Discussion Paper Series – CRC TR 224

Discussion Paper No. 295
Project B 04

Relational Enforcement

Peter Wagner [1]
Jan Knoepfle [2]

April 2021

[1] Dept. of Economics, University of York. Email: peter.wagner@york.ac.uk
[2] Aalto University School of Business. 02150 Espoo, Finland. Email: jan.knoepfle@aalto.fi

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Relational enforcement*

Peter Wagner[†] Jan Knoepfle[‡]

April 20, 2021

Abstract

This paper studies a principal who incentivizes an agent to achieve and maintain *compliance* and voluntarily disclose incidences of non-compliance. Compliance is modeled as a persistent binary process that jumps at random times arriving at a rate that depends on the agent's efforts. The state of compliance is verifiable by the principal only at isolated instances through costly inspections. We show that in principal-optimal equilibria, the principal attains maximum compliance by using deterministic inspections. The optimal equilibrium features periodic inspection cycles which are suspended during periods of self-reported non-compliance, in which the agent is fined. We explain how commitment to random inspections benefits the principal by relaxing the agent's incentive-compatibility constraints, and we discuss possible ways for the principal to overcome her commitment problem through third-party involvement.

1 Introduction

Trust rarely endures at arm's length without oversight. In 2010, Boeing decided to redesign its best-selling 737 to create a newer model, the 737 MAX, in response to its arch rival's fuel-efficient newcomer, the Airbus 320 NEO. The redesign required fitting larger engines on the 737 which was a challenge due to the aircraft's low profile. The necessary adjustments to the position of the engines altered the aerodynamics of the aircraft in a way that could result in a stall under certain flight conditions, a risk that Boeing engineers tried to mitigate with a software fix. This software turned out to be prone to failure, however; a defect that ultimately caused a fatal crash in 2018 and a second one 5 months later. These accidents killed 346 people and led to a worldwide grounding of the remaining aircraft. As a result, Boeing suffered an operational loss in excess of $20 billion, and the estimated impact on the U.S. economy as a whole was a devastating 0.4 percentage points loss in GDP growth (di Giovanni et al., 2020). An investigation by U.S. Congress concluded that the accidents were to a large extent due to "grossly insufficient oversight by the FAA."[1] In the early 2000s, the FAA began to increasingly trust

*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 Econometric Society Summer Meeting in St Louis in 2017, as well as seminar participants at Bonn and HECER. Peter Wagner 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. Jan Knoepfle acknowledges financial support by the German Research Foundation (DFG) through CRC TR 224 (Project B04).

†Dept. of Economics, University of York. Email: peter.wagner@york.ac.uk

‡Aalto University School of Business. 02150 Espoo, Finland. Email: jan.knoepfle@aalto.fi

[1] U.S. Congress. House. Committee on transportation and infrastructure. Subcommittee on aviation. *Final Committee Report on the Design, Development, and Certification of the Boeing 737 MAX.* September 2020. 116th Congress, 2nd session. https://transportation.house.gov/committee-activity/boeing-737-max-investigation

manufacturers to certify their own planes as a cost-saving measure. By 2018, Boeing had self-certified nearly all of its work (Kitroeff et al., 2019). In absence of oversight, Boeing responded to competitive pressure by rushing development and production at the expense of safety.

The case of Boeing's 737 MAX illustrates the potential risks of delegating responsibilities without maintaining sufficient oversight. Oversight is needed to provide incentives for compliance, but in general, it is difficult to say how much oversight is enough. How much should manufacturers be monitored to ensure that they adhere to safety and environmental regulation? How much should service industries be monitored for potential consumer-protection or privacy violations? How closely should banks be watched to ensure that they maintain functioning internal audit systems and control their risk exposure? Questions of adequate oversight are relevant not only for regulatory enforcement but also arise frequently in the private sector. Retail firms, for example, must decide how often to inspect franchise outlets to ensure that they comply with quality and operating requirements (Martin, 1988; Lafontaine and Slade, 1996). Banks must decide how closely to monitor borrowers to verify their appropriate use of funds (Diamond, 1984; Antinolfi and Carli, 2015). Similar questions also arise in the enforcement of international agreements, such as disarmament or financial-aid treaties. A critical issue in all of these settings is the tension between trust and oversight. In a principal-agent relationship, if the principal trusts that the agent is compliant, what incentive is there to maintain oversight? With too little oversight, why would the agent exert effort to remain compliant? Theoretical work predicts that when monitoring is costly, compliance is attainable if the principal has commitment power (Holmstrom, 1979). Without commitment power, however, the relationship must inevitably involve some degree of distrust and cheating (Reinganum and Wilde, 1985; Avenhaus et al., 2002; Dilmé and Garrett, 2019).

