Theoretical Economics 19 (2024), 823–863 1555-7561/20240823

Relational enforcement

PETER ACHIM
Department of Economics, University of York

JAN KNOEPFLE
School of Economics and Finance, Queen Mary University of London

A principal incentivizes an agent to maintain compliance and to truthfully announce any breaches of compliance. Compliance is imperfectly controlled by the agent's private effort choices, is partially persistent, and is verifiable by the principal only through costly inspections. We show that in principal-optimal equilibria, the principal enforces maximum compliance using deterministic inspections. Periodic inspection cycles are suspended during periods of self-reported noncompliance, during which the agent is fined. We show how commitment to random inspections would benefit the principal, and discuss possible ways for the principal to overcome her commitment problem.

KEYWORDS. Relational contracts, dynamic enforcement, persistence, costly inspections.

JEL CLASSIFICATION. C73, D83.

1. INTRODUCTION

In 2018 and 2019 two plane crashes killed 346 people and led to a worldwide grounding of the Boeing 737 MAX.[1] An investigation by the U.S. Congress concluded that the accidents were to a large extent due to "grossly insufficient oversight by the FAA" (Federal Aviation Administration).[2] Starting from the early 2000s, the FAA had increasingly trusted manufacturers to certify their own planes to save costs. By 2018, Boeing had self-certified nearly all of its work (Kitroeff, Gelles, and Nicas (2019)). Boeing rushed the development of the 737 MAX at the expense of safety. This case illustrates the risk in relying on self-reported quality assurances without sufficient oversight.

Peter Achim: peter.wagner@york.ac.uk
Jan Knoepfle: j.knoepfle@qmul.ac.uk

We are thankful to Francesc Dilmé, Daniel Hauser, Florian Hoffmann, Martin Pollrich, Sven Rady, and Alex Smolin for valuable comments, and thank participants at the Canadian Economic Theory Meeting in Vancouver and the Econometric Society Summer Meeting in St Louis in 2017, as well as seminar participants at Bonn and HECER. Achim is grateful to Chris Shannon and the Economics Department at UC Berkeley for their hospitality during a productive visit in the fall of 2016 and to the Hausdorff Center for Mathematics in Bonn for providing financial support. Knoepfle acknowledges financial support by the German Research Foundation through CRC TR224-B04 and from the Academy of Finland (Project 325218).

[1] As a result, Boeing suffered an operational loss of over $20 billion. The estimated impact on the U.S. economy as a whole was a 0.4 percentage points loss in gross domestic product (GDP) growth (di Giovanni et al. (2020)).

[2] See U.S. House of Representatives (2020).

 https://doi.org/10.3982/TE5183

In this paper, we study enforcement relationships in which the agent privately controls and observes the state of compliance and makes reports to a principal without commitment power. Compliance is partially persistent over time and can be observed by the principal only through costly inspections. The principal schedules inspections and imposes fines to incentivize the agent to exert effort and to self-report instances of noncompliance. We show that the principal can induce the agent to exert full effort and report truthfully at all times through relational incentives. The principal carries out inspections despite knowing the result beforehand. Our analysis highlights the importance of the persistent effect of effort. Further, the principal cannot gain from randomized inspections when she lacks commitment, but random inspections would be optimal with commitment.

Public-sector applications of our model include banking supervision to ensure that banks maintain functioning internal risk assessments[3] and environmental protection where the corresponding government agency ensures the enforcement of regulation by firms.[4] Similarly, private-sector organizations must ensure internally that employees follow regulations.[5]

We consider principal-optimal equilibria in which the agent truthfully discloses all instances of noncompliance and exerts maximum effort throughout. The principal-optimal equilibrium we derive in our main result (Theorem 1) entails two phases: a monitoring phase and a penalty phase. The agent is in the monitoring phase when he reports compliance. During the monitoring phase, the agent is not fined, but is subject to periodic inspections that would result in the maximal possible fine in the off-path event that the inspection revealed misreporting. The agent is in the penalty phase when he reports noncompliance. He pays a constant flow fine, but is never inspected. He also pays a lump sum fine each time the state transitions from compliance to noncompliance. Crucially, this transition fine features penalty reductions for early disclosures of noncompliance, an aspect that is consistent with voluntary disclosure schemes commonly used in practice. The penalty reduction prevents the agent from delaying a report of an incidence of noncompliance in the hope that he can regain compliance before the next inspection.[6]

Notably, inspection times in this equilibrium are entirely predictable for the agent, which implies that the principal cannot gain from randomized inspections. Intuitively,

[3]See Section 5 for a brief discussion of banking supervision practices in Germany.

[4]For the United States, Blundell, Gowrisankaran, and Langer (2020) measure the benefits of dynamic procedures used by the EPA.

[5]For instance, the European Commission (2019) supports exporting firms in elaborating internal compliance programs(ICP) to "mitigate risks associated with dual-use trade controls and to ensure compliance" internally. Dual-use goods have civil and military applications and fall under special regulation to promote international security, e.g., by "countering risks associated with the proliferation of Weapons of Mass Destruction" (European Commission (2019, p. 17)).

[6]Blundell, Gowrisankaran, and Langer (2020) point out that when determining the gravity of fines, the EPA takes into account whether a violation was self-reported or not. See also Kapon (2022), who studies optimal design of fine reductions (amnesties) granted for self-reports of illegal activity when detections arrive at an exogenous rate. Focusing on deterministic fine reduction paths, Kapon (2022) also finds a cyclical structure of the optimal mechanism.

the principal"s motive to inspect is derived from her desire to maintain a reputation for vigilance.[7] Predictable inspections provide the strongest incentive for the principal to inspect. As long as the principal inspects as prescribed by her equilibrium strategy, the agent continues to expect to be monitored and, thus, has an incentive to exert effort and report truthfully. However, when the principal delays inspections in a way that is detectable by the agent, then the agent will infer that the principal has become nonvigilant. This in turn induces the agent to shirk, which ultimately leads to a breakdown of the relationship that is costly for the principal. If the principal uses a random strategy and mixtures are unobservable for the agent, deviations by the principal are harder to detect for the agent. This destroys any potential benefit for the principal in equilibrium.

We exploit the optimality of predictable inspection schedules for the construction of the principal-optimal equilibrium in Theorem 1: her equilibrium payoffs coincide with the value of an auxiliary mechanism-design problem in which the principal is restricted to nonrandom inspections. We then transform this optimization into a dynamic programming problem that uses the agent's promised utility as a state variable.

Comparative statics reveal the importance of persistence for relational enforcement. In equilibrium, the persistent effect of effort on compliance allows the principal to deter the agent from deviating through isolated inspections. As the state's persistence vanishes, the inspection costs necessary to enforce compliance grow arbitrarily large.

We then contrast the relational enforcement equilibrium with stochastic inspection mechanisms. The ability to commit to random inspections decreases the principal's inspection costs relative to the deterministic inspections that are required in the noncommitment case. Deterministic inspections are more costly because of delay and noise in the compliance process, and due to the transition penalties that are needed to generate incentives for voluntary disclosure. Comparative statics highlight the contrast between relational enforcement and the commitment case with random inspections. As the persistence of the state of compliance vanishes, the random inspection costs decrease monotonically. We also discuss ways to overcome the principal's commitment problem, including institutional separation of planning and execution of oversight and inspection sampling combined with publicly accessible and verifiable records.

The rest of the paper is organized as follows. After discussing related literature, the model setup is presented in Section 2. Section 3 characterizes the agent's incentive constraints, shows that the principal-optimal equilibrium can be determined by solving an auxiliary mechanism-design problem, and outlines how to solve the auxiliary problem. We present the principal-optimal equilibrium in Section 4, followed by comparative statics. Section 5 discusses random inspections. All proofs are contained in the Appendix.

[7] Here, "maintaining a reputation" means following equilibrium actions because deviating leads to a less favorable continuation value (see Chapter 22 in Ljungqvist and Sargent (2018)). This "history-dependence" notion of reputation is distinct from the "adverse-selection" approach to reputation (Mailath and Samuelson (2006, p. 459)), in which incentives stem from the desire to convince the opponent that you are of a specific type.

Related literature Our paper is closely related to the literature on costly state verification (CSV). Early papers, including Townsend (1979), Gale and Hellwig (1985), Mookherjee and Png (1989), and Border and Sobel (1987), focus on one-shot interactions. One of the main findings in this literature is the optimality of cutoff verification protocols, an insight that has been influential in explaining the use of debt contracts and the role of financial intermediaries. A number of papers consider dynamic extensions. In many of these, the principal's observation reveals the agent"s current private information with no intertemporal link to past actions or states.[8] By contrast, the state in our model is partially persistent, so inspections reveal information about past behavior.

Inspections of a persistent state are analyzed in Ravikumar and Zhang (2012) and Kim (2015). These papers study pure adverse-selection problems with exogenous private information when the principal has commitment. In Ravikumar and Zhang (2012), the contracting friction is driven by risk-sharing concerns. They find that random inspections are optimal, and, after each inspection, there is a grace period without inspections. In Kim (2015), the contracting friction is driven by the agent's limited liability. They find that random inspections are optimal for incentive provision when truthful disclosure is attainable, but periodic inspections are optimal to guide environmental protection activities when the fines are insufficient to attain truth-telling. Our setting features an adverse-selection and moral-hazard problem, the principal lacks commitment power, and the agent is risk-neutral so that the contracting friction stems from limited liability. We find deterministic inspections are optimal when the principals lacks commitment. Our result for the commitment case is in line with their findings that random inspections provide incentives more effectively.

Most closely related is the paper by Varas, Marinovic, and Skrzypacz (2020), which studies a pure moral-hazard model with full commitment and without fines.[9] In their model, the agent is incentivized by the desire to maintain a good reputation and inspections make the agent's type public. Additionally, inspections serve an information-acquisition purpose for the principal. The authors find random inspections are optimal for incentive provision, but deterministic inspections are optimal for information acquisition. In contrast, in our model, the agent discloses the state of compliance, so that inspections do not reduce the uncertainty about the state. Ball and Knoepfle (2023) study

[8]For dynamic moral-hazard problems in which monitoring reveals the current action, see Antinolfi and Carli (2015), Piskorski and Westerfield (2016), Dilmé and Garrett (2019), Chen, Sun, and Xiao (2020), Li and Yang (2020), Dai, Wang, and Yang (2022), Wong (2022). For dynamic adverse-selection problems in which verification reveals the agent's current information that is independent and identically distributed (i.i.d.) across periods, see Chang (1990), Webb (1992), Monnet and Quintin (2005), Wang (2005), Popov (2016), Malenko (2019).

[9]In both papers, state transitions are based on the reputation for quality model (Board and Meyer-ter-Vehn (2013)). In Board and Meyer-ter-Vehn (2013) quality becomes publicly observable at random times. In the present paper and in Varas, Marinovic, and Skrzypacz (2020), the principal chooses the times at which the state becomes publicly observable at a cost. In Halac and Prat (2016) and Dilmé and Garrett (2019), the principal invests in building her persistent monitoring capabilities, and monitoring reveals information about current actions of the agents. In contrast to the setup in the present paper, the principal's actions are private and she cannot perfectly control the time at which she signals vigilance. In Halac and Prat (2016) this leads to a breakdown of the relationship with positive probability after the agent's effort remains unrecognized for too long.

optimal inspections with commitment and show that random inspections are optimal for incentives when the agent must avoid a breakdown and deterministic inspections are optimal when the agent must achieve a breakthrough. The driver of nonrandom inspections in our paper is the principal's lack of commitment.

The persistent effect of effort is important for relational incentives. This is also highlighted for a collaboration problem without commitment in Ramos and Sadzik (2023). Similar to our comparative statics in Section 4.2, the authors show that relational incentives vanish without persistence. Without persistence, commitment is crucial for enforcement with costly inspections: when the agent is expected to comply, the principal has no incentive to pay the inspection cost to reveal information she already knows. Indeed, Reinganum and Wilde (1985) confirm for a nonrepeated setting that compliance is not achievable without full commitment. With repeated interactions, continuation play can provide punishment for insufficient inspection. Ben-Porath and Kahneman (2003) prove a folk theorem, showing that full compliance can be obtained without commitment in the undiscounted limit. In our game, full compliance is attainable even with discounting. This difference stems from the persistence of the state and the observability of inspections by the agent in our model.

2. Model

Players, actions, and state dynamics There are an agent and a principal. Time $t \in [0, \infty)$ is continuous. The agent, at each instant t, privately chooses effort $\eta_t \in [0, 1]$ to comply with exogenously given regulation as best he can. The state of compliance at time t is $\theta_t \in \{0, 1\}$, where we refer to state 0 as *noncompliant* and to state 1 as *compliant*. Effort affects the transitions of the process $\{\theta_t\}_{t\geq 0}$: there are parameters $\lambda > 0$ and $\alpha \in (0, 1)$ such that the state changes from 0 to 1 at Poisson rate $\eta_t\lambda\alpha$ and from 1 to 0 at rate $\lambda(1 - \eta_t\alpha)$. We may interpret λ and α as follows. There is a Poisson process of shocks arriving at rate λ. Whenever there is a shock at time t, the resulting state is $\theta_t = 1$ with probability $\eta_t\alpha$ and it is $\theta_t = 0$ with probability $1 - \eta_t\alpha$; between shocks the state remains unchanged. Thus, λ measures the variability of compliance and α measures the responsiveness to the agent's effort conditional on a shock; $\alpha < 1$ implies that the agent cannot always maintain compliance despite his best efforts. The agent observes θ_t at all times and sends report $\hat{\theta}_t \in \{0, 1\}$ to the principal. The agent can exit the relationship unilaterally at any time.

The principal chooses inspections and fines to incentivize the agent. We denote by N_t^I the cumulative number of inspections and by F_t the cumulative fines up to and including time t. That is, $\mathrm{d}N_t^I \equiv N_t - \lim_{s\nearrow t} N_s^I \in \{0, 1\}$ is equal to 1 if and only if there is an inspection at time t and $\mathrm{d}F_t \geq 0$ is the fine paid by the agent at time t.

Information and timing The agent observes the history of all paths

$$h_t = \left\{\eta_s, \theta_s, \hat{\theta}_s, N_s^I, F_s\right\}_{s\in[0,t]}.$$

The principal never observes the agent's effort and is able to observe the state θ_t only by performing an inspection at time t. To allow for randomized inspections, we equip

the principal with a private random signal π, defined on a sufficiently rich sample space Π. The principal observes histories of the form $h_t^P = \{\pi, \hat{\theta}_s, N_s^I, F_s, \theta_s \colon N_s^I = 1\}_{s\in[0,t]}$. Heuristically, we can describe the timing of events within each instant $[t, t+\mathrm{d}t)$ as follows.[10] First, the agent chooses effort η_t. Subsequently, nature determines whether a shock arrives and, conditional on the arrival of a shock and the effort, draws a new state θ_t. The agent then observes the realized θ_t and sends a report $\hat{\theta}_t \in \{0, 1\}$ to the principal. The principal chooses whether to inspect, $\mathrm{d}N_t^I \in \{0, 1\}$, and sets a fine $\mathrm{d}F_t$ incurred immediately by the agent, where the fine can be contingent on the true state θ_t if and only if the principal chose to inspect.

Payoffs and equilibrium The principal and the agent are risk-neutral and discount future payoffs at a common rate $r > 0$. The principal is tasked with ensuring that the agent complies with the regulation. She incurs a lump-sum cost $\kappa > 0$ from each inspection. For a realized history $h = \{\eta_t, \theta_t, \hat{\theta}_t, N_t^I, F_t\}_{t\in[0,\infty)}$, the discounted net present cost of the principal at time t is

$$k_t = \int_t^\infty e^{-r(s-t)}\kappa\,\mathrm{d}N_s^I. \tag{1}$$

The principal does not benefit directly from compliance or from fining the agent. Fines are interpreted as remedial actions that negatively impact the agent.[11] To ensure that the principal is willing to bear the inspection costs, assume that when the relationship breaks down, i.e., the agent exits or ceases to exert effort, the principal suffers cost $\bar{K}$. We assume throughout that the bound $\bar{K}$ is large enough such that it exceeds the expected inspection costs necessary to incentivize the agent.[12]

The agent incurs effort cost of $c\eta_t\,\mathrm{d}t$ with $c > 0$ and disutility $\mathrm{d}F_t$ from fines. His discounted net present payoff at time t is given by

$$u_t = \int_t^\infty e^{-r(s-t)}(-c\eta_s\,\mathrm{d}s - \mathrm{d}F_s). \tag{2}$$

The agent is protected by limited liability. If he chooses to exit, the relationship ends permanently, which results in a continuation payoff of $-B$. This implies a constraint on the severity of fines the principal can impose. We assume that the exogenously given

[10] We outline the sequentiality at a given instant to give an intuition about the order of moves. Formally, the order is captured by continuity properties of the respective action and state paths. It is well known that in continuous-time games with observable actions, strategies may not produce well defined action paths. To focus the exposition in the main text on the main economic forces, we defer a more formal treatment to Supplemental Appendix A (available at http://econtheory.org/supp/5183/supplement.pdf), where we adopt an approach by Kamada and Rao (2023) to impose restrictions on strategies that guarantee well defined action paths.

