{-\rho s} h(s) ds \\ &= e^{-\rho\hat{\tau}}\left(\int_{\hat{\tau}}^{\infty} \rho e^{-\rho(s-\hat{\tau})} h(s) ds - h(\hat{\tau})\right) \end{aligned}$$

Therefore $f(t)$ is weakly positive for all $t \geq \hat{\tau}$ if and only if

$$\int_{\hat{\tau}}^{\infty} \rho e^{-\rho(s-\hat{\tau})} h(s) ds \leq h(\hat{\tau})$$

‖

Lemma B.4. *The first order condition*

$$V'(\hat{\tau}) = \frac{(r+\lambda)}{1-\underline{q}}\left(\int_{\hat{\tau}}^{\infty} m e^{-m(\tau-\hat{\tau})} V(\tau) \mathrm{d}\tau - V(\hat{\tau})\right)$$

is equivalent to the condition

$$h(\hat{\tau}) = \int_{\hat{\tau}}^{\infty} \rho e^{-\rho(\tau-\hat{\tau})} h(\tau) \mathrm{d}\tau.$$

Moreover,

1. If $h'(\tau) < 0$ *on* $(0,\infty)$ *then*

$$h(\tau) > \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s$$

for all $\tau \geq 0$.

2. If

$$h(\tilde{\tau}) \leq \int_{\tilde{\tau}}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s$$

then

$$h(\tau) < \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s$$

for all $\tau > \tilde{\tau}$.

3. Suppose that $h'(0) < 0$ *and*

$$h(0) > \int_{0}^{\infty} \rho e^{-\rho s} h(s) \mathrm{d}s,$$

and let let $\tilde{\tau} = \inf\{\tau : h'(\tilde{\tau}) \geq 0\}$. *Then, there is a unique* $\hat{\tau} < \tilde{\tau}$ *satisfying the first order condition such that*

$$\begin{aligned} h(\tau) &\geq \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s, \forall \tau \in [0,\hat{\tau}) \\ h(\tau) &\leq \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s, \forall \tau \in (\hat{\tau},\tilde{\tau}]. \end{aligned}$$

Proof. Letting

$$\rho = \frac{r+\lambda}{1-\underline{q}} = r+\lambda+m$$

we can write the first order condition as

$$e^{-\rho\hat{\tau}} h(\hat{\tau}) = \frac{(r+\lambda)}{1-\underline{q}} \left(\int_{\hat{\tau}}^{\infty} m e^{-m\tau} V(\tau) \mathrm{d}\tau - e^{-m\hat{\tau}} V(\hat{\tau}) \right)$$

Using integration by parts

$$\int_{\hat{\tau}}^{\infty} m e^{-m\tau} V(\tau) \mathrm{d}\tau = e^{-m\hat{\tau}} V(\tau) + \int_{\hat{\tau}}^{\infty} e^{-m\tau} V'(\tau) \mathrm{d}\tau$$

which yield

$$e^{-\rho\hat{\tau}} h(\hat{\tau}) = \frac{(r+\lambda)}{1-\underline{q}} \int_{\hat{\tau}}^{\infty} e^{-m\tau} V'(\tau) \mathrm{d}\tau.$$

Part 1:. If $h'(\tau) \leq 0$ is negative for all τ, then it is immediate that

$$h(\tau) > \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s).$$

Part 2:. We consider two cases, $h'(\tilde{\tau}) \geq 0$ and $h'(\tilde{\tau}) < 0$. In the first case, because $h(\tau)$ is quasi-convex we have that $h(\tau)$ is increasing on $(\tilde{\tau},\infty)$, which immediately implies that

$$h(\tau) < \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s).$$

In the second case, let's define

$$H(\tau) \equiv h(\tau) - \int_{\tau}^{\infty} \rho e^{-\rho(s-\tau)} h(s) \mathrm{d}s,$$

which satisfies the following ODE.

$$H'(\tau) = \rho H(\tau) + h'(\tau). \tag{B.17}$$

Suppose by contradiction, that there is some $\tau' > \tilde{\tau}$ such that $H(\tau') > 0$, and let $\tau^{\dagger} = \inf\{\tau \in (\tilde{\tau},\tau') : H(\tau) > 0\}$. $H(\tau^{\dagger}) = 0$ because $H(\tau)$ is continuous, and $h'(\tau^{\dagger}) < 0$ as $h'(\tau) \geq 0$ implies $H(\tau) > 0$. But then, equation (B.17) implies that $H'(\tau^{\dagger}) < 0$, which contradicts $H(\tau^{\dagger} + \epsilon) > 0$ for a sufficiently small ϵ. Hence, it must be the case that $H(\tau) < 0$ for all $\tau > \tilde{\tau}$.