We explore the tension between trust and oversight in a continuous-time relational contracting framework between a principal and an agent. Our setup entails two critical modeling assumptions. First, we assume that the agent's effort has a delayed and persistent effect on compliance. To achieve this, we model compliance as a binary Markov process which jumps at random times arriving at a rate that depends on the agent's effort choices. Prolonged periods of shirking will inevitably lead to a state of non-compliance, which the principal can detect by performing costly inspections. Our model thus exhibits both hidden action as well as hidden information (Levin, 2003). Second, we consider an environment with extreme conflict of interest between the principal and the agent in that only the principal benefits from the agent's compliance, and compliance is hard to observe and non-contractible, so that the information asymmetry cannot be easily mitigated through performance contracts. To generate incentives, we assume that the principal can impose limited sanctions —reductions in the agent's promised utilities, which we can interpret literally as monetary payments (i.e., fines), or as resulting from other, non-monetary measures such as increasing bureaucratic burden, reputational harm, limits to trade, or restricting access to resources.

We focus on principal-optimal equilibria of this environment. Our main result is that there exist equilibria in which the agent truthfully discloses all changes in compliance and exerts maximum effort throughout. The novel insight is that the principal can induce the agent to work and report truthfully despite her lack of commitment, provided the agent's effort has a delayed and persistent effect on compliance. The persistence allows the principal to deter the agent from deviations through isolated inspections and threats of penalties. The principal's motive to monitor in these equilibria is derived from her desire to maintain a reputation for vigilance. Our notion of reputation follows Barro and Gordon (1983), Canzoneri (1985), and Ljungqvist and Sargent (2018), and it is in the sense of maintaining reputation by following equilibrium actions.[2] As long as the principal monitors as prescribed by her equilibrium strategy, the agent will continue to expect to be monitored in the future, and therefore have an incentive to exert effort and report changes in compliance truthfully. When the principal monitors insufficiently, in such a way that the agent detects a deviation, the agent will infer that the principal has become non-vigilant, which in turn induces the agent to shirk and ultimately leads to a breakdown of the relationship.

In the class of principal-optimal equilibria we construct, inspections are entirely predictable for the agent. This result implies in particular that the principal cannot gain from randomized inspections. Intuitively, inspecting at random without commitment is ineffective because it renders deviations by the principal undetectable for the agent. To deter the principal from delaying inspections, later inspections must be followed by continuation play which is increasingly less favorable for the principal. In equilibrium, the principal is willing to randomize if she receives the same expected payoff from inspecting at any date in the support of her inspection strategy. This also means that she would be equally well off when inspecting at the earliest possible date in this support. We use this observation and show that we can modify the equilibrium strategies successively to obtain an alternative equilibrium in which the principal uses a non-random inspection strategy that involves performing an inspection with certainty at the earliest possible realization of the original (mixed) strategy while maintaining the agent's incentive-compatibility conditions.

To construct the principal-optimal equilibrium, we show that equilibrium payoffs coincide with the value of an auxiliary mechanism-design problem in which the principal is restricted to non-random inspections. We then transform this optimization into a dynamic programming problem which uses the agent's *promised utility* as state variable. The principal-optimal equilibrium we derive entails two phases that depend on the agent's report: a penalty phase and a monitoring phase. The agent is in the penalty phase when he reports a failure of compliance. He then pays a constant flow fine but is never inspected. The agent enters the monitoring phase when he reports compliance. During the monitoring phase, the agent is not fined but subject to periodic inspections. Crucially, the equilibrium naturally features penalty reductions for early disclosures of non-compliance, which is an aspect that is consistent with voluntary disclosure schemes that are commonly used in practice. The penalty reduction is needed to prevent the agent from delaying a report of an incidence of non-compliance in the hope that he can remain undetected and regain compliance in time for the next inspection.