[11] Our results do not rely on this assumption. When the principal benefits from fines, her preferred equilibrium differs from the one we present only in an initial fine paid by the agent (see Section 6).

[12] This serves as a concise way to deliver incentives for inspection for the principal when analyzing the equilibrium problem without commitment. Alternatively, we could explicitly incorporate an (unobserved) flow reward $\theta_t R$ or $\eta_t R$ in our model so that, for $R > 0$ large enough, the principal's expected payoff from inducing effort by the agent exceeds the necessary inspection costs. Our results would be unaffected.

bound B is large enough: $B > \bar{B} \equiv c(r+\lambda)/(\lambda\alpha r)$. Otherwise, the maximal punishment is insufficient to incentivize effort even if θ_t were public at all times.

Given a strategy profile, the principal and the agent form expectations about history h based on their past observations. For strategies that induce measurable action processes on path, we denote the expected cost of the principal and payoff of the agent at time t by

$$K_t = \mathbb{E}^P_{t-}[k_t] \quad \text{and} \quad U_t = \mathbb{E}^A_{t-}[u_t].$$

The expectation is with respect to the process $\{\theta_s\}_{s\in[0,\infty)}$ and the randomization device π, and it is conditional on the information that is available to the principal and the agent, respectively.[13] In continuous-time games with observable actions and stochastic environments, players' behavior may be nonmeasurable. We do not impose restrictions on strategies that rule out nonmeasurable behavior. Instead, our equilibrium definition below requires that strategies lead to measurable actions on path. Histories away from the equilibrium path may lead to nonmeasurability. Payoffs at such histories can be assigned freely within the feasible bounds. In our game, the lower bounds on payoffs can be reached by either player unilaterally through exit or by imposing the maximal fine. Therefore, potential nonmeasurabilities off path and the assigned payoffs cannot be used as a threat to enlarge the equilibrium set (see also the discussion of this approach in Kamada and Rao (2023)).

We define a strategy profile, together with processes $\{K_t, U_t\}_{t\geq 0}$, to be a perfect Bayesian equilibrium if the following statements hold.

- The strategies of the principal and the agent are sequentially rational.
- Along the equilibrium path, K_t and U_t are equal to the conditional expectations given above. Away from the equilibrium path, K_t and U_t are equal to the conditional expectations whenever these are well defined.
- At all histories and all times, $K_t \in [0, \bar{K}]$ and $U_t \in [-B, 0]$.

We say that the agent's strategy is *truthful* if $\hat{\theta}_t = \theta_t$ at all histories along the equilibrium path. Further, we call the agent's strategy *maximally compliant* if $\eta_t = 1$ at all histories along the equilibrium path. Note that with full effort by the agent, the probability of compliance at any given time is maximized. We refer to an equilibrium as truthful or maximally compliant if the agent's strategy in this equilibrium has the respective property. Henceforth, we restrict attention to such equilibria (see the discussion in Section 6).

[13] For the principal, the expectation is with respect to the natural filtration generated by the process $\{N^I_s, F_s, \theta_s : \mathrm{d}N_s = 1\}_{s\in[0,t)} \cup \{\hat{\theta}_s\}_{s\in[0,t]}$ when taking her inspection decision, and with respect to the natural filtration generated by $\{\hat{\theta}_s, N^I_s, F_{s-}, \theta_s : \mathrm{d}N_s = 1\}_{s\in[0,t]}$ for the fine. For the agent, the expectation is with respect to the natural filtration generated by the process $\{\eta_s, \theta_s, \hat{\theta}_s, N^I_s, F_s\}_{s\in[0,t)}$ for his effort choice, and with respect to the natural filtration generated by $\{\hat{\theta}_s, N^I_s, F_s\}_{s\in[0,t)} \cup \{\eta_s, \theta_s\}_{s\in[0,t]}$ for his report. As mentioned above, see Supplemental Appendix A for a formal treatment of permissible strategies.

We say that inspections are *predictable* for the agent if he knows for certain whether or not his current report will lead to an inspection at any history.[14] Henceforth, we refer to inspections as random whenever they are nonpredictable for the agent.

3. Agent's and principal's problem

3.1 *Agent: Incentive compatibility*

Fix an arbitrary principal strategy of fines and inspections, and let U_t be the agent's associated expected discounted continuation payoff at time t under the assumption that he exerts full effort and reports truthfully. We characterize recursively the conditions under which truthful reporting and maximal compliance are a best response for the agent in terms of the evolution of his promised utility at all times. Due to the persistence in the agent's private information, the recursive characterization of incentive compatibility requires tracking two state variables:[15] the agent's expected continuation utility given that $\theta_t = 0$ and given that $\theta_t = 1$. Formally, fix a principal strategy and define for any strict history at time t,

$$U_t^0 = \mathbb{E}_{t-}^A[U_t|\theta_t = 0] \quad \text{and} \quad U_t^1 = \mathbb{E}_{t-}^A[U_t|\theta_t = 1]. \tag{3}$$

These are the agent's expected continuation utilities when history h_{t-} is followed by the realization of $\theta_t = 0$ or $\theta_t = 1$. Here, $\mathbb{E}_{t-}^A$ represents the expectation conditional on all available information before time t. Following Zhang (2009), we call U_t^1 the *persistent payoff* if $\theta_{t-} = 1$ and the *transitional payoff* in case $\theta_{t-} = 0$, and vice versa for U_t^0.

Our first result provides a complete characterization of the agent's incentive-compatibility constraints in terms of the evolutions of U_t^0 and U_t^1. The construction is based on the martingale representation for marked point processes (Last and Brandt (1995)), which is presented in detail in Appendix A. We exploit the fact that the agent's time-t expectation of his total discounted lifetime utility is a martingale. For the inspection counting process N^I, the *compensator* is a predictable process $\nu^I = \{\nu_t^I\}_{t\geq 0}$ such that the compensated process $N_t^I - \nu_t^I$ is a martingale. The compensator exists under very general conditions and can be interpreted as the predictable drift of the underlying (nonpredictable) stochastic process. We can think of the compensator as a generalization of the cumulative hazard function, and, consequently, think of $\mathrm{d}\nu^I/\mathrm{d}t$ as the hazard rate of inspections (whenever it exists). Furthermore, let the predictable process $\Delta^I = \{\Delta_t^I\}_{t\geq 0}$ measure the jump in the persistent payoff if an inspection is performed at time t.[16]

Lemma 1. *A principal's strategy induces maximal compliance and truthful reporting if and only if it generates the processes* $\{U_t^1, U_t^0\}_{t\geq 0}$ *of promised utilities satisfying for* $i = \theta_{t-}$ *and* $j = 1 - \theta_{t-}$, *and at all* t *with* $\mathrm{d}N_t^I = 0$ *and* $\theta_{t-} = \theta_t$,

[14] Formally, predictability means that the process N^I is measurable with respect to the information available to the agent (see Davis (1993, p. 67, for a definition in the context of jump processes)).

[15] This is based on Fernandes and Phelan (2000), who introduce a recursive approach with serially correlated states in discrete time. See Zhang (2009) for a treatment in continuous time.

[16] That is, given that an inspection occurs at time t (and $\theta_{t-} = 1$), then $\Delta_t^I = U_t^1 - U_{t-}^1$, supposing that the inspection confirms that the agent reported truthfully, which is the case along the equilibrium path.

$$\textit{(Pk)}\quad \mathrm{d}U_t^i = rU_t^i\,\mathrm{d}t + \lambda(i-\alpha)(U_t^1 - U_t^0)\,\mathrm{d}t + c\,\mathrm{d}t + \mathrm{d}F_t - \Delta_t^I\,\mathrm{d}\nu_t^I$$

$$\textit{(H)}\quad \mathrm{d}U_t^j \le rU_t^j\,\mathrm{d}t + \lambda(j-\alpha)(U_t^1 - U_t^0)\,\mathrm{d}t + c\,\mathrm{d}t + \mathrm{d}F_t + (B + U_t^j)\,\mathrm{d}\nu_t^I$$

$$\textit{(O)}\quad U_t^1 - U_t^0 \ge c/\lambda\alpha$$

$$\textit{(P)}\quad U_t^0, U_t^1 \in [-B, 0],$$

We now explain the role of the four constraints: promise keeping (Pk), honesty (H), obedience (O), and participation (P).

First, the promise-keeping constraint (Pk) ensures that the agent's expectation of his discounted lifetime utility is indeed a martingale, so that $U_t^{\theta_{t-}}$ represents the continuation utilities consistently. For illustration, suppose $\theta_{t-} = 1$ and rearrange (Pk) for $i = 1$:

$$rU_t^1\,\mathrm{d}t = -c\,\mathrm{d}t - \mathrm{d}F_t + \lambda(1-\alpha)\big(U_t^0 - U_t^1\big)\,\mathrm{d}t + \mathrm{d}\nu_t^I\Delta_t^I + \mathrm{d}U_t^1.$$

This formulation has the familiar asset price interpretation where the return rU_t^1 is equal to the current flow payoff (dividends) plus expected capital gains. The agent incurs flow cost $c\,\mathrm{d}t$ from effort $\eta_t = 1$ and suffers the fine $\mathrm{d}F_t$. With full effort, there is a transition from state 1 to 0 with probability $\lambda(1-\alpha)\,\mathrm{d}t$ at which the agent's payoff changes by $(U_t^0 - U_t^1)$, an inspection arrives with probability $\mathrm{d}\nu_t^I$ and changes the agent's payoff by Δ_t^I, and $\mathrm{d}U_t^1$ is the change in the current payoff if no transition or inspection arrives.

Second, the honesty constraint (H) makes sure that the agent truthfully reports any state transitions immediately.[17] For a heuristic illustration, suppose again that $i = \theta_{t-} = 1$ and consider the agent's reporting incentives when a transition to state 0 occurs at time t. For exposition, assume also that the inspection distribution has no mass point at time t and that the density is $\tilde{\nu}_t^I = \lim_{\mathrm{d}t\to 0}\mathrm{d}\nu_t^I/\mathrm{d}t$. The agent is willing to report the decline without delay only if he cannot gain from misreporting $\hat{\theta} = 1$ for a small interval $[t, t+\mathrm{d}t)$ and reverting to truth-telling afterward. Using a first-order Taylor approximation, we must have

$$U_t^0 \ge -c\,\mathrm{d}t - \mathrm{d}F_t + \lambda\alpha U_t^1\,\mathrm{d}t + \tilde{\nu}_t^I\,\mathrm{d}t(-B) + \big(1 - \lambda\alpha\,\mathrm{d}t - \tilde{\nu}_t^I\,\mathrm{d}t\big)(1 - r\,\mathrm{d}t)U_{t+\mathrm{d}t}^0.$$

On the left-hand side we have the value from reporting truthfully. On the right-hand side, the agent incurs $c\,\mathrm{d}t$ and $\mathrm{d}F_t$. With probability $\lambda\alpha\,\mathrm{d}t$, the state changes back from 0 to 1 and the agent gets U_t^1. With probability $\tilde{\nu}_t^I\,\mathrm{d}t$, the agent is inspected and caught misreporting; by standard arguments, it is optimal for the principal to enforce the most severe punishment. This leaves the agent a payoff equal to his outside option $-B$. With probability $(1 - \lambda\alpha\,\mathrm{d}t - \tilde{\nu}^I{}_t\,\mathrm{d}t)$, the state remains 0 and there is no inspection, in which case the agent gets the discounted payoff $(1 - r\,\mathrm{d}t)U_{t+\mathrm{d}t}^0$ from reporting state 0 (truthfully) at $t + \mathrm{d}t$. Substituting the approximation $\mathrm{d}U_t^0 := U_{t+\mathrm{d}t}^0 - U_t^0$ and ignoring higher-order terms, this necessary condition is equivalent to

$$\mathrm{d}U_t^0 \le rU_t^0\,\mathrm{d}t - \lambda\alpha\big(U_t^1 - U_t^0\big)\,\mathrm{d}t + c\,\mathrm{d}t + \mathrm{d}F_t + \tilde{\nu}_t^I\,\mathrm{d}t\big(B + U_t^0\big).$$

[17] This corresponds to the threat-keeping constraint in Fernandes and Phelan (2000).

This inequality is precisely condition (H) for the case $i = 1$ and $j = 0$ when $\mathrm{d}\nu_t^I = \tilde{\nu}_t^I \,\mathrm{d}t$. By ensuring that the transitional utility U_t^j decreases quickly enough, the agent is deterred from delaying the report of any transition. While the heuristic derivation above generates a necessary condition, the general result in Lemma 1 is also sufficient and covers the possibility of inspections with positive probability mass.

Third, the obedience constraint (O) ensures that $\eta_t = 1$ is a best response for the agent. The marginal cost of effort is c. The marginal benefit is $\lambda\alpha(U_t^1 - U_t^0)$, where λ is the arrival rate of a shock, α is the sensitivity of the realization to the agent's effort, and $U_t^1 - U_t^0$ is the utility gain from the high state.

Fourth and finally, the participation constraint (P) makes sure that the agent's payoff is not below $-B$ so he does not withdraw from the contract. Whereas the agent only incurs costs from effort and fines, only payoffs below 0 are feasible.

3.2 *Principal: Sequential rationality and predictability*

A principal-optimal equilibrium can be characterized by solving an auxiliary mechanism-design problem in which the principal minimizes her inspection costs over all strategies with nonrandom inspections subject to the incentive compatibility (IC) conditions in Lemma 1.

LEMMA 2. *Let $\{N_t^{I*}, F_t^*\}_{t\geq 0}$ be a solution to the auxiliary mechanism-design problem*

$$\min_{\{N_t^I, F_t\}_{t\geq 0}} \mathbb{E}^P\left[\int_0^\infty e^{-rt}\kappa \,\mathrm{d}N_t^I\right],$$

subject to the requirements that (i) the corresponding utility paths $\{U_t^1, U_t^0\}_{t\geq 0}$ satisfy incentive-compatibility conditions (Pk), (H), (O), and (P) in Lemma 1, and (ii) $\{N_t^I\}_t$ is predictable whenever the agent reports compliance. Then $\{N_t^{I*}, F_t^*\}_{t\geq 0}$ describes the principal's strategy on the path of a principal-optimal truthful and maximally compliant equilibrium.

The result requires that inspections are predictable only during compliance, which is sufficient here as it is never optimal to inspect when the agent admits noncompliance.

Lemma 2 is the result of two essential insights. First, the minimal costs for the principal in Lemma 2 cannot exceed the principal's optimal equilibrium costs. That is, in equilibrium, she cannot gain from any randomness in the timing of inspections. To prove this first insight, we exploit the idea that in any equilibrium in which an inspection arrives at random, the principal must be indifferent between all times in the support of the inspection-time distribution.[18] By replacing a random inspection with a deterministic inspection time in the support, any equilibrium strategy can be transformed into a predictable strategy that generates the same expected costs. We show that this can be achieved while preserving incentives for the agent.

[18]Given that the agent always exerts full effort, any inspection that occurs at the later of two times in the support must be followed by a continuation equilibrium with strictly higher expected costs compared to the continuation equilibrium after the earlier possible inspection time.

Second, the the minimal costs for the principal in Lemma 2 do not lie strictly below the principal's optimal equilibrium costs. That is, any nonrandom inspection strategy that solves the problem can be supported in a perfect Bayesian equilibrium. Predictability makes it easy to incentivize the principal because the agent immediately detects when an inspection does not take place as anticipated. In the equilibrium constructed formally below, the agent immediately stops exerting effort and exits if the principal deviates by not inspecting as expected. After detecting such a deviation, the agent infers that the principal has become nonvigilant. The agent would thus begin to shirk, and the principal would retaliate by setting large fines, which forces the agent to exit.

Note that the exact continuation play after a deviation by the principal is not essential. What is needed for our construction is that there is a continuation play that is sufficiently undesirable for the principal that it incentivizes her to inspect. The agent's exit thus represents, in reduced form, a possibly richer continuation play, which could involve periods of shirking by the agent, and intensified inspection regimes by the principal in an effort to reestablish a reputation for vigilance.

3.3 *Solving the auxiliary problem*

We now illustrate the solution of the principal's problem in Lemma 2. It is intuitive that inspections are unnecessary when the agent admits that $\theta_t = 0$. It follows from the obedience constraint (O) that the agent has no incentive to misreport noncompliance. We therefore focus on histories for which the state is in compliance and turn to noncompliance later.