Part 3:. Suppose that $H(0) > 0$ and let $\tilde{\tau} = \inf\{\tau : h'(\tau) \geq 0\}$, which means that $H(\tilde{\tau}) > 0$. By continuity there is $\hat{\tau} \in (0,\tilde{\tau})$ such that $H(\hat{\tau}) = 0$. Moreover, because $h'(\tau) < 0$ for all $\tau < \tilde{\tau}$, equation (B.17) implies that $H(\tau) > 0$ on $[0,\hat{\tau})$ and $H(\tau) < 0$ on $(\hat{\tau},\tilde{\tau}]$. ||

C. PROOF BROWNIAN LINEAR-QUADRATIC MODEL

Proof of Proposition 5.1

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. ‖

D. ANALYSIS USING OPTIMAL CONTROL

In this section of the appendix, we sketch an alternative proof of Theorem 4.1 that uses optimal control techniques instead of the weak duality approach. The main advantage of this approach is that it does not require a guess of the shape of the optimal monitoring policy in advance. To simplify the number of cases to consider, we assume here that $u(x)$ is strictly convex and that the solution of the relaxed problem (ignoring incentive compatibility constraints) is not incentive compatible. Proofs of all lemmas are relegated to the online appendix.

This proof is based on the analysis of necessary conditions that any optimal policy must satisfy. Therefore, the first step in the analysis is to show that an optimal policy actually exists. The following lemma establishes that a fixed point in (4.3) exists , is unique, and that the supremum is attained.

Lemma *D.1 (Existence)* *Suppose the* max *in* (4.3) *is replaced by a* sup. *Then 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* (4.3) *(so the* sup *is attained).*

The next step is to 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}, conditional on reaching time τ. That is, q_τ, represents the expected discounted time until the next review, where the effective discount rate incorporates the depreciation rate λ. The incentive compatibility constraint in Proposition 3.1 becomes $q_\tau\geq\underline{q}$.

We can derive the law of motion of $(q_\tau)_{\tau\geq0}$ 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 into its continuous and its discrete part[34]

$$M_\tau=M_\tau^c+M_\tau^d.$$

34. 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

$$1-F(\tau)=\begin{cases}e^{-M_\tau^c}\prod_{0<s<\tau}e^{-\Delta M_s^d} & \text{if } \tau<\bar{\tau}\\ 0 & \text{if } \tau\geq\bar{\tau}\end{cases}$$

Thus, we can write

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

With this notation, at any point of continuity of the monitoring policy we have that

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

while at any point of discontinuity we have that

$$q_{\tau-} = e^{-\Delta M_\tau^d} q_\tau + (1-e^{-\Delta M_\tau^d}). \tag{D.2}$$

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

Lemma *D.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* (D.1) *and* (D.2) *with terminal condition $q_{\bar{\tau}} = 1$. Full effort is incentive compatible if and only if $q_\tau \geq \underline{q}$, for all $\tau \in [0,\bar{\tau}]$.*

To formulate the principal problem as an optimal control with state constraints, we also use the principal's continuation value, $U_\tau(\theta,\mathbf{U})$, as an additional state variable. To simplify notation, we simply write U_τ and omit the dependence on $(\theta,\mathbf{U})$. The continuation value for the principal given a monitoring policy M_τ and post-monitoring continuation payoffs $\mathbf{U}$ is

$$U_\tau = \int_\tau^\infty e^{-r(s-\tau)-(M_{s-}-M_\tau)} u(x_s^\theta)\mathrm{d}s + \int_\tau^\infty e^{-r(s-\tau)-(M_{s-}-M_\tau)} \mathcal{M}(\mathbf{U},x_s^\theta)\mathrm{d}M_s^c$$
$$+ \sum_{s>\tau} e^{-r(s-\tau)-(M_{s-}-M_\tau)}(1-e^{-\Delta M_s^d})\mathcal{M}(\mathbf{U},x_s^\theta)$$