We begin by characterizing the agent's promised and transitional utility during compliance in terms of a pair of coupled differential equations. The restriction to nonrandom policies implies $\mathrm{d}\nu_t^I = 0$ everywhere, except at isolated times at which an inspection occurs with probability 1. In this case, it is without loss for the principal to set $\mathrm{d}F_t = 0$ between inspections. This is because, in the absence of inspections, the fines cannot depend on the true state (conditional on the report). Furthermore, we verify in the proof that at the optimum, constraint (H) holds with equality between inspections. Thus, if we start at $t = 0$ with some initial payoff pair (U_0^0, U_0^1), the trajectories of the transitional payoff U_t^0 and the persistent payoff U_t^1 up until the next inspection are pinned down by constraints (Pk) and (H) in Lemma 1 (with $\mathrm{d}\nu_t^I = \mathrm{d}F_t = 0$ and (H) holding with equality). This pair of coupled first-order differential equations has the closed-form solution

$$U_t^0 = e^{rt}\big(U_0^0 - \alpha\big(e^{\lambda t} - 1\big)\big(U_0^1 - U_0^0\big)\big) + c\big(e^{rt} - 1\big)/r \tag{4}$$

$$U_t^1 = e^{rt}\big(U_0^1 + (1-\alpha)\big(e^{\lambda t} - 1\big)\big(U_0^1 - U_0^0\big)\big) + c\big(e^{rt} - 1\big)/r. \tag{5}$$

Combining these equations yields the identity $U_t^1 - U_t^0 = (U_0^1 - U_0^0)e^{(r+\lambda)t}$. Whenever there is no inspection, the difference $U_t^1 - U_t^0$ must increase to guarantee that the agent cannot profit from delaying the report of a transition to noncompliance. Thus, given an initial payoff pair with $U_0^1 - U_0^0 \geq c/(\lambda\alpha)$, the paths U_t^0 and U_t^1 satisfy constraints (Pk), (H), and (O) at all times $t \geq 0$. The remaining constraint is (P); specifically,

$U_t^0 \geq -B$ and $U_t^1 \leq 0$.[19] To make the promised utilities satisfy (P) at all times, the principal performs inspections that allow her to increase the transitional utility U_t^0 without violating the honesty constraint so as to push U_t^0 and U_t^1 back together.

We solve for the optimal strategy using a recursive approach due to Davis (1993), using the promised utilities (U_t^0, U_t^1) as state variables. This approach involves restricting the principal to perform a bounded number of inspections and then solving for the optimal strategy iteratively, letting the maximal number of inspections go to infinity.

Suppose the principal can perform only one inspection and consider the choice of initial values $(U_0^0, U_0^1) = (u^0, u^1)$ to ensure that $U_t^1 \leq 0$ and $U_t^0 \geq -B$ for as long as possible. First, what is the optimal u^0 for any given value of u^1? From (4) and (5) we see that, for all t, the value U_t^0 is increasing and U_t^1 is decreasing in u^0. Thus, to satisfy $U_t^1 \leq 0$ and $U_t^0 \geq -B$, it is optimal to choose u^0 as large as (O) permits, i.e., equal to $u^1 - c/(\lambda\alpha)$. This is intuitive: whereas the honesty constraint (H) requires U_t^0 to fall quickly enough, it is optimal to start off from the highest possible value. Second, with $u^0 = u^1 - c/(\lambda\alpha)$, what is the optimal level of u^1? Substituting for u^0 in (4) and (5), we see that, for all t, U_t^0 and U_t^1 are increasing in u^1. Thus, an increase in u^1 makes U_t^0 hit the lower bound $-B$ later while it makes U_t^1 hit the upper bound 0 earlier. Given that the objective is to satisfy both constraints for as long as possible, the optimal choice of u^1 when there is one inspection makes U_t^0 and U_t^1 hit their respective boundary at the same time. Figure 1 illustrates this. For any other choice of initial value u^1, the minimum of both hitting times would be lower.

Going to multiple inspections, note that hitting $U_t^1 = 0$ at any time implies that no further incentives can be provided and the agent would stop exerting effort.[20] The optimal initial utility u^1 with multiple inspections is lower so that at the first inspection, i.e., when U_t^0 reaches $-B$, the value U_t^1 is below 0 to incentivize the agent in the future and

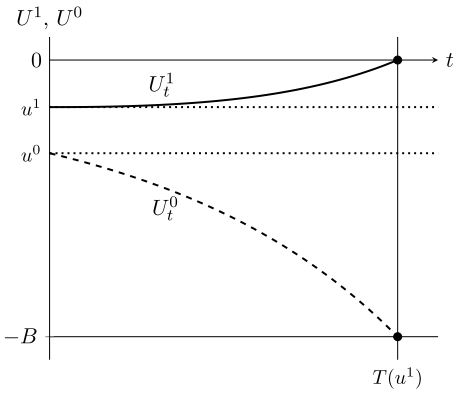

FIGURE 1. The evolution of promised utilities over time conditional on continued compliance with a single inspection. Persistent utility is shown as a solid line; transitional utility is dashed.

[19] By (O), these are two relevant inequalities implied by (P).

[20] Clearly, if the agent is expected to exert effort at all times, his expected utility cannot lie above $-c/r$. In fact it will lie strictly below this level as breaches of compliance, and thereby fines, cannot be fully ruled out even with full effort.

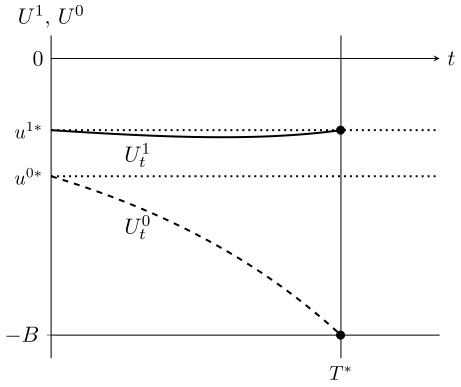

FIGURE 2. The evolution of promised utilities over time conditional on continued compliance with repeated inspections. Persistent utility is shown as a solid line; transitional utility is dashed.

allow time until the following inspection. When iterating over the number of inspections, let $u^1(n)$ denote the optimal initial value of the trajectory of U_t^1 when the maximal number is n. With each additional inspection, the optimal value $u^1(n)$ decreases. This implies that the time until the first inspection, T^n, decreases as U_t^0 reaches $-B$ earlier. As n grows large, $u^1(n)$ converges to a unique limit u^{1*} and T^n converges to a unique limit T^*, the length of the inspection cycle. In the optimal mechanism, the persistent utility U_t^1 is u-shaped and it returns to the initial value u^{1*} at each inspection (see Figure 2). The limit values u^{1*} and T^* are characterized by (4) and (5) with boundaries $(U_0^0, U_0^1) = (u^{1*} - c/(\lambda\alpha), u^{1*})$ and $(U_{T^*}^0, U_{T^*}^1) = (-B, u^{1*})$. This yields

$$T^* = \sup\{T > 0 \mid 0 = (B - c/r)\big(1 - e^{-rT}\big)\lambda\alpha - ce^{\lambda T}\big(e^{rT} - \alpha\big) + c(1-\alpha)\} \tag{6}$$

$$u^{1*} = -B + e^{(r+\lambda)T^*}\frac{c}{\lambda\alpha}, \quad \text{and} \quad u^{0*} = u^{1*} - \frac{c}{\lambda\alpha}. \tag{7}$$

So far, we abstracted from transitions to state $\theta_t = 0$. If such a breach of compliance occurs at time t, the persistent utility becomes U_t^0 in (4). The agent then pays a lump-sum fine $P(t) = u^{0*} - U_t^0$ to increase the persistent utility to u^{0*}. Using constant flow fines (and no inspections), the promised utilities are held constant at $U_t^0 = u^{0*}$ and $U_t^1 = u^{1*}$ as long as $\theta_t = 0$.[21] This way, upon another transition to compliance, the promised utilities are already at the optimal initial values.

4. EQUILIBRIUM

4.1 *The principal-optimal equilibrium*

We now characterize a principal-optimal equilibrium. The equilibrium we consider alternates between two phases: First, while the agent reports compliance, he pays no fine and is subject to periodic inspections with inspection cycle length T^*. Formally, let the

[21] Verify with equations (Pk) and (H) for the case $i = 0$ and $j = 1$, and $\mathrm{d}\nu_t^I = 0$.

clock

$$\tau_t \equiv t - \sup\{s \in [0,t) \mid \hat{\theta}_s = 0 \vee \mathrm{d}N_s^I = 1\}$$

count the time in compliance since the last transition or inspection. While in compliance ($\theta_t = 1$), the clock τ_t increases linearly with slope 1, and it drops to 0 during noncompliance ($\theta_t = 0$) or at each inspection ($\mathrm{d}N_t^I = 1$). At each time t with $\tau_t = T^*$, an inspection is performed. Second, while the agent reports noncompliance, he pays a lump-sum fine at the time of the transition and a constant flow fine at all times.

THEOREM 1. *There is a principal-optimal truthful and maximally compliant equilibrium with inspection cycle length T^* and initial expected payoff pair (u^{0*}, u^{1*}) such that on the equilibrium path, the following statements hold.*

- *Inspections are performed only during compliance ($\theta_t = 1$), with a periodic inspection whenever the clock τ_t reaches T^*.*
- *Fines are levied only during noncompliance ($\theta_t = 0$), with a constant flow fine $f^* = -ru^{0*}$ at all times t with $\theta_t = 0$, and a lump-sum transition fine $P(\tau_t) = u^{0*} - U^0_{\tau_t}$ at all times t with $\theta_{t-} = 1$ and $\theta_t = 0$, where U^0_t is given in (4) with initial value u^{0*}.*

Off the equilibrium path, the following statements hold.

- *If an inspection reveals noncompliance, i.e., that the agent misreported, then the agent pays the maximal fine, so that his continuation utility is $-B$.*
- *If $\tau_t = T^*$ but the principal fails to inspect, then the agent exits.*

Figure 3 illustrates the equilibrium for a sample path with initial state $\theta_0 = 1$. While in compliance, the agent pays no fines, and an inspection is performed at time t_1, where $\tau_{t_1} = T^*$. During compliance, the agent's persistent payoff evolves according to $U_t = U^1_{\tau_t}$, which is equal to u^{1*} initially and at the inspection time. At time t_2 in Figure 3, a breach of compliance occurs. In a first step, the agent's utility drops to the current level of the transitional utility $U^0_{\tau_{t_2}}$, the dashed blue line; at the same time, the agent pays the transition fine $P(\tau_{t_2})$, so that his continuation utility increases by that amount to u^{0*}. While in noncompliance, the agent pays a constant flow fine that keeps the continuation utility constant at u^{0*} until the transition back to compliance at t_3. At this transition, the persistent utility jumps up to u^{1*} and the evolution takes the same course as at $t = 0$ and at $t = t_1$.

Note that the persistent utility of the agent during compliance is u-shaped. This is the result of two opposing forces. On one hand, the agent faces a mounting threat from the increasing transition fine he must pay when becoming noncompliant. The rise in the transition fee is necessary to maintain truth-telling incentives. On the other hand, the likelihood of having to pay this fine falls as he approaches the next inspection. Early on in the inspection cycle, the first force is dominant, resulting in a decrease in persistent utility, while the latter force is dominant toward the end of the inspection cycle.

A crucial feature of this equilibrium is that inspections are predictable from the perspective of the agent. With nonrandom inspections, each inspection is a signal to the

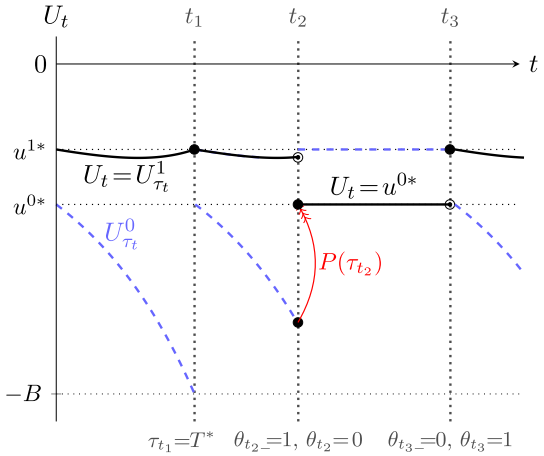

FIGURE 3. The evolution of an example path realization starting in the compliant state. Solid curves depict the agent's persistent payoff in the current state; dashed curves depict the transitional payoff, to which the agent's payoff jumps when the state changes.

agent of the principal's continued oversight. Demonstrated vigilance shapes the agent's perception that he will eventually be detected if he were to deviate. While random inspections may be supported in a relational contract, such arrangements require strong deterrents for the principal to ensure her adherence to the equilibrium strategy. This requirement ultimately renders randomization nonbeneficial for the principal (Lemma 2).

The equilibrium in Theorem 1 naturally features penalty reductions for early disclosures of noncompliance. This is consistent with voluntary disclosure schemes that are commonly used in practice. The U.S. environmental protection agency (EPA) employs a self-reporting program called "Incentives for Self-Policing" that requires firms voluntarily disclose any violations that are detected internally. Similar to the way the agent is incentivized in the above equilibrium, firms that disclose violations early are rewarded by a reduction in penalties and a suspension of inspections until compliance is restored. Theorem 1 provides insights into how enforcement agencies can benefit from offering regulated firms incentives for voluntary disclosure. Voluntary disclosure allows the principal to limit inspection to periods of compliance and, thus, lowers the overall inspection costs. The EPA points out that the advantage of these incentives lies in "making formal EPA investigations and enforcement actions unnecessary."[22] In the theoretical literature, the observation that voluntary disclosure reduces monitoring costs dates back to Kaplow and Shavell (1994), who introduced self-reporting into the enforcement model by Becker (1968). Without the agent's disclosure, the principal in our model would not be able to consistently avoid inspections during phases of noncompliance.[23]

[22] https://www.epa.gov/compliance/how-we-monitor-compliance.

[23] See Varas, Marinovic, and Skrzypacz (2020) for a model without reports. Our results confirm the conjecture in that paper that voluntary disclosure can avoid unnecessary inspections (see Varas, Marinovic, and Skrzypacz (2020, p. 2921)).

4.2 *Comparative statics*

How do variations in the parameters affect the length of inspection cycles and the inspection costs? As one would expect, if the penalty bound B increases or the effort cost c decreases, the agency problem becomes less severe: the inspection cycle T^* becomes larger and the expected costs decrease.[24] The effect of a change in the arrival rate of shocks λ is more intricate. An increase in λ decreases the state's persistence and has a non-monotone effect on the length of inspection cycles and the overall costs. The following result makes these statements precise. To ensure that the equilibrium in Theorem 1 always exists, we require that $\lambda > \underline{\lambda} \equiv cr/(Br\alpha - c) > 0$, fixing all other parameters.[25]

LEMMA 3. *As the arrival rate of shocks λ increases, the following statements hold.*

- *The inspection cycle length increases for low λ and decreases for high λ, with* $\lim_{\lambda \downarrow \underline{\lambda}} T^*(\lambda) = \lim_{\lambda \uparrow \infty} T^*(\lambda) = 0$.
- *The discounted inspection costs decrease for low λ and increase for high λ, with* $\lim_{\lambda \downarrow \underline{\lambda}} K_0^*(\lambda) = \lim_{\lambda \uparrow \infty} K_0^*(\lambda) = \infty$.

For the cycle length T^*, there are two opposing effects if λ increases. First, at any given instance, the state is more likely to change in response to current effort. The marginal benefit from effort is higher and it is easier to incentivize the agent, allowing for an increase in T^*. Second, the state becomes more fragile: the link between current effort and future compliance weakens. Delayed inspections have less incentive power, forcing the principal to shorten inspection cycles. Lemma 3 shows that the first effect dominates for low λ and the second effect dominates for high λ.

For any fixed T^*, the total inspection costs decrease in λ as any cycle of fixed length is more likely to be interrupted by a breach of compliance, so the inspection is less likely to be carried out. Thus, for low λ, this effect and the increase in T^* work in the same direction. Inspection costs decrease in λ. For high λ, the two effects work in opposite directions. Lemma 3 shows that the decrease in T^* is fast enough to outdo the second effect; the inspection costs increase in λ. Both inspection intensity and inspection costs grow arbitrarily large at both extremes.

As λ goes to infinity and state persistence vanishes, inspections must be immediate to deter deviations. This highlights a key disadvantage of nonrandom inspections and the absence of commitment. Intuitively, a shirking agent faces an "effective" discount rate of $r + \lambda$ when considering the impact of the next inspection. This is because the state today determines the state at the next inspection only with probability $e^{-\lambda T}$. To ensure inspection effectiveness, $T^*(\lambda)$ must approach zero fast enough so as to keep $\lim_{\lambda \to \infty} \lambda e^{-(r+\lambda)T^*(\lambda)}$ strictly positive. For the principle, in contrast, the effective discount rate is $r + \lambda(1 - \alpha)$, which is smaller than that for the agent. The limit of the

[24] A formal proof for the changes in B and c, as well as comparative statics with respect to α, can be found in a working paper version, which is available upon request.

[25] Observe that the lower bound on B required for feasibility of effort, $\bar{B} = \frac{c(r+\lambda)}{\lambda \alpha r}$, grows arbitrarily large as λ goes to 0.

principal's cost is proportional to $\lim_{\lambda\to\infty} \lambda e^{-(r+(1-\alpha)\lambda)T^*(\lambda)}$. It is then easy to see that this cost must be infinite for $\alpha < 1$ when $\lim_{\lambda\to\infty} \lambda e^{-(r+\lambda)T^*(\lambda)} > 0$.

The high compliance cost for large λ arises from the agent's opportunity to regain compliance with high probability unless the next inspection is imminent. Imminent inspections (T^* near 0) inflate costs. Random inspections may be valuable, allowing the principal to threaten immediate inspections without performing them constantly. We now confirm that random inspection schedules outperform predictable ones when feasible. With randomization, the principal strictly prefers higher arrival rates λ.

5. Commitment

Our results show that without commitment, the principal cannot benefit from randomization. In this section, we confirm that if the principal could commit to follow through with random inspection schedules, this would decrease inspection costs.

One way to enhance commitment to a profitable random procedure in arm's-length enforcement is to separate planning and execution of inspections, as seen in German banking supervision. The European Central Bank or the supervisory agency at the Finance Ministry (BaFin) schedules audits, while the German Bundesbank executes them (BaFin (2016)). The inspection cost is not incurred by the party making the inspection decision, eliminating the temptation to delay or skip inspections. This separation differs significantly from two seemingly similar alternatives: outsourcing all oversight or compensating the principal for inspection costs. Outsourcing only shifts the problem one layer further; compensation requires precise knowledge of the cost to avoid inefficient inspections.[26]

Alternatively, the lack of detectability, which hinders profitable randomization, can be overcome if the principal is responsible for overseeing a large pool of independent agents and there are public records. The principal can then regularly inspect a fixed proportion of agents and make the results publicly available to create a verifiable signal of continued vigilance. For example, the EPA's database "Enforcement and Compliance History Online" collects over 44,000 inspected facilities within the 12 months up to April 2021;[27] the Public Company Accounting Oversight Board (PCAOB) publicizes approximately 100–300 inspection reports per year.[28]

To confirm the benefit of randomization, consider the following mechanism, which is optimal in the class of stationary random mechanisms.[29]

- Inspections are performed only during compliance with constant Poisson arrival rate

$$m_R^* = r\frac{\bar{B}}{B-\bar{B}}.$$

[26]If the compensation falls short of the actual cost and effort required for an inspection, the incentive to skip it persists. If the compensation exceeds the cost, this creates an incentive to inspect inefficiently often.

[27]https://echo.epa.gov.

[28]https://pcaobus.org/oversight/inspections/firm-inspection-reports.

[29]For a proof of this claim, see Appendix C. Note that we do not claim that the mechanism presented here is the optimal commitment mechanism.

- Fines are levied only during noncompliance with a constant flow fine

$$f_R^* = r\bar{B}.$$

- If an inspection reveals noncompliance, then the agent pays the maximal fine.

Similar to the equilibrium with predictable inspections, there are two phases. Inspections but no fines during compliance, and fines but no inspections during noncompliance. The differences are that inspections arrive at random and there is no lump-sum fine at transitions to noncompliance.

Inserting into Lemma 1 the values $\mathrm{d}F_t = 0$ and $\mathrm{d}\nu_t^I = m_R^* \,\mathrm{d}t$ in case $i = 1$, and the values $\mathrm{d}F_t = f_R^* \,\mathrm{d}t$ and $\mathrm{d}\nu_t^I = 0$ in case $i = 0$, it is straightforward to verify that the payoffs of the agent are constant at

$$U_R^1 = -\frac{c}{r\alpha}, \qquad U_R^0 = -\frac{c}{r\alpha} - \frac{c}{\lambda\alpha}. \tag{8}$$

Here, U_R^1 is the persistent payoff and U_R^0 is the transitional payoff when the agent reports compliance, and vice versa when the agent reports noncompliance. It is straightforward that all constraints are satisfied at all times, with (H) binding in $i = 1$ and (O) binding in both states. The next result shows that the principal's inspection costs with predictable inspections are generally higher than with random inspections. In contrast to the predictable inspection schedule, a high arrival rate λ benefits the principal in this random mechanism.

THEOREM 2. *The inspection costs in the stationary random mechanism defined above are strictly lower than in the principal-optimal equilibrium in Theorem 1. Furthermore, the costs in this random mechanism are decreasing in λ for all λ, with*

$$\lim_{\lambda \downarrow \underline{\lambda}} K_R(\lambda) = \infty \quad \textit{and} \quad \lim_{\lambda \uparrow \infty} K_R(\lambda) = \frac{c\alpha}{Br\alpha - c}\kappa.$$

Random inspections dominate predictable inspection procedures for two reasons. First, by the argument at the end of Section 4.2, noise and delay make periodic inspections less effective. The threat of an imminent inspection at all times is more effective in our setting, even when holding the payoff impact of each inspection fixed.[30] That is, if the agent's initial utility is fixed at some level u, the costs from the random mechanism implementing this utility level are strictly below the costs in the predictable equilibrium implementing the same level. Second, in our setting with fines and self-reported compliance, random inspections allow for a greater payoff impact of an inspection on the deviating agent: with predictable inspections, self-reporting requires a transition fine whenever a breach of compliance occurs. The risk of the transition fine reduces the agent's overall equilibrium payoff. Since the lower bound on payoffs is fixed at B, this reduction decreases the maximum loss that the principal can impose when an inspection

[30] See also Varas, Marinovic, and Skrzypacz (2020), who show that a constant inspection rate provides incentives most effectively under commitment when the payoff consequence of each inspection is fixed.

reveals a misreport. Thus, predictable inspections have a smaller payoff impact, making them overall less powerful.

Our finding that random inspections provide incentives more effectively is consistent with Varas, Marinovic, and Skrzypacz (2020), who study a setting without voluntary disclosure. They show that (partially) predictable inspections can be optimal when the principal derives direct value from information, i.e., when her flow payoff is convex in the posterior belief. In our model, the principal induces honest self-disclosure. Along the equilibrium path, the principal always knows the true state. Therefore, introducing convexity in the principal's value as a function of her belief would not affect our results; her belief is always 0 or 1. Varas, Marinovic, and Skrzypacz (2020) identify a trade-off according to which incentive provision recommends randomization while information acquisition makes predictable inspections more profitable. Our analysis suggests that self-reporting can resolve this trade-off in favor of randomization when the current state is known to the agent and monetary incentives are feasible.

6. Conclusion

We study enforcement through inspections and fines. In relational enforcement, maximum compliance and truthful disclosure are attained through nonrandom inspections. A fully committed principal would benefit from random inspections.

The persistent effect of the agent's effort on compliance makes it possible to create incentives through isolated inspections. An intermediate level of persistence is optimal in the case of relational enforcement. If the principal can commit to random inspections, inspection costs are increasing in the level of persistence as compliance becomes less responsive to effort. This highlights the importance of persistence in relational enforcement.

Throughout the analysis, we assume that the principal does not benefit from the fines imposed on the agent. This assumption is innocuous. Given that the agent exerts full effort at all times, when his initial promised utility is u, the expected discounted sum of fines paid by the agent is equal to $-u - c/r$. If the principal were to benefit from fines at rate $\beta \in (0, 1]$, her objective would be to maximize $-K(u) + \beta(-u - c/r)$ instead of maximizing $-K(u)$ in the baseline model. Denoting the maximizer of $-K(u)$ by u^*, it is easy to see that the optimal equilibrium consists of an initial fine $B + u^*$ paid to the principal, followed by the equilibrium of Theorem 1.[31]

A possible variation to our model is to allow the principal to pay subsidies to the agent when successfully passing inspections. If the principal could reward the agent for passed inspections, the upper bound on the agent's continuation utility would increase. The principal could then decrease the inspection frequency as the maximal punishment increases. With commitment to random inspections, the principal could essentially avoid all inspection costs if rewards were unbounded. She could offer an arbitrarily large reward after inspecting with vanishing probability.

[31]The agent's initial utility (before paying the fine) is at his outside option $-B$ and then jumps to u^*. The principal's payoff $-K(u^*) + \beta(B - c/r)$ is clearly an upper bound for $-K(u) + \beta(-u - c/r)$ among $u \geq -B$.

The assumptions that the principal implements full effort is natural in many situations, for example, when the principal, tasked with monitoring compliance, is not the same institution as the one designing the regulation. The assumption is also important for tractability. To let the principal choose effort, the model would need to account explicitly for the principal's benefit from compliance.[32] More importantly, the optimization problem would become substantially more complex. Maximizing over the effort level would add a continual control at all times.[33]

Similarly, we focus on equilibria with truthful self-reports. This focus is natural in many applications in which it is essential for regulators to accurately identify compliance violations. In the auxiliary mechanism design problem with principal commitment, having only one agent ensures that the revelation principle applies. With commitment, the principal can replicate the outcome of any combination of mechanism and reporting strategy with the corresponding direct mechanism and a truthful reporting strategy. However, in the equilibrium problem without commitment, we do not rule out potential benefits from nontruthful behavior. Since we find the optimal predictable equilibrium via the auxiliary mechanism design problem, the only remaining concern is whether the principal could exploit nontruthful reporting to benefit from random inspections in equilibrium. We suspect this is not the case, but verifying the conjecture is beyond this article's scope.

Appendix A: Proofs for Section 3

Proof of Lemma 1. The proof of Lemma 1 consists of two intermediate results. Lemma A provides a martingale representation for the agent's lifetime expected utility, and Lemma B provides necessary and sufficient conditions for the path of expected payoffs such that full effort and truthful reporting are a best response for the agent.

Define W_t to be the agent's lifetime expected utility, with expectations taken with respect to the information that is available at time t:

$$W_t = \int_0^t e^{-rs}(-\mathrm{d}F_s - c\eta_s\,\mathrm{d}s) + e^{-rt}U_t.$$

By construction, the process $\{W_t\}_{t\geq 0}$ is a martingale (Davis (1993, p. 20)). There are three types of events: changes in the state, changes in reports, and inspections. Inspections are governed by the process N^I given by the principal's strategy. For consistency, we introduce the counting processes $N^\theta = \{N_t^\theta\}_{t\geq 0}$ and $N^{\hat\theta} = \{N_t^{\hat\theta}\}_{t\geq 0}$ that count the number of changes in the state of compliance and in the reports, respectively. For each process

[32] In some cases, when the principal's benefit is large enough, implementing full effort is optimal and the analysis would be unaffected.

[33] Indeed, for predictable inspections, we sidestep the problem of continual controls by showing that it is without loss to levy no fines between inspections. This difficulty is also the reason why we do not claim that the random mechanism in Section 5 is optimal among all inspection mechanisms. While it is optimal among any Markovian procedure, confirming that it is optimal generally would require a verification argument that deals with continual controls that can change both continuously or impulsively. We are not aware of existing dynamic programming results to verify that the recursive approach we employ remains valid in this class of optimization problems.

N^a with $a \in \{\theta, \hat{\theta}, I\}$, define the *compensator* to be a predictable process $\nu^a = \{\nu_t^a\}_{t\geq 0}$ such that the compensated process $N_t^a - \nu_t^a$ is a martingale. The compensator exists under very general conditions and can be interpreted as the predictable drift of the underlying (nonpredictable) stochastic process. For the hazard rate of transitions in compliance, we shall write $q_t(\eta_t) := \mathrm{d}\nu_t^\theta/\mathrm{d}t$ or, more explicitly,

$$q_t(\eta_t) = \theta_{t-}\lambda(1-\alpha\eta_t) + (1-\theta_{t-})\lambda\alpha\eta_t. \tag{9}$$

The martingale representation theorem for marked point processes (Last and Brandt (1995)) implies the following result.[34]

Lemma A. *There exist predictable processes Δ^θ, $\Delta^{\hat{\theta}}$, and Δ^I such that the evolution of the agent's expected utility is given by*

$$\mathrm{d}U_t = rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\eta_t\,\mathrm{d}t + \sum_{a\in\{\theta,\hat{\theta},I\}} \Delta_t^a\big(\mathrm{d}N_t^a - \mathrm{d}\nu_t^a\big). \tag{10}$$

The processes Δ^θ, $\Delta^{\hat{\theta}}$, and Δ^I have an intuitive interpretation: They represent the jump in utility at time t that results from a change in compliance, a change in reported compliance, or an inspection.

The following lemma will complete the proof of Lemma 1.

Lemma B. *A mechanism that induces the payoffs $\{U_t\}_{t\geq 0}$ is incentive compatible with full effort and truthful reporting if and only if for all $t \geq 0$,*

(i) $(r + q_t(1))\Delta_t^{\hat{\theta}} - \mathrm{d}\nu_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}}) \geq \mathrm{d}\Delta_t^{\hat{\theta}}$ *when* $\theta_t \neq \hat{\theta}_t$

(ii) $(1-2\theta_{t-})\lambda\alpha(\Delta_t^\theta + \Delta_t^{\hat{\theta}}) \geq c$ *when* $\theta_t = \hat{\theta}_t$

(iii) $U_t \in [-B, 0]$.

Proof. Define

$$W_t = \int_0^t e^{-rs}(-\mathrm{d}F_s - c\eta_s\,\mathrm{d}s) + e^{-rt}\tilde{U}_t$$

to be the agent's expected payoff from choosing effort $\{\tilde{\eta}_s\}$ and report $\{\hat{\theta}_s\}$ up to time t with maximum effort and truthful reporting thereafter. Here $\tilde{U}_t$ is the expected continuation payoff. We may have $\tilde{U}_t \neq U_t$ if the agent has reported nontruthfully, i.e., $\hat{\theta}_{t-} \neq \theta_{t-}$. Consider first the case in which the agent's report regarding his type at time t is truthful, so that $\tilde{U}_t = U_t$. Differentiating with respect to t yields

$$\mathrm{d}W_t = e^{-rt}(-\mathrm{d}F_t - c\eta_t\,\mathrm{d}t) - re^{-rt}U_t\,\mathrm{d}t + e^{-rt}\,\mathrm{d}U_t.$$

[34]A formal proof for our setting, which is a straightforward adaptation of the proof of Theorem 1.13.15 in Last and Brandt (1995) p. 25, is provided in Supplemental Appendix B.