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

$$\mathrm{d}U_\tau = \Big(rU_\tau - u(x_\tau^\theta)\Big)\mathrm{d}\tau + \Big(U_\tau - \mathcal{M}(\mathbf{U},x_\tau^\theta)\Big)\mathrm{d}M_\tau^c, \tag{D.3}$$

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

$$U_{\tau-} = (1-e^{-\Delta M_\tau^d})\mathcal{M}(\mathbf{U},x_\tau^\theta) + e^{-\Delta M_\tau^d} U_\tau. \tag{D.4}$$

Combining these definitions with Lemma D.2 allows us to represent the optimal monitoring policy recursively, with q_τ and U_τ as state variables. In particular, the optimal control problem associated with (4.3) is:

$$\begin{cases} \mathscr{G}^\theta(\mathbf{U}) = \max_{M_\tau} U_0 \\ \text{subject to} \\ \mathrm{d}U_\tau = \big(rU_\tau - u(x_\tau^\theta)\big)\mathrm{d}\tau + \big(U_\tau - \mathcal{M}(\mathbf{U},x_\tau^\theta)\big)\mathrm{d}M_\tau^c, U_{\bar{\tau}} = \mathcal{M}(\mathbf{U},x_{\bar{\tau}}^\theta) \\ U_{\tau-} = (1-e^{-\Delta M_\tau^d})\mathcal{M}(\mathbf{U},x_\tau^\theta) + e^{-\Delta M_\tau^d} U_\tau \\ \mathrm{d}q_\tau = (r+\lambda)q_\tau \mathrm{d}t - (1-q_\tau)\mathrm{d}M_\tau^c, q_{\bar{\tau}} = 1 \\ q_{\tau-} = e^{-\Delta M_\tau^d} q_\tau + (1-e^{-\Delta M_\tau^d}) \\ q_\tau \in [\underline{q},1] \end{cases} \tag{D.5}$$

Solving this problem is challenging due to the presence of state constraints. The formal proof relies on necessary conditions from Pontryagin's maximum principle for a 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 4.1. As we explained in the text, the fixed-point problem is then greatly simplified because the maximization problem is reduced to a one-dimensional problem.

The analysis of the principal problem follows five 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 4.1. Using dynamic programming to solve the principal problem is difficult because it requires solving a nonlinear 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.

We start deriving some necessary conditions for optimality using Pontryagin's maximum principle for problems with state constraints. In order to guarantee existence, we rely on 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 reparametrize 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 restricting the candidate policies to a simple family of distributions. The final solution is found maximizing over this family, which is done in equation (4.12). At this point, we only need to solve a one-dimensional optimization problem to find the optimal policy.

As it is usual in optimal control problem, we have to write the Hamiltonian for the problem. The statement of the maximum principle in Theorem 4.1 in Arutyunov et al. (2005) is quite involved, so next, we present the subset of necessary conditions in Theorem 4.1 that we will use for our analysis. Let's define the Hamiltonian $\tilde{H}(\tau)$ and the switching function $\tilde{S}(\tau)$

$$\tilde{H}(\tau) = \tilde{\zeta}_\tau \left(rU_\tau - u(x^\theta_\tau) \right) + \tilde{\nu}_\tau (r+\lambda) q_\tau \tag{D.6a}$$

$$\tilde{S}(\tau) = \tilde{\zeta}_\tau \left(U_\tau - \mathcal{M}(\mathbf{U}, x^\theta_\tau) \right) - \tilde{\nu}_\tau (1 - q_\tau), \tag{D.6b}$$

where $(\tilde{\zeta}_\tau, \tilde{\nu}_\tau)$ are the adjoint variables.[35] The Lagrange multiplier for the incentive compatibility constraint, Φ_τ, is a positive nondecreasing function satisfying

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

35. $S(\tau)$ corresponds to the function $Q(\tau)$ defined in (Arutyunov et al., 2005, p. 1816), which corresponds to the so-called switching function in linear optimal control problems.