Using Lemma A to replace $\mathrm{d}U_t$ yields

$$\mathrm{d}W_t = e^{-rt}\Bigg((1-\eta_t)c\,\mathrm{d}t + \sum_{a\in\{\theta,\hat{\theta}\}} \Delta_t^a\big(\mathrm{d}N_t^a - q_t(1)\,\mathrm{d}t\big) + \Delta_t^I\big(\mathrm{d}N_t^I - \mathrm{d}\nu_t^I\big)\Bigg).$$

If the agent deviates for an additional instant (but still reports truthfully), then

$$\mathrm{d}N_t^\theta = \mathrm{d}N_t^{\hat{\theta}} = \begin{cases} 1 & \text{with probability } q_t(\tilde{\eta}_t)\,\mathrm{d}t \\ 0 & \text{with probability } 1 - q_t(\tilde{\eta}_t)\,\mathrm{d}t. \end{cases}$$

Taking expectations therefore yields

$$\mathbb{E}_t^A[\mathrm{d}W_t] = e^{-rt}\mathbb{E}^A\big[(1-\eta_t)c\,\mathrm{d}t + \big(\Delta_t^\theta + \Delta_t^{\hat{\theta}}\big)\big(q_t(\tilde{\eta}_t) - q_t(1)\big)\,\mathrm{d}t\big].$$

It follows from condition (ii) that

$$\big(\Delta_t^\theta + \Delta_t^{\hat{\theta}}\big)q(\tilde{\eta}_t) - c\eta_t \le \big(\Delta_t^\theta + \Delta_t^{\hat{\theta}}\big)q_t(1) - c.$$

Thus $\mathbb{E}_t^A[\mathrm{d}W_t] \le 0$. We thus obtain the chain of inequalities

$$\begin{aligned}\mathbb{E}_0^A[W_t] &= \mathbb{E}_0^A\bigg[\int_0^t \mathrm{d}W_s + W_0\bigg] = \int_0^t \mathbb{E}_0^A[\mathrm{d}W_s] + \mathbb{E}_0^A[W_0] \\ &= \int_0^t \mathbb{E}_0^A\big[\mathbb{E}_s^A[\mathrm{d}W_s]\big] + W_0 \le W_0. \end{aligned} \tag{11}$$

Now consider the case in which the agent's most recent report at time t is false, that is, $\theta_{t-} \neq \hat{\theta}_{t-}$, and he continues the nontruthful strategy for an additional moment at time t. If no change in the state occurs at the additional moment, then the agent must correct his report immediately thereafter. If a change occurs, then the previously false statement becomes truthful, and thus his report does not change. Therefore, we have

$$\begin{aligned}\mathrm{d}\tilde{U}_t &= \tilde{U}_t - \tilde{U}_{t-\mathrm{d}t} \\ &= \mathrm{d}N_t^\theta\big(U_t - U_{t-\mathrm{d}t} - \Delta_{t-\mathrm{d}t}^{\hat{\theta}}\big) + \mathrm{d}N_t^I\big(U_t + \Delta_t^I - U_{t-\mathrm{d}t} - \Delta_{t-\mathrm{d}t}^{\hat{\theta}}\big) \\ &\quad + \big(1 - \mathrm{d}N_t^\theta - \mathrm{d}N_t^I\big)\big(U_t + \Delta_t^{\hat{\theta}} - U_{t-\mathrm{d}t} - \Delta_{t-\mathrm{d}t}^{\hat{\theta}}\big) \\ &= \mathrm{d}N_t^\theta\big(\mathrm{d}U_t + \mathrm{d}\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}}\big) + \mathrm{d}N_t^I\big(\mathrm{d}U_t + \mathrm{d}\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}} + \Delta_t^I\big) \\ &\quad + \big(1 - \mathrm{d}N_t^\theta - \mathrm{d}N_t^I\big)\big(\mathrm{d}U_t + \mathrm{d}\Delta_t^{\hat{\theta}}\big) \\ &= \mathrm{d}U_t + \mathrm{d}\Delta_t^{\hat{\theta}} - \mathrm{d}N_t^\theta\Delta_t^{\hat{\theta}} + \mathrm{d}N_t^I\big(\Delta_t^I - \Delta_t^{\hat{\theta}}\big). \end{aligned}$$

Again using Lemma A to replace $\mathrm{d}U_t$, we obtain

$$\begin{aligned}\mathrm{d}W_t &= e^{-rt}(-\mathrm{d}F_t - c\eta_t\,\mathrm{d}t) - re^{-rt}\big(U_t + \Delta_t^{\hat{\theta}}\big) \\ &\quad + e^{-rt}\big(rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t + \Delta_t^\theta\big(\mathrm{d}N_t^\theta - q_t^*\big) + \mathrm{d}\Delta_t^{\hat{\theta}} - \mathrm{d}N_t^\theta\Delta_t^{\hat{\theta}} + \mathrm{d}N_t^I\big(\Delta_t^I - \Delta_t^{\hat{\theta}}\big)\big). \end{aligned}$$

It follows from the honesty constraint (i) that, in expectation, $\mathrm{d}\Delta_t^{\hat\theta} \le (r + q_t(1))\Delta_t^{\hat\theta} - \mathrm{d}\nu_t(\Delta_t^I - \Delta_t^{\hat\theta})$. Substituting it into $\mathrm{d}W_t$ and simplifying, again using $\tilde U_t = U_t + \Delta_t^{\hat\theta}$, gives

$$\mathbb{E}_t^A[\mathrm{d}W_t] = e^{-rt}\big((1-\eta_t)c\,\mathrm{d}t + \big(\Delta_t^\theta - \Delta_t^{\hat\theta}\big)q(\tilde\eta_t) - q_t(1)\big(\Delta_t^\theta - \Delta_t^{\hat\theta}\big)\big).$$

Now $\Delta_t^\theta - \Delta_t^{\hat\theta} = (\Delta_t^\theta + U_t) - (\Delta_t^{\hat\theta} + U_t)$ is the payoff difference from a change in the state without a change in report and a change in report without a change in the state. Since $\theta_{t-} \neq \hat\theta_{t-}$ by hypothesis, this is identical to $\tilde\Delta_t^\theta + \tilde\Delta_t^{\hat\theta}$ after the history in which the true state was identical to his report. Thus (ii) implies that $\eta_t = 1$ maximizes the right-hand side, so that $\mathbb{E}_t^A[\mathrm{d}W_t] \le 0$. By the same argument as in (11), we have $\mathbb{E}_0^A[W_t] \le W_0 = U_0$, so that the agent cannot profit from deviating. Taking the limit, we find that

$$\lim_{t\to\infty} \mathbb{E}_0^A[W_t] \le U_0,$$

which implies that the agent cannot gain from deviating from maximum effort and truthful reporting. Conversely, if the incentive constraint (i) is violated, then the above inequalities are inverted, so that the agent has a strict incentive to be dishonest. Likewise, if (ii) is violated, the agent has a strict incentive to exert no effort, and a violation of (iii) leads to exit by the agent. □

To complete the proof of Lemma 1, we show that condition (Pk) follows from Lemma A, and (H), (O), and (P) are equivalent to conditions (i), (ii), and (iii) in Lemma B.

Consider a mechanism and a strategy for the agent that jointly generate the payoff process $\{U_t\}_{t\ge0}$ for the agent, and denote by $\{U_t^1, U_t^0\}_{t\ge0}$ the associated pair of promised utilities defined in (3).

Step 1. By the definition of U_t^0, U_t^1, we have

$$\begin{aligned}\Delta_t^\theta + \Delta_t^{\hat\theta} &= \begin{cases} U_t^1 - U_t^0 & \text{if } \theta_{t-} = \hat\theta_{t-} = 0 \\ U_t^0 - U_t^1 & \text{if } \theta_{t-} = \hat\theta_{t-} = 1, \end{cases} \\ q_t(1) = q_t(1) &= \begin{cases} \lambda\alpha & \text{if } \theta_{t-} = 0 \\ (1-\alpha)\lambda & \text{if } \theta_{t-} = 1. \end{cases}\end{aligned} \tag{12}$$

Combining these two expressions, we can write more succinctly

$$q_t(1)\big(\Delta_t^\theta + \Delta_t^{\hat\theta}\big) = \lambda(\theta_{t-} - \alpha)\big(U_t^1 - U_t^0\big).$$

Lemma A then implies that, conditioning on the event that $\mathrm{d}N_t^\theta = \mathrm{d}N_t^{\hat\theta} = \mathrm{d}N_t^I = 0$, we get exactly condition (Pk) in Lemma 1.

Step 2. Next, suppose that the agent is not truthful after some history at time t. Let $i = \theta_t$ be the true state and suppose the agent reports $j = 1 - \theta_t$. Then $U_t^i = U_t + \Delta_t^{\hat\theta}$ and

$$\begin{aligned}\mathrm{d}U_t^i &= \big(U_{t+\mathrm{d}t} + \Delta_{t+\mathrm{d}t}^{\hat\theta}\big) - \big(U_t + \Delta_t^{\hat\theta}\big) = rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t - q_t(1)\Delta_t^\theta + \mathrm{d}\Delta_t^{\hat\theta} \\ &\le rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t - q_t(1)\Delta_t^\theta + \big(r + q_t(1)\big)\Delta_t^{\hat\theta} - \mathrm{d}\nu_t\big(\Delta_t^I - \Delta_t^{\hat\theta}\big) \\ &= rU_t^i + \lambda(i-\alpha)\big(U_t^1 - U_t^0\big) - \mathrm{d}\nu_t\big(\Delta_t^I - \Delta_t^{\hat\theta}\big) + \mathrm{d}F_t + c\,\mathrm{d}t.\end{aligned} \tag{13}$$

The second line follows from Lemma A, and the inequality in the third line follows from condition (i) in Lemma B, where we take expectations conditional on the event that $\mathrm{d}N_t^\theta = \mathrm{d}N_t^{\hat\theta} = 0$. The last equality in (13) holds since

$$q_t(1)\big(\Delta_t^\theta - \Delta_t^{\hat\theta}\big) = q_t(1)\big(U_t + \Delta_t^\theta - \big(U_t + \Delta_t^{\hat\theta}\big)\big) = q_t(1)\big(U_t^j - U_t^i\big) = \lambda(i-\alpha)\big(U_t^1 - U_t^0\big).$$

Punishment is without cost for the principal and, therefore, it is optimal to impose the most severe punishment after an inspection reveals a dishonest report. The severity of punishments is restricted by the limits of enforcement that require the agent's continuation value not to fall below the lower bound $-B < 0$. Therefore, we have

$$\Delta_t^I - \Delta_t^{\hat\theta} = \underbrace{U_t + \Delta_t^I}_{=-B} - \big(\underbrace{U_t + \Delta_t^{\hat\theta}}_{=U_t^i}\big) = -\big(B + U_t^i\big).$$

Substituting this last equation into (13) yields

$$\mathrm{d}U_t^i = rU_t^i + \lambda(i-\alpha)\big(U_t^1 - U_t^0\big)\,\mathrm{d}t + \mathrm{d}\nu_t\big(B + U_t^i\big) + \mathrm{d}F_t + c\,\mathrm{d}t,$$

which is equal to condition (H) in Lemma 1. Conversely, if (i) does not hold at some t, then using the same steps as above, the inequality is reversed, so that (H) is violated.

Step 3. Substituting (12) into the obedience constraint (ii), we obtain, for each θ_{t-},

$$\big(\Delta_t^\theta + \Delta_t^{\hat\theta}\big)(1 - 2\theta_{t-})\lambda\alpha = \lambda\alpha\big(U_t^1 - U_t^0\big) \geq c.$$

The last inequality is identical to (O) in Lemma 1. Conversely, if (ii) is violated at some t, then the inequality is reversed, so that (O) is violated. □

Proof of Lemma 2. The proof of Lemma 2 consists of two results, stated and proven formally below.

Lemma C. *For any truthful and maximally compliant equilibrium, there exists a principal strategy such that truthful reporting and maximal compliance are a best response for the agent and*

(i) inspections are predictable for the agent whenever he reports compliance

(ii) it generates weakly lower inspection costs for the principal.

Proof. Take any truthful maximally compliant equilibrium. Let U_t^0 and U_t^1 be the continuation payoffs of the agent in this equilibrium. The following steps present a modified inspection schedule satisfying the properties stated in Lemma C. As the original equilibrium is truthful and maximally compliant, U_t^0 and U_t^1 satisfy the constraints from Lemma 1. First, we argue that for any inspection following a high report, it is without loss to assume that truthful reports are never punished more than misreports to the low state at the time of an inspection. That is, if the persistent payoff U_t^1 jumps downward on the path after an inspection, it will do so by less than the distance from the transitional utility to the lower bound $-B$. Formally, let $\bar U_t^1$ be the agent's persistent payoff right after

an inspection performed at time t and let U^1_{t-} be the payoff just prior to time t. Recall that by definition $\Delta^I_t = \bar{U}^1_t - U^1_{t-}$. We show that, without loss, $\Delta^I_t > -B - U^0_t$.

Suppose, to the contrary, that $\Delta^I_t \leq -B - U^0_t \leq 0$. Then we can construct another truthful maximally compliant equilibrium in which the principal's inspection costs are weakly lower by removing instant t from the support of the inspection distribution. To satisfy the agent's incentives for truth-telling and compliance, we compensate for the change in utility resulting from eliminating the inspection. To this end, introduce an additional fine at t, such that the new fine is $\mathrm{d}\hat{F}_t = \mathrm{d}F_t - \mathrm{d}\nu^I_t \Delta^I_t$, where $\mathrm{d}F_t$ denotes the fine specified in the original equilibrium. Hence, the agent's expected loss from the inspection caused by $\Delta^I_t < 0$ is paid as a fine at time t. This way, the path of persistent payoff U^1_t remains unchanged for all $s \leq t$. Similarly, the path of transitional utility, U^0_s, remains unchanged as the continuation equilibria after a transition remain the same. Whereas both paths U^1_s and U^0_s are as before, the obedience constraint remains satisfied.

To see that the honesty constraint is not violated by this change, consider the constraint (H) in case $j = 0$:

$$\mathrm{d}U^0_t \leq r\big(U^0_t\big)\,\mathrm{d}t - \lambda\alpha\big(U^1_t - U^0_t\big)\,\mathrm{d}t + \mathrm{d}\nu^I_t\big(B + U^0_t\big) + \mathrm{d}F_t + c\,\mathrm{d}t.$$

The effect of the proposed change on the right-hand side of this constraint is $-\mathrm{d}\nu^I_t(B + U^0_t) - \mathrm{d}\nu^I_t\Delta^I_t$. Whereas $\Delta^I_t \leq -B - U^0_t$, this effect is positive and the path of U^0_s still satisfies the honesty constraint. To randomize at time t in the original equilibrium, the principal must have been indifferent between inspecting and continuing without inspection, so that removing instant t from the support weakly lowers inspection costs. Now, with $\Delta^I_t > -B - U^0_t$, we prove Lemma C. Suppose, toward a contradiction, that the statement in the result is false. Then there must be some time t and history h_t with $\hat{\theta}_t = 1$ such that any inspection schedule with the first inspection after t being predictable for the agent must create higher inspection costs for the principal. We show that this cannot be the case by replacing the random inspection with a nonrandom inspection at the earliest realization of the random inspection schedule.

Without loss, take the time t above to be $t = 0$ and $\hat{\theta}_0 = 1$. Let $\mathcal{T}$ be the support of the first inspection time for this history and denote its infimum by $t^0 = \inf \mathcal{T}$. If $\mathcal{T} = \{t^0\}$, the inspection strategy for this history is already predictable, and we continue with the next instance, interpreting 0 as the last time of inspection after the high report or the time of transition to the high report.

When the support is not a singleton, consider first the case in which $t^0 \in \mathcal{T}$, i.e., the infimum is contained in the support. In Supplemental Appendix C we extend the argument to the case $t^0 \notin \mathcal{T}$, i.e., when t^0 is an accumulation point. Let $t^0 \in \mathcal{T}$ and consider the inspection schedule with a certain inspection at t^0 in case time t^0 is reached without prior transition. If $\Delta^I_{t^0} \geq 0$, introduce an additional fine at t^0 so that the new fine is given by $\mathrm{d}\hat{F}_{t^0} = \mathrm{d}F_{t^0} + (1 - \mathrm{d}\nu^I_{t^0})\Delta^I_{t^0}$, where $\mathrm{d}F_{t^0}$ denotes the fine in the original equilibrium. The payoff paths U^0_s and U^1_s remain unchanged for $s \leq t^0$ and, thus, the obedience constraint is unaffected. The honesty constraint at t^0 is relaxed since both the increase in inspection probability and the additional fine increase the right-hand side of (H). If $\Delta^I_{t^0} < 0$, increasing the inspection probability from $\mathrm{d}\nu^I_{t^0}$ to 1 decreases the

persistent payoff path U^1_s for all $s \leq t^0$ by $|\Delta^I_{t^0}|(1 - \mathrm{d}\nu^I_{t^0})e^{-(r+(1-\alpha)\lambda)(t^0-s)}$. This change in persistent payoff cannot be compensated by an additional fine at the high report as it would reduce the expected persistent payoffs further. Instead, we ensure obedience and truth-telling by lowering the transitional payoff by the necessary amount. To this end, introduce an additional transition fine of $|\Delta^I_{t^0}|(1 - \mathrm{d}\nu^I_{t^0})e^{-(r+(1-\alpha)\lambda)(t^0-s)}$ to be paid at time $s \leq t^0$ if a transition to the bad state occurs. This additional fine ensures that the difference $U^1_s - U^0_s$ is as in the original equilibrium, so the obedience and honesty constraints will still be satisfied. To ensure that this additional transition fine is feasible, we need to verify for all $s \leq t^0$, that $U^0_s - |\Delta^I_{t^0}|(1 - \mathrm{d}\nu^I_{t^0})e^{-(r+(1-\alpha)\lambda)(t^0-s)} \geq -B$. This term is decreasing in s, so it is sufficient to verify that $U^0_{t^0} + \Delta^I_{t^0}(1 - \mathrm{d}\nu^I_{t^0}) \geq -B$. Recall that we have shown that for any inspection time, $\Delta^I_{t^0} > -B - U^0_{t^0}$. Feasibility follows since $\mathrm{d}\nu^I_{t^0} < 1$.

This concludes the proof of the result by constructing an inspection schedule in which the next inspection following a good report is predictable, the agent's incentive constraints are satisfied, and the principal's inspections costs have not increased. □

The next result implies that predictability of inspections is the only restriction implied by the principal's sequential rationality.

LEMMA D. *Take a strategy profile such that the following statements hold.*

(i) The inspection schedule is predictable for the agent.

(ii) The agent's strategy is truthful, maximally compliant, and a best response to the principal's strategy.