It follows from the system of Equations (4.1) in Arutyunov et al. (2005) that at an optimal solution the adjoint variables $(\tilde{\zeta}_\tau,\, \tilde{\nu}_\tau)$ and Hamiltonian satisfy

$$\tilde{\zeta}_\tau = \tilde{\zeta}_0 - \int_0^\tau r\tilde{\zeta}_s \mathrm{d}s - \int_0^\tau \tilde{\zeta}_s \mathrm{d}M_s^c - \sum_k (1-e^{-\Delta M_{\tau_k}^d})\tilde{\zeta}_{\tau_k-} \tag{D.7a}$$

$$\tilde{\nu}_\tau = \tilde{\nu}_0 - \int_0^\tau (r+\lambda)\tilde{\nu}_s \mathrm{d}s - \int_0^\tau \tilde{\nu}_s \mathrm{d}M_s^c - \tilde{\Phi}_\tau - \sum_k (1-e^{-\Delta M_{\tau_k}^d})\tilde{\nu}_{\tau_k-} \tag{D.7b}$$

$$\tilde{H}(\tau) = \tilde{H}(0) - \int_0^\tau \tilde{\zeta}_s u'(x_s^\theta)\dot{x}_s^\theta \mathrm{d}s - \int_0^\tau \tilde{\zeta}_s \dot{x}_s^\theta (U_H - U_L)\mathrm{d}M_s^c \tag{D.7c}$$

$$- \sum_k (1-e^{-\Delta M_{\tau_k}^d})\tilde{\zeta}_{\tau_k-}\dot{x}_{\tau_k}^\theta (U_H - U_L).$$

Equations (D.7a)-(D.7c) look quite complicated; however, they correspond to the generalized integral representation of the ordinary differential equations for the co-state variables in traditional optimal control theory. Because the multipliers and maximized Hamiltonian are not necessarily absolutely continuous, we need to write the system as integral equations rather than ordinary differential equations.

The adjoint variables $\tilde{\zeta}_t$ and $\tilde{\nu}_t$ also have to satisfy the transversality conditions

$$\tilde{\zeta}_0 = -1$$

$$\tilde{\nu}_0 \le 0$$

$$\tilde{\nu}_0(q_0 - \underline{q}) = 0.$$

Finally, the optimization of the Hamiltonian requires that the following optimality and complementary slackness conditions are satisfied:

$$\tilde{S}(\tau) \le 0 \tag{D.8a}$$

$$M_\tau = \int_0^\tau \mathbf{1}_{\{S(u)=0\}} \mathrm{d}M_u. \tag{D.8b}$$

Condition (D.8a) is required for the Hamiltonian to be finite, while Condition (D.8b) states that there is positive probability of monitoring only if $S(\tau)=0$. It can be noticed that condition (D.8b) also coincide with the optimality conditions from the Hamiltonian maximization in (Seierstad and Sydsaeter, 1986, Theorem 2, p. 332).

The adjoint variables are expressed in terms of their time 0 discounted value. As it is common in the analysis of discounted optimal control problems, it is convenient to express the variables in terms of their current value counterparts. Thus, we define $\zeta_\tau \equiv e^{r\tau+M_\tau}\tilde{\zeta}_\tau$, $\nu_\tau \equiv e^{r\tau+M_\tau}\tilde{\nu}_\tau$, $H(\tau) \equiv e^{r\tau+M_\tau}\tilde{H}(\tau)$, 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

$$H(\tau) = \zeta_\tau \left(rU_\tau - u(x_\tau^\theta)\right) + \nu_\tau (r+\lambda) q_\tau \tag{D.9a}$$

$$S(\tau) = \zeta_\tau \left(U_\tau - \mathcal{M}(\mathbf{U}, x_\tau^\theta)\right) - \nu_\tau (1-q_\tau). \tag{D.9b}$$

Similarly, we can write the current value counterpart for the Lagrange multiplier, which is defined as