(iii) The expected cost to the principal along any history is below $\bar{K}$.

(iv) Every action path generated by the strategy profile is measurable.

Then there exists a perfect Bayesian equilibrium that generates the same distribution over action paths.

PROOF. We show that any predictable principal strategy that generates costs $K_t \leq \bar{K}$ for the principal for all t can be implemented in equilibrium. First, note that after any history, we can construct a continuation equilibrium in which the agent chooses to exit the relationship with probability 1. To support exit by the agent as a best response, the principal's strategy is such that whenever the agent fails to exit although he was supposed to do so, the principal implements the harshest possible fine, leading to a continuation payoff of $-B$ for the agent. This bad continuation equilibrium can be leveraged to support any principal strategy as an equilibrium given that its inspections are predictable for the agent.

Let $\{N_t, F_t\}$ be the paths induced by the strategy profile in the result. By hypothesis (ii), compliance is incentive compatible for the agent. Let $(\tilde{n}, \tilde{f})$ be an alternative strategy for the principal (with possibly random inspection) and denote by $\tilde{N}^I$ the resulting inspection path if the agent follows the compliant strategy. Adapt the agent's strategy

such that he exits after any history h_t with $\mathrm{d}N_t^I \neq \mathrm{d}\tilde{N}_t^I$, that is, whenever the agent observes that the principal deviated from the original inspection strategy. Define the set $D = \{t|\mathrm{d}N_t^I \neq \mathrm{d}\tilde{N}_t^I\}$ containing the dates at which the agent observes that the principal deviates from her original inspection strategy. Since the cost to the principal from the original strategy is below $\bar{K}$ at each t and the payoff from any deviating strategy is equal for all $t < \inf D$, her deviation cannot be profitable as it results in a cost $\bar{K}$ from $\inf D$ onward. Finally, adapt the principal's strategy from the result such that he fines the agent as harshly as possible whenever the agent was expected to exit but failed to do so. This way, for the agent the strategy that leads to exit at $t = \inf D$ is incentive compatible, and the constructed equilibrium differs from the initial strategy profile in Lemma D at most off the equilibrium path. □

Lemmas C and D in combination imply Lemma 2: to characterize principal-optimal equilibria, it is sufficient to find a strategy for the principal with nonrandom inspections that induces truthfulness and maximum effort and minimizes the principal's monitoring costs. □

Appendix B: Proofs for Section 4

Proof of Theorem 1

We first solve the auxiliary problem with additional constraints below and then verify that the solution indeed constitutes an optimal mechanism.

B.1 *Auxiliary control problem*

Consider the auxiliary control problem

$$\min_{\{N_t^I, F_t\}_{t\geq 0}} \mathbb{E}^P\left[\int_0^\infty e^{-rt}\kappa\,\mathrm{d}N_t^I\right] \tag{14}$$

subject to the incentive-compatibility conditions (H), (O), and (P), and the following additional conditions:

(A) When $\hat{\theta}_{t-} = 1$, there are no fines between inspections, that is, $\mathrm{d}N_t^I = 0$ implies $\mathrm{d}F_t = 0$ and the honesty constraint (H) binds.

(B) When $\hat{\theta}_{t-} = 0$, the evolution of U_t^1 is not limited by the honesty constraint (H).

Condition (A) is a restriction on the set of strategies for the principal so that she levies no fines when the agent reports compliance unless an inspection is performed. Condition (B) relaxes the incentive-compatibility restrictions, saying that the honesty constraint is imposed only while the agent reports compliance. It is intuitive that a principal-optimal relationship satisfies these properties as the agent must be incentivized to exert effort and truthfully report states of compliance. We now solve for the optimal mechanism. We first derive the optimal Markovian mechanism in the auxiliary

problem using an iteration argument. We then confirm that (i) fines between inspections cannot decrease the principal's costs, (ii) there is no mechanism in non-Markovian strategies that performs better in the relaxed problem than the optimal Markovian mechanism, and (iii) that the solution to the relaxed problem is achievable in the original problem.

Under condition (A), the honesty constraint (H) holds with equality during compliance, so that when $\hat{\theta}_t = 1$, conditions (Pk) and (H) yield a pair of simple first-order differential equations that can be solved in closed form. The inspection problem thus becomes a standard deterministic impulse-control problem with state constraints. We solve this by first deriving the optimal mechanism when the principal can inspect at most n times and continue iteratively to consider the limit as the total number of available inspections n goes to infinity. More specifically, for any integer $n \geq 0$, consider the problem of maximizing the objective in (14) subject to $\lim_{t\to\infty} N_t^I \leq n$ pathwise, and to the incentive-compatibility conditions (H), (O), (P), (A), and (B) at all $t \geq 0$ at which $N_t^I < n$. Denote by K_n the solution to the problem with n available inspections. It then follows from Proposition 54.18 in Davis (1993) that the value function for the auxiliary problem K is the limit of K_n, i.e., $K = \lim_{n\to\infty} K_n$.

Evolution of promised utilities during compliance We begin by establishing an upper bound for the promised utility for the agent.

Claim 1. *Along the path of any maximally compliant mechanism, we have* $U_t^1 \leq -c/(r\alpha)$.

Proof. Let $\bar{U}^1$ be the supremum of U_t^1 that exists by (P). By obedience (O), we have that $\bar{U}^1 - c/(\lambda\alpha)$ is an upper bound for U_t^0. Therefore, in a maximally compliant equilibrium, we must have

$$\bar{U}^1 \leq \int_0^\infty e^{-(r+\lambda(1-\alpha))s}\left[-c + \lambda(1-\alpha)\left(\bar{U}^1 - \frac{c}{\lambda\alpha}\right)\right] \mathrm{d}s.$$

Solving the integral yields

$$\bar{U}^1 \leq \frac{-c + \bar{U}^1\lambda\alpha(1-\alpha)}{r\alpha + \lambda\alpha(1-\alpha)} \quad \Rightarrow \quad \bar{U}^1 \leq -\frac{c}{r\alpha}.$$

Since $\bar{U}^1$ is the supremum for U_t^1, we have $U_t^1 \leq -\frac{c}{r\alpha}$ as required. □

By assumption (A), the promise-keeping and truth-telling constraints in state $\hat{\theta}_t = 1$ yield a system of coupled first-order differential equations with the solution given in (4) and (5). For given initial values $U_0^1 = u^1$ and $U_0^0 = u^0$, (4) and (5) reveal immediately that for $u^1 < -\frac{c}{r\alpha}$, U_t^1 is u-shaped in t and strictly decreasing in u^0, whereas U_t^0 is strictly decreasing in t and strictly increasing in u^0. We will show below that it is optimal to set $u^0 = u^1 - \frac{c}{\lambda\alpha}$ and, therefore, it is sufficient to specify the promised utility $u = u^1$. We define

$$\phi_1(t,u) := e^{rt}\left(u + \left(\frac{1-\alpha}{\alpha}\right)\left(e^{\lambda t} - 1\right)\frac{c}{\lambda}\right) - c\left(1 - e^{rt}\right)/r \tag{15}$$

$$\phi_0(t,u) := e^{rt}\left(u - \left(\frac{1-\alpha}{\alpha}\right)\frac{c}{\lambda} - e^{\lambda t}\frac{c}{\lambda}\right) - c\big(1-e^{rt}\big)/r. \tag{16}$$

Define the boundary hitting times

$$T^{\theta}(u) = \min_{t\geq 0}\big\{t|\phi_{\theta}(t,u)\in\{0,-B\}\big\},$$

denoting the length of time until U_t^{θ} hits the boundary, where $\theta\in\{0,1\}$.

Claim 2. *The boundary hitting times T^0 and T^1 are differentiable in u and their minimum is quasi-convex.*

Proof. It follows from the implicit function theorem that T^1 and T^0 are differentiable. Define

$$T(u) = \min\big\{T^0(u), T^1(u)\big\}.$$

It is immediate that ϕ_1 and ϕ_0 are increasing in u. Therefore, an increase in u decreases $T^1(u)$ and increases $T^0(u)$. It follows that, T is quasi-convex, and T assumes its maximum at the point u_1^* at which $T^0(u) = T^1(u)$. Hence, $T^{0'}(u) < 0$ and $T^{1'}(u) > 0$ imply that

$$T'(u)\begin{cases} > 0 & \text{if } u < u_1^* \\ < 0 & \text{if } u > u_1^*. \end{cases} \tag{17}$$

□

Claim 3. *There is a unique value $\bar{u} < -\frac{c}{r\alpha}$ such that $\phi_1(T(\bar{u}),\bar{u}) = \bar{u}$.*

Proof. We show that for $u < -c/(r\alpha)$ there is a unique t solving $\phi_1(t,u) = u$ and that the solution is strictly decreasing in u. After a few simple operations, the identity $\phi_1(t,u) = u$ becomes

$$\underbrace{\frac{\alpha}{1-\alpha}\left(-\frac{\lambda}{r} - \frac{\lambda u}{c}\right)}_{=:\,\mathrm{LHS}} = \underbrace{\frac{e^{(r+\lambda)t}-1}{e^{rt}-1} - 1}_{=:\,\mathrm{RHS}}. \tag{18}$$

It is easy to see that RHS is increasing and convex in t, and that $\lim_{t\to 0}\frac{e^{(r+\lambda)t}-1}{e^{rt}-1} - 1 = \lambda/r$. LHS is clearly strictly decreasing in u, and for $u < -c/(r\alpha)$, we have

$$\frac{\alpha}{(1-\alpha)}\left(-\frac{\lambda}{r} - \frac{\lambda u}{c}\right) > \frac{\alpha}{(1-\alpha)}\left(-\frac{\lambda}{r} + \frac{\lambda}{c}\frac{c}{r\alpha}\right) = \frac{\lambda}{r}.$$

Therefore, for any $u < -c/(r\alpha)$, there is a unique time $T_s(u)$ such that $\phi_1(T_s(u),u) = u$. Moreover, inspection of (18) reveals that this time is continuous and strictly decreasing in u. Note that for $u\to -c/(r\alpha)$, we have $T_0(u) > T_s(u)\to 0$ and for $u\to -B + c/(\lambda\alpha)$, we have $T_s(u) > T^0(u)\to 0$. Because $T^0(u)$ is continuous and strictly increasing, and $T_s(u)$ is continuous and strictly decreasing, there must then exist a unique value $\bar{u}$ such that $T_s(\bar{u}) = T^0(\bar{u})$ and $\phi_1(T^0(\bar{u}),\bar{u}) = \bar{u}$. □

CLAIM 4. *We have $\phi_1(T(u), u) > u$ if $u > \bar{u}$ and $\phi_1(T(u), u) < u$ if $u < \bar{u}$.*

PROOF. Note that T_s is decreasing while T^0 is increasing. Moreover, $T_s(\bar{u}) = T^0(\bar{u})$ by construction. Thus, for $u > \bar{u}$, we have $T_s(u) < T^0(u)$, so that $\phi_1(T^0(u), u) > u$. Similarly, for $u < \bar{u}$, we have $T_s(u) > T^0(u)$, so that $\phi_1(T^0(u), u) < u$. □

Evolution of promised utilities during noncompliance We show that during reports of noncompliance, the utility of the agent is held constant Define $\beta_1 = \lambda\alpha/(r + \lambda\alpha)$.

CLAIM 5. *Let (K_n^0, K_n^1) be the principal's cost functions in an optimal mechanism when there are $n \geq 1$ available inspections. Denote the pair of initial promised utilities in this mechanism by $u^* = (u^{0*}, u^{1*})$. Then*

$$K_n^0(U_t) = \beta_1 K_n^1(u^*).$$

PROOF. Without loss, assume $\theta_0 = 1$. We establish the claim via contradiction. Suppose to the contrary that $K_n^0(U_t) > \beta_1 K_n^1(u^*)$, and consider the following alternative mechanism. For $\theta_t = 1$, let the new mechanism be identical to the original one. For $\theta_t^0 = 0$, we set $\mathrm{d}F_t = u^{0*} - U_t^0$ for $U_t^0 < u^{0*}$ and $\mathrm{d}F_t/\mathrm{d}t = ru^{0*}$ for $U_t^0 \geq u^{0*}$. In this new mechanism, for $\theta_t = 1$, the paths of promised utilities are identical to those in the original mechanism by construction, so that all incentive-compatibility constraints hold when $\theta_t = 1$. Moreover, since U_t^0 is strictly decreasing in t when $\theta_t = 1$, we have $U_t^0 < u^{0*}$ and, thus, $\mathrm{d}F_{t_1} > 0$. The promised utilities at $\theta_t = 0$ in the new mechanism are constant and equal to u^*, so that the obedience constraint is satisfied. Along the equilibrium paths, the expected cost for the principal at time t in state $\theta_t = 0$ in the new mechanism is, therefore,

$$\hat{K}_n^0(U_t) = \int_0^\infty e^{-(r+\lambda\alpha)s}\lambda\alpha K_n^1(u^*)\,\mathrm{d}s = \frac{\lambda\alpha}{r+\lambda\alpha}K_n^1(u^*) = \beta_1 K_n^1(u^*).$$

Since the new mechanism is identical to the original mechanism for $\theta_t = 1$, the expected costs for the principal in the new mechanism are strictly lower than in the original mechanism, contradicting optimality of the original mechanism. □

Derivation of the optimal mechanism in the auxiliary problem We now solve for the principal's value function iteratively by solving a sequence of impulse-control problems where the number of available inspections is bounded by a number n. We derive the optimal initial promised utility u_n^* for each n, and we show that the sequence $\{u_n^*\}$ converges to $\bar{u}$ as $n \to \infty$. For the case in which the agent reports $\hat{\theta}_t = 0$, Claim 5 implies that without loss the expected costs for the principal in state $\theta_t = 0$ with n available inspections can be written as $K_n^0(u_0, u_1) = \beta_1 K_n^1(u_1)$. We show that for all $n \geq 0$, the obedience constraint (O) binds at the outset.

CLAIM 6. *Suppose the total number of available inspections is n. Then there is an optimal policy such that at the initial pair of promised utility (u_n^0, u_n^1), the obedience constraint (O) binds.*

PROOF. Using Claim 5, there is no loss in generality in assuming that $\hat{\theta}_0 = 1$. Consider the optimal initial utilities (u^0, u^1), where we assume to the contrary that $u^1 - u^0 > c/(\lambda\alpha)$. Denote by t^* the minimizer of U_t^1. Let T be the first inspection time conditional on no transition, and let the promised utilities at that time be $\hat{u}^1$ and $\hat{u}^0$. Now fix $\epsilon > 0$ sufficiently small, and consider an alternative mechanism identical to the original mechanism, except that the first time of inspection is $(T + \epsilon)$, and with initial utilities $(\tilde{u}^1, \tilde{u}^0)$. If $T < t^*$, then let $\tilde{u}^0 = \tilde{u}^1 - u^1 + u^0$ and let $\tilde{u}^1$ solve

$$\hat{u}^1 = e^{r(T+\epsilon)}\big(\tilde{u}^1 + (1-\alpha)\big(e^{\lambda(T+\epsilon)} - 1\big)\big(u^1 - u^0\big)\big) - c\big(1 - e^{r(T+\epsilon)}\big)/r.$$

Thus, by shifting the initial promised utilities up, the first inspection date is postponed, while maintaining incentive compatibility and keeping the terminal values constant. Consequently, the initial utilities could not have been optimal. If $T \geq t^*$, then let $\tilde{u}^1 = u^1$ and let $\tilde{u}^0$ solve

$$\hat{u}^1 = e^{r(T+\epsilon)}\big(\tilde{u}^1 + (1-\alpha)\big(e^{\lambda(T+\epsilon)} - 1\big)\big(\tilde{u}^1 - \tilde{u}^0\big)\big) - c\big(1 - e^{r(T+\epsilon)}\big)/r.$$

Thus, by shifting u^0 up while keeping u^1 constant, the first inspection date can be postponed while maintaining incentive compatibility and keeping the terminal values constant. In either case, a pair of initial utilities with $u^1 - u^0 > c/(\lambda\alpha)$ cannot be optimal. □

Without loss, we can now restrict attention to initial pairs of utility (u^0, u^1) such that $u^1 - u^0 = c/(\lambda\alpha)$. Let $u = u^1$ denote the initial utility for the agent in the high state. The paths of promised utilities are then described by $\phi_0(t, u)$ and $\phi_1(t, u)$. Define

$$K_n^1(u) = \min_{\substack{0 \leq t \leq T(u) \\ u' \geq \phi_1(t,u)}} \int_0^t e^{-(r+\lambda(1-\alpha))s}\big(\lambda(1-\alpha)K_n^0\big)\,\mathrm{d}s + e^{-(r+\lambda(1-\alpha))t}\big(K_{n-1}^1\big(u'\big) + \kappa\big) \tag{19}$$

to be the maximum payoff for the principal at initial utility u for the agent, where the principal maximizes over stopping times and the post inspection utility u' resulting from the terminal promised utility $\phi_1(t, u)$ and a potential fine at the time of an inspection. Let u_n^* be a minimizer of K_n^1 and denote by t_n^* the associated first inspection date.