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

The previous equation can be inverted to get

$$\tilde{\Phi}_\tau = \Phi_0 + \int_0^\tau e^{-rs-M_s} \mathrm{d}\Phi_s.$$

Rewriting equations (D.7a)-(D.7b) we get

$$e^{-r\tau-M_\tau}\zeta_\tau=\zeta_0-\int_0^\tau re^{-rs-M_s}\zeta_s\mathrm{d}s-\int_0^\tau e^{-rs-M_s}\zeta_s\mathrm{d}M_s^c-\sum_k(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\zeta_{\tau_k-} \tag{D.10a}$$

$$e^{-r\tau-M_\tau}\nu_\tau=\nu_0-\int_0^\tau(r+\lambda)e^{-rs-M_s}\nu_s\mathrm{d}s-\int_0^\tau e^{-rs-M_s}\nu_s\mathrm{d}M_s^c-\Phi_0-\int_0^\tau e^{-rs-M_s}\mathrm{d}\Phi_s \tag{D.10b}$$
$$-\sum_k(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\nu_{\tau_k-}.$$

Equation (D.10a) implies that ζ_τ must be constant, and together with the transversality condition we get that $\zeta_\tau=\zeta_0=-1$. It can be verified form equation (D.10b) that the adjoint variable ν_τ must satisfy

$$\nu_\tau=\nu_0-\int_0^\tau\lambda\nu_s\mathrm{d}s-\Phi_\tau. \tag{D.11}$$

Equation (D.11) corresponds to the traditional differential equation for the adjoint variable (Seierstad and Sydsaeter, 1986, Equation (91) in Theorem 2, p. 332). Next, equation (D.7c) implies that at any jump time τ_k we must have

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

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

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

Combining the previous two equation and using the definition of the current value variables we get that

$$\begin{aligned}(1-e^{-\Delta M^d_{\tau_k}})e^{-r\tau_k-M_{\tau_k-}}\dot{x}^\theta_{\tau_k}(U_H-U_L)&=-e^{-r\tau_k-M_{\tau_k}}\left(rU_{\tau_k}-u(x^\theta_{\tau_k})\right)+e^{-r\tau_k-M_{\tau_k}}(\nu_{\tau_k-}-\Delta\Phi_{\tau_k})(r+\lambda)q_{\tau_k}\\&\quad+e^{-r\tau_k-M_{\tau_k-}}\left(rU_{\tau_k-}-u(x^\theta_{\tau_k-})\right)-e^{-r\tau_k-M_{\tau_k-}}\nu_{\tau_k-}(r+\lambda)q_{\tau_k-}\\&\Longrightarrow\\(1-e^{-\Delta M^d_{\tau_k}})e^{\Delta M^d_{\tau_k}}\dot{x}^\theta_{\tau_k}(U_H-U_L)&=-rU_{\tau_k}+u(x^\theta_{\tau_k})+(\nu_{\tau_k-}-\Phi_{\tau_k})(r+\lambda)q_{\tau_k}\\&\quad+e^{\Delta M^d_{\tau_k}}rU_{\tau_k-}-e^{\Delta M^d_{\tau_k}}u(x^\theta_{\tau_k-})-e^{\Delta M^d_{\tau_k}}\nu_{\tau_k-}(r+\lambda)q_{\tau_k-}\\&\Longrightarrow\\(e^{\Delta M^d_{\tau_k}}-1)\dot{x}^\theta_{\tau_k}(U_H-U_L)&=r(e^{\Delta M^d_{\tau_k}}U_{\tau_k-}-U_{\tau_k})-u(x^\theta_{\tau_k})\left(e^{\Delta M^d_{\tau_k}}-1\right)\\&\quad-\nu_{\tau_k-}(r+\lambda)\left(e^{\Delta M^d_{\tau_k}}q_{\tau_k-}-q_{\tau_k}\right)-\Delta\Phi_{\tau_k}(r+\lambda)q_{\tau_k}.\end{aligned} \tag{D.12}$$

Substituting the expressions for the jump in U_τ and q_τ, which are 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

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

Simplifying terms we get the following necessary condition that must be satisfied at any time τ_k in which there is an atom:

$$r\mathcal{M}(\mathbf{U},x^\theta_{\tau_k})=u(x^\theta_{\tau_k})+\dot{x}^\theta_{\tau_k}(U_H-U_L)+(r+\lambda)\nu_{\tau_k}+(r+\lambda)\frac{q_{\tau_k}}{e^{\Delta M^d_{\tau_k}}-1}\Delta\Phi_{\tau_k}. \tag{D.13}$$