CLAIM 7. *Let u_{n-1}^* be a minimizer of $K_{n-1}^1(u)$ and suppose ${K_{n-1}^1}'(u) > 0$ for all $u > u_{n-1}^*$. Then $t_n^* = T^0(u_n^*)$ and $\phi_1(t_n^*, u_n^*) > u_{n-1}^*$.*

PROOF. First we show that $t_n^* = T^0(u_n^*)$. Suppose, to the contrary, that $t_n^* < T(u_n^*)$. If $\phi_1(t_n^*, u_n^*) > u_{n-1}^*$, then because ϕ_1 is strictly increasing in its second argument, we can find a lower initial utility $u < u_n^*$ such that $\phi_1(t_n^*, u) < \phi_1(t_n^*, u_n^*)$. Since ${K_{n-1}^1}'(\tilde{u}) > 0$ for $\tilde{u} > u_{n-1}^*$, we have $K_n^1(u) < K_n^1(u_n^*)$, contradicting optimality of u_n^*. If $\phi_1(t_n^*, u_n^*) \leq u_{n-1}^*$, then the optimal initial utility in step $n-1$ is $u' = -u_{n-1}^*$. We can thus find $t > t_n^*$ such that $\phi_1(t, u_n^*) < u_{n-1}^*$. Thus, the first inspection was delayed, while the continuation utility for the agent remains constant, contradicting optimality of u_n^*. Thus, we have $t_n^* = T^0(u_n^*)$. Now suppose $\phi_1(T^0(u_n^*), u_n^*) < u_{n-1}^*$. Then we can find a new initial utility $u > u_n^*$ such that $\phi_1(T^0(u), u) = u_{n-1}^*$. Since $T^0(\cdot)$ is increasing we have $T^0(u) > T^0(u_n^*)$, contradicting the optimality of u_n^*. □

In light of the result of Claim 7, there will be no loss in limiting our attention to the case $t = T(u)$ and $u' = \phi_1(t, u)$. The principal's expected costs for given utility u are, therefore,

$$K_n^1(u) = \int_0^{T(u)} e^{-(r+\lambda(1-\alpha))s}\big(\lambda(1-\alpha)K_n^0\big)\,\mathrm{d}s + e^{-(r+\lambda(1-\alpha))T(u)}\big(K_{n-1}^1\big(\phi_1(t,u)\big)+\kappa\big).$$

Define $\beta_0 = \lambda\alpha/(r+\lambda\alpha)$ and $\beta_1 = \lambda(1-\alpha)/(r+\lambda(1-\alpha))$. Solving the integrals and rearranging, the principal's payoff can be expressed more succinctly as

$$K_n^1(u) = a(u) + b(u)K_{n-1}^1\big(\phi_1\big(T(u),u\big)\big),$$

where

$$a(u) = \frac{e^{-(r+\lambda(1-\alpha))T(u)}}{1-\beta_0\beta_1+\beta_0\beta_1 e^{-(r+\lambda(1-\alpha))T(u)}}\kappa$$
$$b(u) = \frac{e^{-(r+\lambda(1-\alpha))T(u)}}{1-\beta_0\beta_1+\beta_0\beta_1 e^{-(r+\lambda(1-\alpha))T(u)}}.$$

Simple calculus reveals

$$a'(u) = -\frac{\big(e^{(r+\lambda-\alpha\lambda)T(u)}(r+\lambda-\alpha\lambda)^2(r+\alpha\lambda)\big(r(r+\lambda)\big)\big)}{\big((1-\alpha)\alpha\lambda e^{(r+\lambda-\alpha\lambda)T(u)}r(r+\lambda)\big)^2}\kappa T'(u) \quad \text{and}$$
$$b'(u) = -\frac{e^{(r+\lambda-\alpha\lambda)T(u)}r(r+\lambda)(r+\lambda-\alpha\lambda)^2(r+\alpha\lambda)}{\big((1-\alpha)\alpha\lambda^2+e^{(r+\lambda-\alpha\lambda)}r(r+\lambda)\big)^2}T'(u),$$

so that $\operatorname{sign} a'(u) = \operatorname{sign} b'(u) = \operatorname{sign} T'(u)$. From (17), it follows that

$$a'(u)\begin{cases}<0 & \text{if } u<u_1^*\\ >0 & \text{if } u>u_1^*\end{cases}, \qquad b'(u)\begin{cases}<0 & \text{if } u<u_1^*\\ >0 & \text{if } u>u_1^*.\end{cases}$$

Step 0 Consider the case $n = 0$, so the principal cannot perform any inspections. The principal has no way to incentivize the agent so that her value function is equal to the lower bound

$$K_0^\theta = \bar{K}.$$

Step 1 *Suppose the principal can inspect at most once, so that* $n = 1$. Let t be the first inspection if no transition occurs; let u be the initial utility for the agent. The expected costs for the principal when inspecting at time t are

$$K_1^1(u) = a(u) + b(u)\bar{K}.$$

The marginal cost increase in utility u is

$$K_1^{1'}(u) = a'(u) + b'(u)\bar{K}.$$

We have $K_1^{1'}(u) < 0$ for $u < u_1{}^*$ and $K_1^{1'}(u) > 0$ for $u > u_1^*$, with $u_1^* > \bar{u}$. Thus $u_1{}^*$ minimizes K_1^1.

Step 2 Suppose there are two inspections left to be performed. The principal's payoff can be written as

$$K_2^1(u) = a(u) + b(u)K_1^1\big(\phi_1\big(T(u), u\big)\big).$$

When $u > u_1^*$, then $(K_2^1)'(u) > 0$ and, therefore, the optimizer does not exceed u_1^*. Because $K_2^1(u)$ is maximal when u lies at the participation boundary and is continuous in between, there must be a minimizer u_2^*. The marginal cost increase is

$$K_2^{1'}(u) = a'(u) + b'(u)K_1\big(\phi_1\big(T(u), u\big)\big) + b(u)D_u\phi_1\big(T(u), u\big))K_1^{1'}\big(\phi_1\big(T(u), u\big)\big).$$

Here $D_u\phi_1(T(u), u)$ is the total derivative of $\phi_1(T(u), u)$ with respect to u, which can be shown to be

$$D_u\phi_1\big(T(u), u\big) = e^{rT(u)}\left(1 + T'(u)\left(c\big(e^{\lambda T(u)} - 1\big)\frac{1-\alpha}{\alpha}\frac{r}{\lambda} + ru + c\left(e^{\lambda T(u)}\frac{1-\alpha}{\alpha} + 1\right)\right)\right) > 0.$$

Thus, for $u > u_1^*$ $(> \bar{u})$,

$$\begin{aligned} K_2^{1'}(u) &= a'(u) + b'(u)K_1\big(\phi_1\big(T(u), u\big)\big) + b(u)D_u\phi_1\big(T(u), u\big))K_1'\big(\phi_1\big(T(u), u\big)\big) \\ &> a'(u) + b'(u)K_1\big(\phi_1\big(T(u), u\big)\big) > a'(u) + b'(u)\bar{K} = K_1^{1'}(u). \end{aligned}$$

In particular, this means $u_2^* < u_1^*$.

Step n $K_n^1(u) = a(u) + b(u)K_n^1(\phi_1(T(u), u))$ *has a minimum at* u_n^*. The marginal cost at $u > u_{n-1}^*$ $(> \bar{u})$ is

$$\begin{aligned} K_n^{1'}(u) &= a'(u) + b'(u)K_{n-1}\big(\phi_1\big(T(u), u\big)\big) + b(u)D_u\phi_1\big(T(u), u\big))K_{n-1}'\big(\phi_1\big(T(u), u\big)\big) \\ &< a'(u) + b'(u)K_{n-1}^1(u) + b(u)D_u\phi_1\big(T(u), u\big))K_{n-2}'\big(\phi_1\big(T(u), u\big)\big) \\ &< a'(u) + b'(u)K_{n-2}^1(u) + b(u)D_u\phi_1\big(T(u), u\big))K_{n-2}'\big(\phi_1\big(T(u), u\big)\big), \end{aligned}$$

where the first line follows from our induction hypothesis. Therefore, $u \geq u_{n-1}^*$ implies $K_n^{1'}(u) < K_{n-1}^{1\,\prime}(u) < 0$.

The induction shows that $u_n^* < u_{n-1}^*$ for all $n \geq 0$. It follows immediately from the definition of $\bar{u}$ that $u_n^* > \bar{u}$ for all n. Hence, $\{u_n^*\}$ is a decreasing and bounded sequence, so that by the monotone convergence theorem, the sequence converges to a limit $\hat{u} \geq \bar{u}$. Since $\{u_n^*\}$ is convergent, it is a Cauchy sequence, so that by Claim 3,

$$\lim_{n\to\infty}\big|u_n^* - u_{n-1}^*\big| = \lim_{n\to\infty}\big|u_k^* - \phi_1\big(T\big(u_n^*\big), u_n^*\big)\big| = 0 \quad \Rightarrow \quad \hat{u} = \bar{u}. \qquad \square$$

This establishes the mechanism characterized in Theorem 1 as the optimum in the auxiliary problem. We now verify that it is also optimal in the original problem.

B.2 *Verification for the proof of Theorem 1*

B.2.1 *No fines between inspections* We now show that the mechanism described in the previous section remains optimal when we remove assumption (A). To this end, we show

that when performing the iteration over the number of available inspections n, the principal cannot gain from imposing fines between inspections when n inspections are left. Consider again Step n of the iteration in the previous section. By the same argument as before, we have $u^1 - u^0 = c/(\lambda\alpha)$ and the first time of inspection is at the first time t at which $U_t^0 = -B$. The evolution of the paths of promised utilities are given by

$$\begin{aligned} \mathrm{d}U_t^1 &= rU_t^1\,\mathrm{d}t - \lambda(1-\alpha)\big(U_t^1 - U_t^0\big)\,\mathrm{d}t + c\,\mathrm{d}t + \mathrm{d}F_t \\ \mathrm{d}U_t^0 &= rU_t^0\,\mathrm{d}t - \lambda\alpha\big(U_t^1 - U_t^0\big)\,\mathrm{d}t + c\,\mathrm{d}t + \mathrm{d}F_t - \mathrm{d}\mu_t, \end{aligned}$$

where we let $\mathrm{d}\mu_t \geq 0$ denote the slacking in the honesty constraint. The evolution of the difference in utilities is

$$\mathrm{d}\big(U_t^1 - U_t^0\big) = (r+\lambda)\big(u^1 - u^0\big) + \mathrm{d}\mu_t,$$

which implies that the utility paths diverge at least exponentially, and are independent of any fines and increasing in threats. If the first inspection takes place at t, conditional on no transition before t, this means that $U_t^0 = -B$ and

$$U_t^1 = -B + e^{(r+\lambda)t}\frac{c}{\lambda\alpha} + \int_0^t e^{(r+\lambda)s}\,\mathrm{d}\mu_s.$$

The last term has to be zero because otherwise we could find a pair of initial promised utilities with $\hat{u}^1 < u^1$ and set $\mathrm{d}\mu_s = 0$ for all $s \in (0, t)$, and a time $t' > t$ such that the promised utilities at time t' under the new initial conditions are as with the original pair at time t, thus increasing the principal's payoff. Therefore, at the first time of inspection,

$$U_t^1 = -B + e^{(r+\lambda)t}\frac{c}{\lambda\alpha}.$$

Given that U_1^t is independent of any fines in step n, and there no fines in step $n-1$ onward, we must have $U_t^1 = \phi_1(T(u_n^*), u_n^*)$. This means that the policy of the previous section with initial promised utilities $(u_n^*, u_n^* - c/(\lambda\alpha))$ remains optimal even when fines between inspections are available.

B.2.2 *General mechanisms in the relaxed problem* Parts B.1 and B.2.1 demonstrate that the mechanism described in the theorem is an optimal Markovian mechanism under the relaxing of assumption (B). It remains to verify that no (non-Markovian) mechanism can do better. Let $K^{\theta_t}(U)$ denote the expected costs for the principal in our mechanism that delivers the agent with promised payoffs of $U = (U^0, U^1)$. We show that the expected value in state θ_t from any incentive-compatible mechanism that delivers the initial promised payoff $U_0 = (U_0^0, U_0^1)$ to the agent cannot exceed $K^{\theta_t}(U_0)$. Since both the inspection cost and the set of feasible continuation utilities do not depend on their values prior to inspection, we can apply Proposition 54.18 and Theorem 54.28 in Davis (1993, pp. 235 & 242) to conclude that K_n, the value function with no more than n inspections, converges to value function K of the problem without bound on the number

of inspections, and that K is the unique bounded and continuous function that solves the quasi-variational inequality

$$\begin{aligned}\mathcal{U}K^\theta(u) - rK^\theta(u) &\geq 0\\ WK^\theta(u) - K^\theta(u) &\geq 0\\ \big(\mathcal{U}K^\theta(u) - rK^\theta(u)\big)\big(WK^\theta(u) - K^\theta(u)\big) &= 0\end{aligned}$$

on the state space $\{(\theta, u^0, u^1) | \theta \in \{0, 1\}, (u^0, u^1) \in [-B, 0]^2, u^1 - u^0 \geq c/(\lambda\alpha)\}$. Here, $\mathcal{U}$ denotes the extended generator of the piecewise deterministic Markov process that is defined by the relationship[35]

$$\mathbb{E}_0^P\big[K^{\theta_t}(U_t)\big] = K^{\theta_0}(u) + \mathbb{E}_0^P\left[\int_0^t \mathcal{U}K^{\theta_s}(u_s)\,\mathrm{d}s\right]$$

in case no inspection occurs before t, and W is the expected cost at an inspection time:

$$WK^\theta = \min_{u_0, u_1} K^\theta(u_0, u_1) + \kappa.$$

Consider an arbitrary incentive-compatible mechanism with inspection process $\{\mathrm{d}N_t^I\}_t$ and define the expected value at time t by

$$G_t = \int_0^t e^{-rs}\big(\kappa\,\mathrm{d}N_s^I\big) + e^{-rt}K^{\theta_t}(U_t).$$

For $t = 0$, we have $G_0 = K^{\theta_0}(U_0)$. For $t > 0$, we can represent G_t by the differential formula (see Theorem 31.3 in Davis (1993, p. 83)) as

$$\begin{aligned}\mathbb{E}_s[G_t] - G_s = &\int_s^t e^{-r(z-s)}\big(\mathcal{U}K^{\theta_z}(U_z) - rK^{\theta_z}(U_z)\big)\,\mathrm{d}z\\ &+ \mathbb{E}_s\left[\int_s^t e^{-r(z-s)}\big(WK^{\theta_z}(U_z) - K^{\theta_z}(U_z)\big)\,\mathrm{d}N_z^I\right].\end{aligned}$$

By the variation inequality above, both integrals are positive so that the process $(G_t)_t \geq 0$ is a submartingale bounded by 0. This implies that $E_0[G_t] \geq G_0$ for any $t \geq 0$. In particular, taking the limit as t approaches infinity, we get $E_0[\int_0^\infty e^{-rs}(\kappa\,\mathrm{d}N_s^I)] = E_0[\lim_{t\to\infty} G_t] \geq G_0 = K^{\theta_0}(U_0)$. Hence, any incentive-compatible maximal-compliance mechanism leads to weakly higher inspection costs.

B.2.3 *Optimality in the original problem* We now consider the original model, in which we remove assumption (A) so that the honesty constraint must hold in both states. We show that during noncompliance, the honesty constraint does not bind and, therefore, the solution of the relaxed problem is also a solution to our original problem. The proof is constructive. In the optimal mechanism of the relaxed problem, the pair of promised utilities at the outset and during noncompliance is $(u^0, u^1) := (\bar{u}, \bar{u} - c/(\lambda\alpha))$.

[35]See Davis (1993, pp. 27–33).