We show in the proof of Lemma D.4 that $\Delta\Phi_{\tau_k}=0$, which means that ν_τ is continuous. This means that (D.13) can be simplified to

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

This expression coincides with the optimality condition at free time $\bar{\tau}$ in Equation (4.4) in Arutyunov et al. (2005), and this condition also coincides with the condition for free final time problems in (Seierstad and Sydsaeter, 1986, Equation (152) in Theorem 16, p. 398).

We can summarize the necessary condition that we use in the proof:

$$S(\tau)=-\left(U_\tau-\mathcal{M}(\mathbf{U},x^\theta_\tau)\right)-\nu_\tau(1-q_\tau)\leq 0 \tag{D.14}$$

$$M_\tau=\int_0^\tau \mathbf{1}_{\{S(u)=0\}}\mathrm{d}M_u \tag{D.15}$$

$$\nu_\tau=\nu_0-\int_0^\tau \lambda\nu_s\mathrm{d}s-\Phi_\tau \tag{D.16}$$

$$\Phi_\tau=\int_0^\tau \mathbf{1}_{\{q_u=\underline{q}\}}\mathrm{d}\Phi_u,\quad \mathrm{d}\Phi_\tau\geq 0 \tag{D.17}$$

$$r\mathcal{M}(\mathbf{U},x^\theta_{\tau_k})=u(x^\theta_{\tau_k})+\dot{x}^\theta_{\tau_k}(U_H-U_L)+(r+\lambda)\nu_{\tau_k}+(r+\lambda)\frac{q_{\tau_k}}{e^{\Delta M^d_{\tau_k}}-1}\Delta\Phi_{\tau_k}. \tag{D.18}$$

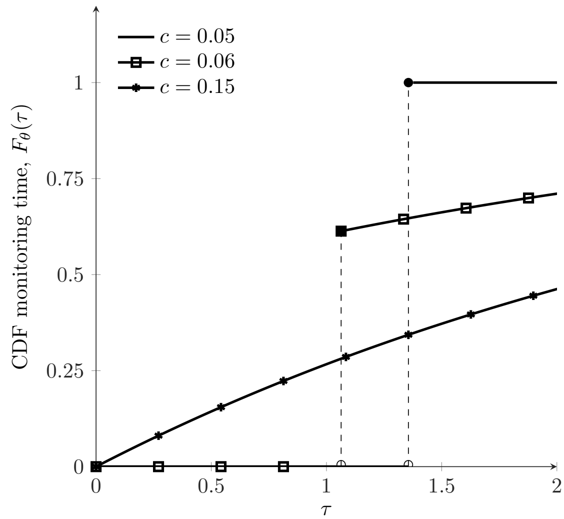
The next step is to show that, in any optimal policy, the principal never monitors using a positive hazard rate if the incentive compatibility constraint is slack. If the incentive compatibility constraint is binding over a period of time, then q_τ is constant and the constant monitoring rate is determined by the condition that $q_\tau=\underline{q}$. On the other hand, if the incentive compatibility constraint were slack and $m_\tau>0$, then the necessary condition (D.8b) would require that $S(\tau)=0$. In the following lemma, we show that this leads to a contradiction due to the convexity of $u(x)$, which means that the monitoring rate must be zero in this case.

Lemma D.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, $\int_B \mathrm{d}M^{c*}_\tau=0$.*

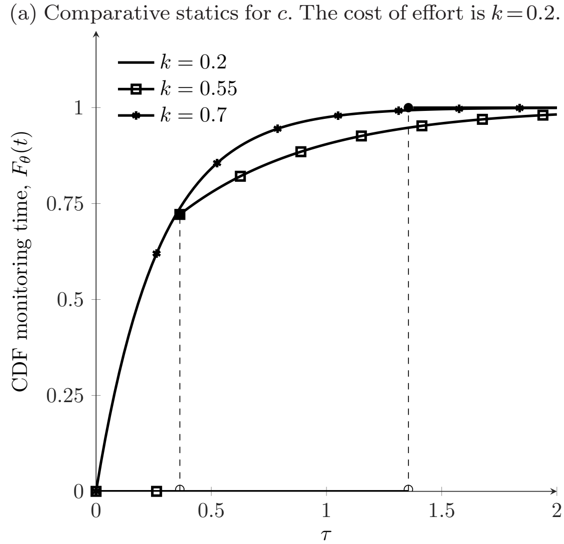
Lemma D.3 provides a partial characterization of the continuous part of the monitoring distribution. However, we also need to take care of the atoms. Relying on the convexity of $u(x)$ once again, we show that Equation (D.13) satisfies a single crossing condition implying that (D.13) can hold at most at one point in the optimal path of q_τ, which means that any optimal policy can have at most one atom. Formally, we prove the following result:

Lemma D.4. *Let M^*_τ be an optimal policy, M^{d*}_τ its discrete part. Then, there is at most one time $\hat{\tau}$ such that $\Delta M^{d*}_{\hat{\tau}}>0$.*

The final step is to verify that these results imply that any optimal policy must take the form in Theorem 4.1. Any policy consistent with Lemmas D.3 and D.4 must look as the one in Figure A1a, and the trajectory of q_τ must look like the one in Figure A1b: that is, 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 after that. As it is shown in Figure A1a, the monitoring policy associated with the trajectory of q_τ is such the incentive compatibility constraint is binding before time $\tilde{\tau}$, which requires a monitoring rate equal to m^* (where m^* is the same as in Proposition 4.5). 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. By means of optimizing over $\tilde{\tau}$, for an arbitrary fixed $\hat{\tau}$, we can show that $\tilde{\tau}$ is either zero or infinity, which allow us to conclude that the optimal policy must take the form in Theorem 4.1 and completes the proof.



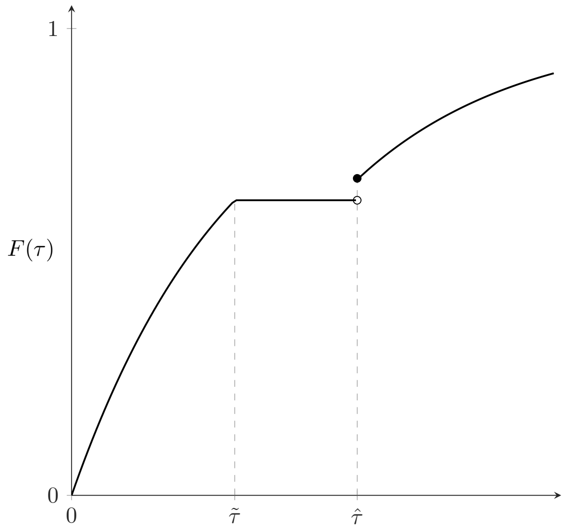
(a) Comparative statics for c. The cost of effort is $k = 0.2$.



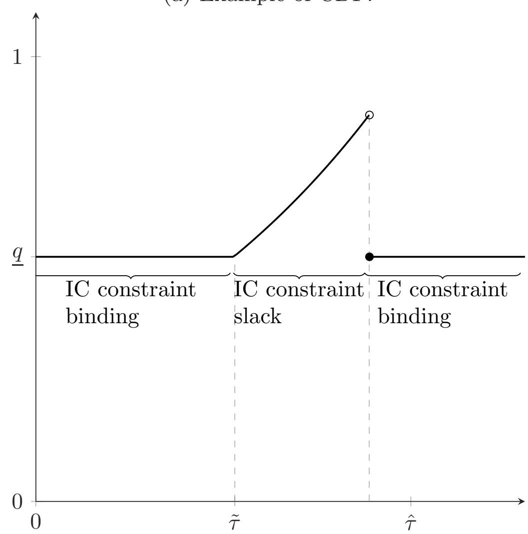
(b) Comparative statics for k. The cost of monitoring is $c = 0.05$.

Figure 2

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.



(a) Example of CDF.



(b) Path of q_τ conditional on not monitoring.

FIGURE A1

Cumulative density function and path of q_τ implied by Lemmas D.3 and D.4.