Since $\mathrm{d}U_t^0 \le 0$, we have $U_t^0 \le u^0$. Set $\mathrm{d}F_t = u^0 - U_t^0$ and $\mathrm{d}\mu_t = u^1 - U_t^1 + u^0 - U_t^0$. Next, while $\theta = 0$, set $\mathrm{d}F_t = -ru^0 + \lambda\alpha(u^1 - u^0)$ and $\mathrm{d}\mu_t = c + (r+\lambda)(u^1 - u^0)$. Substituting into the promise-keeping and truth-telling constraints, it follows that $\mathrm{d}U_t^i = 0$ for each $i = 0, 1$ and $\mathrm{d}N_t^I = 0$ while $\theta_t = 0$, which is identical to the solution in the relaxed problem. □

Proof of Lemma 3

Define

$$\Psi(T) \equiv (B - c/r)\big(1 - e^{-rT}\big) - c/(\lambda\alpha)e^{\lambda T}\big(e^{rT} - \alpha\big) + c/(\lambda\alpha)(1-\alpha). \tag{20}$$

By Theorem 1, we have $T^* = \inf\{T > 0 : \Psi(T) = 0\}$. This exists and is unique whenever $B > \bar{B}$ (Ψ is increasing from 0 at $T = 0$ and crosses 0 from above exactly once). The function Ψ is continuously differentiable in λ and T on a neighborhood of T^*. By the implicit function theorem, we have

$$\frac{\partial T^*}{\partial \lambda} = -\frac{\Psi_\lambda}{\Psi_T}\bigg|_{T=T^*},$$

where Ψ_x denotes the partial derivative of Ψ with respect to x. As mentioned above, $\Psi(T)$ crosses 0 from above at $T = T^*$ so that $\Psi_T|_{T=T^*} < 0$. Hence, for all parameters, we have

$$\operatorname{sign}\left(\frac{\partial T^*}{\partial \lambda}\right) = \operatorname{sign}(\Psi_\lambda|_{T=T^*}).$$

Consider Ψ in (20) as $\lambda \searrow \underline{\lambda} = cr/(Br\alpha - c)$, which is the lower bound on λ such that the feasibility assumption $B > \bar{B} = \frac{c(r+\lambda)}{r\lambda\alpha}$ is fulfilled. Then Ψ is equal to

$$\left(B - \frac{c}{r}\right)\big(1 - e^{-rT}\big) - \left(B - \frac{c}{r\alpha}\right)\big(e^{\underline{\lambda}T}\big(e^{rT} - \alpha\big) - (1-\alpha)\big).$$

This can be equal to 0 only if $T = 0$ because it is concave in T and the T derivative is 0 at $T = 0$. Hence, $\lim_{\lambda \downarrow \underline{\lambda}} T^*(\lambda) = 0$ and T^* is initially increasing in λ.

Finally, to show that $T^*(\lambda) \overset{\lambda\to\infty}{\longrightarrow} 0$, consider (20) and observe that $\Psi(T^*) = 0$ implies

$$\lim_{\lambda\to\infty} \frac{e^{(r+\lambda)T^*(\lambda)}}{\lambda} = 0.$$

This implies that $\lambda T^*(\lambda)$ is either finite or grows at lower than logarithmic rate as λ becomes arbitrarily large. In particular, $T^*(\lambda)$ must go to 0.

Considering the costs, let Let K_{EQ}^0 and K_{EQ}^1 denote the expected discounted inspection cost when starting in state 0 or 1, respectively. For fixed inspection cycle length T, they follow the nested equations

$$K_{\mathrm{EQ}}^0 = \int_0^\infty e^{-(r+\lambda\alpha)t}\lambda\alpha K_{\mathrm{EQ}}^1\,\mathrm{d}t = \frac{\lambda\alpha}{r+\lambda\alpha}K_{\mathrm{EQ}}^1$$

and

$$K^1_{\mathrm{EQ}} = \int_0^T e^{-(r+\lambda(1-\alpha))t}\lambda(1-\alpha)K^0_{\mathrm{EQ}}\,\mathrm{d}t + e^{-(r+\lambda(1-\alpha))T}\big(\kappa + K^1_{\mathrm{EQ}}\big)$$
$$= \big(1 - e^{(r+\lambda(1-\alpha))T}\big)\frac{\lambda(1-\alpha)}{r+\lambda(1-\alpha)}K^0_{\mathrm{EQ}} + e^{-(r+\lambda(1-\alpha))T}\big(\kappa + K^1_{\mathrm{EQ}}\big).$$

Inserting K^0_{EQ} and solving for K^1_{EQ} gives

$$K^1_{\mathrm{EQ}} = \frac{r+\lambda\alpha}{r(r+\lambda)} \cdot \big(r+\lambda(1-\alpha)\big)\frac{e^{-(r+\lambda(1-\alpha))T}}{1-e^{-(r+\lambda(1-\alpha))T}} \cdot \kappa.$$

Note that K^1_{EQ} is decreasing in T and decreasing in λ for fixed T. Thus, given that T^* is increasing in λ for low λ, it follows immediately that the costs decrease for low λ. Further, K^1_{EQ} approaches ∞ as T goes to zero for any positive and finite λ. Since

$$\lim_{\lambda \searrow \underline{\lambda}} T^*(\lambda) = 0,$$

it follows that

$$\lim_{\lambda \searrow \underline{\lambda}} K_{\mathrm{EQ}}(\lambda) = 0.$$

For the limit as λ grows arbitrarily large, note that the total cost in the limit is given by

$$\lim_{\lambda\to\infty} K^1_{\mathrm{EQ}} = \frac{(1-\alpha)\alpha}{r} \lim_{\lambda\to\infty} \frac{\lambda}{e^{(1-\alpha)\lambda T^*(\lambda)}}.$$

Recall from above that $\lambda T^*(\lambda)$ grows to ∞ at lower than logarithmic rate so that the term above must be ∞. □

Appendix C: Proofs for Section 5

Proof of Theorem 2. We first argue that as the mechanism described in the main text is optimal among the class of stationary mechanisms, consider an alternative stationary stochastic mechanism that delivers some given promised utility u. From the promise-keeping constraint (Pk) and the honesty constraint (H) for $i = 1$, it is straightforward to obtain that the constant rate $m_R(u)$ of inspection that keeps the promised utility in state 1 stationary at the level $u \in [-B + c/(\lambda\alpha), -c/(r\alpha)]$ is $m_R(u) = r(c - \alpha\lambda u)/(\alpha\lambda(B+u) - c)$.

The principal's expected monitoring costs in the stationary random mechanism that provides promised utility $U^1_t = u$ throughout can be determined recursively. Denoting by $K^1_R(u)$ the expected costs while in compliance, we have

$$K^1_R(u) = \frac{r+\lambda\alpha}{r}\frac{m_R(u)}{r+\lambda} = \frac{r+\lambda\alpha}{r}\frac{r(c-\alpha\lambda u)}{(r+\lambda)\big(\alpha\lambda(B+u)-c\big)}. \tag{21}$$

It is easy to see that $K_R(u)$ is decreasing in u. Given that $-\frac{c}{r\alpha}$ is an upper bound on the promised utility for the agent during compliance (it is the maximum payoff for the agent subject to satisfying the obedience constraint) and it is the promised utility delivered by the mechanism characterized above, it follows that this mechanism is indeed the optimal stationary mechanism.

To show that the costs of the random mechanism are strictly below the equilibrium costs with predictable inspections, express the latter as

$$K^0_{\mathrm{EQ}} = \int_0^\infty e^{-(r+\lambda)t}\lambda\big(\alpha K^1_{\mathrm{EQ}} + (1-\alpha)K^0_{\mathrm{EQ}}\big)\,\mathrm{d}t$$
$$K^1_{\mathrm{EQ}} = \int_0^\infty e^{-(r+\lambda)t}\lambda\big(\alpha \tilde{K}^1(\tau_t) + (1-\alpha)K^0_{\mathrm{EQ}}\big)\,\mathrm{d}t + \sum_{n=1}^\infty e^{-(r+\lambda)kT^*}\kappa,$$

where K^0_{EQ} denotes the expected costs while in noncompliance and

$$\tilde{K}^1(\tau) = e^{-(r+\lambda(1-\alpha))(T^*-\tau)}\big(\kappa + K^1_{\mathrm{EQ}}\big) + \int_\tau^{T^*} e^{-(r+\lambda(1-\alpha))(s-\tau)}\lambda(1-\alpha)K^0_{\mathrm{EQ}}\,\mathrm{d}s$$

denotes the expected costs while in compliance and time $\tau \in [0, T^*]$ has passed since the last inspection or transition. Note that $\tilde{K}^1(\tau)$ is increasing in τ with $\tilde{K}^1(0) = K^1_{\mathrm{EQ}}$ and $\tilde{K}^1(T^*) = \kappa + K^1_{\mathrm{EQ}}$. Thus, replacing $\tilde{K}^1(\tau_t)$ by K^1_{EQ} in the recursive expression above, and solving the system gives an upper bound on the equilibrium costs K^1_{EG}:

$$K^1_{\mathrm{EQ}} \le \tilde{K} = \frac{r+\lambda\alpha}{r}\frac{e^{-(r+\lambda)T^*}}{1-e^{-(r+\lambda)T^*}}\kappa. \tag{22}$$

To see that K^1_R in (21) is lower, use (7) to write T^* in (22) as a function of u^{1*}:

$$\tilde{K}\big(u^{1*}\big) = \frac{r+\lambda\alpha}{r}\frac{e^{-(r+\lambda)T^*}}{1-e^{-(r+\lambda)T^*}}\kappa = \frac{r+\lambda\alpha}{r}\frac{c}{\alpha\lambda\big(B+u^{1*}\big)-c}\kappa.$$

Now it is immediate to check that $\tilde{K}(-\frac{c}{r\alpha}) = K_R(-\frac{c}{r\alpha})$ and $\tilde{K}'(u) < K'_R(u)$ for $u < -\frac{c}{r\alpha}$ and $B > \bar{B}$. Since $-c/(r\alpha)$ is an upper bound on u, it follows that $K_R(u^{1*}) < \tilde{K}(u^{1*})$. It follows from (22) that $K^1_{\mathrm{EQ}} > K^1_R$.

For the comparative statics in λ, consider (21) and observe that K^1_R is decreasing in λ for fixed m_R and decreasing in $m_R(u)$. The optimal promised utility $-c/(r\alpha)$ does not change with λ and $m_R(u)$ is decreasing in λ for any u.

For the limit, observe that

$$\lim_{\lambda\to\infty} m_R\left(-\frac{c}{r\alpha}\right) = \frac{cr}{Br\alpha - c}$$

and, thus,

$$\lim_{\lambda\to\infty} K^*_R = \frac{\alpha}{r}\frac{cr}{Br\alpha - c}\kappa = \frac{c\alpha}{Br\alpha - c}\kappa. \qquad \square$$

References

Antinolfi, Gaetano and Francesco Carli (2015), "Costly monitoring, dynamic incentives, and default." *Journal of Economic Theory*, 159, 105–119. [826]

BaFin (2016), "Richtlinie zur Durchführung und Qualitätssicherung der laufenden Überwachung der Kredit- und Finanzdienstleistungsinstitute durch die Deutsche Bundesbank". [839]

Ball, Ian and Jan Knoepfle (2023), "Should the timing of inspections be predictable?" Working paper, arXiv:2304.01385. [826]

Becker, Gary S. (1968), "Crime and punishment: An economic approach." *Journal of Political Economy*, 76, 169–217. [837]

Ben-Porath, Elchanan and Michael Kahneman (2003), "Communication in repeated games with costly monitoring." *Games and Economic Behavior*, 44, 227–250. [827]

Blundell, Wesley, Gautam Gowrisankaran, and Ashley Langer (2020), "Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations." *American Economic Review*, 110, 2558–2585. [824]

Board, Simon and Moritz Meyer-ter-Vehn (2013), "Reputation for quality." *Econometrica*, 81, 2381–2462. [826]

Border, Kim C. and Joel Sobel (1987), "Samurai accountant: A theory of auditing and plunder." *The Review of Economic Studies*, 54, 525–540. [826]

Chang, Chun (1990), "The dynamic structure of optimal debt contracts." *Journal of Economic theory*, 52, 68–86. [826]

Chen, Mingliu, Peng Sun, and Yongbo Xiao (2020), "Optimal monitoring schedule in dynamic contracts." *Operations Research*, 68, 1285–1314. [826]

Dai, Liang, Yenan Wang, and Ming Yang (2022), "Dynamic contracting with flexible monitoring." Working paper. SSRN 3496785. [826]

Davis, Mark H. (1993), *Markov Models and Optimization*, volume 49 of Monographs on Statistics and Applied Probability. Chapman & Hall. [830, 834, 842, 850, 856, 857]

di Giovanni, Julian et al. (2020), "Firm-level shocks and GDP growth: The case of Boeing's 737 max production pause." Technical report, Federal Reserve Bank of New York. [823]

Dilmé, Francesc and Daniel F. Garrett (2019), "Residual deterrence." *Journal of the European Economic Association*, 17, 1654–1686. [826]

European Commission (2019), "Commission Recommendation (EU) 2019/1318 on internal compliance programmes for dual-use trade controls under Council Regulation (EC) No 428/2009." European Commission, 30 July 2019. [824]

Fernandes, Ana and Christopher Phelan (2000), "A recursive formulation for repeated agency with history dependence." *Journal of Economic Theory*, 91, 223–247. [830, 831]

Gale, Douglas and Martin Hellwig (1985), "Incentive-compatible debt contracts: The one-period problem." *The Review of Economic Studies*, 52, 647–663. [826]

Halac, Marina and Andrea Prat (2016), "Managerial attention and worker performance." *American Economic Review*, 106, 3104–3132. [826]

Kamada, Yuichiro and Neel Rao (2023), "Strategies in stochastic continuous-time games." Technical report, Working Paper UTMD-040, University of Tokyo. [828, 829]

Kaplow, Louis and Steven Shavell (1994), "Optimal law enforcement with self-reporting of behavior." *Journal of Political Economy*, 102, 583–606. [837]

Kapon, Sam (2022), "Dynamic amnesty programs." *American Economic Review*, 112, 4041–4075. [824]

Kim, Sang-Hyun (2015), "Time to come clean? Disclosure and inspection policies for green production." *Operations Research*, 63, 1–20. [826]

Kitroeff, Natalie, David Gelles, and Jack Nicas (2019), "The roots of Boeing's 737 max crisis: A regulator relaxes its oversight." [823]

Last, Günter and Andreas Brandt (1995), *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer. [830, 843]

Li, Anqi and Ming Yang (2020), "Optimal incentive contract with endogenous monitoring technology." *Theoretical Economics*, 15, 1135–1173. [826]

Ljungqvist, Lars and Thomas J. Sargent (2018), *Recursive Macroeconomic Theory*. MIT Press. [825]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford university press. [825]

Malenko, Andrey (2019), "Optimal dynamic capital budgeting." *The Review of Economic Studies*, 86, 1747–1778. [826]

Monnet, Cyril and Erwan Quintin (2005), "Optimal contracts in a dynamic costly state verification model." *Economic Theory*, 26, 867–885. [826]

Mookherjee, Dilip and Ivan Png (1989), "Optimal auditing, insurance, and redistribution." *The Quarterly Journal of Economics*, 104, 399–415. [826]

Piskorski, Tomasz and Mark M. Westerfield (2016), "Optimal dynamic contracts with moral hazard and costly monitoring." *Journal of Economic Theory*, 166, 242–281. [826]

Popov, Latchezar (2016), "Stochastic costly state verification and dynamic contracts." *Journal of Economic Dynamics and Control*, 64, 1–22. [826]

Ramos, Joao and Tomasz Sadzik (2023), "Partnership with persistence." Working paper. [827]

Ravikumar, B. and Yuzhe Zhang (2012), "Optimal auditing and insurance in a dynamic model of tax compliance." *Theoretical Economics*, 7, 241–282. [826]

Reinganum, Jennifer F. and Louis L. Wilde (1985), "Income tax compliance in a principal-agent framework." *Journal of Public Economics*, 26, 1–18. [827]

Townsend, Robert M. (1979), "Optimal contracts and competitive markets with costly state verification." *Journal of Economic Theory*, 21, 265–293. [826]

U.S. House of Representatives (2020), "Final Committee Report on the Design, Development, and Certification of the Boeing 737 MAX. Committee on transportation and infrastructure. Subcommittee on aviation". *116th Congress, 2nd session*, September, 2020. [823]

Varas, Felipe, Iván Marinovic, and Andrzej Skrzypacz (2020), "Random inspections and periodic reviews: Optimal dynamic monitoring." *The Review of Economic Studies*, 87, 2893–2937. [826, 837, 840, 841]

Wang, Cheng (2005), "Dynamic costly state verification." *Economic Theory*, 25, 887–916. [826]

Webb, David C. (1992), "Two-period financial contracts with private information and costly state verification." *The Quarterly Journal of Economics*, 107, 1113–1123. [826]

Wong, Yu F. (2022), "Dynamic monitoring design." Working paper. [826]

Zhang, Yuzhe (2009), "Dynamic contracting with persistent shocks." *Journal of Economic Theory*, 144, 635–675. [830]

Co-editor Simon Board handled this manuscript.

Manuscript received 24 January, 2022; final version accepted 19 July, 2023; available online 25 July, 2023.