We then contrast this equilibrium with stochastic inspection mechanisms and show that the ability to commit to random inspections decreases the expected inspection costs. Delay and noise in the detection of non-compliance, as well as additional penalties needed to generate incentives for voluntary disclosure, make deterministic inspections more costly than random inspection. We conclude by discussing alternative ways to use randomization without commitment power and highlight mechanisms that are used in practice: institutional separation of planning and execution of oversight and inspection sampling, combined with publicly accessible and verifiable inspection records.

1.1 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),

[2] An interpretation which is distinct from the different strands of reputation literature following Kreps and Wilson (1982), Mailath and Samuelson (2001), Holmström (1999), Board and Meyer-ter-Vehn (2013). See also Mailath and Samuelson (2006), pg. 459.

focus on static models of adverse selection. One of the main findings in this literature is the optimality of cut-off 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 of the static models. These include models with a risk-neutral agent and deterministic monitoring schemes (Webb, 1992; Chang, 1990), with a risk-neutral agent and random verification under wealth constraints (Monnet and Quintin, 2005; Antinolfi and Carli, 2015), and models with randomized verification and a risk-averse agent (Wang, 2005; Popov, 2016). These papers focus on discrete-time settings in which the agent's information is i.i.d. across periods. More similar to the setup in this paper is Ravikumar and Zhang (2012), which also considers a continuous-time model with persistent private information. However, they consider a pure adverse-selection framework in which the private information results from a single exogenous jump in income that is unobservable to the principal. In contrast, our analysis includes a moral-hazard problem and our information structure allows for oscillations between states and effort-dependent transitions.

The above papers focus on optimal mechanisms when the principal has full commitment power. Less attention has been given to the question of limited commitment in costly state verification models. A notable exception is the paper by Krasa and Villamil (2000) which considers an extension of the standard CSV model in which the principal is free to choose whether to go to court to enforce a mechanism. Their results are similar in spirit to ours: randomization is optimal when the principal can commit ex-ante, but deterministic verification is optimal without commitment. However, as they study a static model, they rely on an external party that enforces the mechanism. In our dynamic model without an external party, enforcement is ensured by contingent continuation play, and the strategic considerations in deciding *when* to inspect are quite different.

Commitment plays an important role in costly inspections in two fundamental and intertwined ways: First, when truthful reports are expected from the agent, the principal has no incentive to pay the inspection cost to reveal the agent's private information. Indeed, for static games, Reinganum and Wilde (1985) show that full compliance is not achievable without commitment. Second, randomized inspections —which may be more effective in providing incentives— are more demanding on the principal's commitment power. With repeated interactions, in which the principal's choices affect the continuation play, there is scope to provide punishment for insufficient inspection. Indeed, Ben-Porath and Kahneman (2003) prove a folk-theorem, showing that full compliance can be obtained without commitment in the undiscounted limit through random inspections. In our game, full compliance is attainable even with discounting. However, a non-committed principal cannot lower inspection costs through randomization. These differences stem from the persistence of information and the observability of inspections and. In Ben-Porath and Kahneman (2003), instances of non-compliance can only be detected by a contemporaneous inspection and inspections are not observable to the agent, so some instances of non-compliance are required to identify and incentivize inspections for the principal. When these incentives are provided tightly, that is, such that the principal is indifferent between inspecting or not, and inspections are unobserved, the principal may randomize. In the equilibrium constructed in Ben-Porath and Kahneman (2003), the frequency of non-compliance vanishes as the discount factor approaches one.

This paper is also related to the theories of crime deterrence and enforcement through policing and punishment following the seminal paper by Becker (1968). See also Dye (1986), Bassetto and Phelan (2008), and Bond and Hagerty (2010). The primary focus of these papers is the enforcement of an agent's hidden action. This stands in contrast to the CSV literature, which primarily focuses on the problem of eliciting hidden information. Showing that the agent need not be inspected while reporting non-compliance, our dynamic model of enforcement with voluntary disclosure confirms the insights from Malik (1993); Kaplow and Shavell (1994); Pfaff and Sanchirico (2000); Innes (1999a,b, 2001) who show in static models that self-reporting can reduce monitoring costs. For the case of limited commitment, there is an extensive literature on so-called *inspection games*, which have been applied to various problems, among them pollution or arms control. The contributions to this literature, including static as well as dynamic models, are surveyed by Avenhaus et al. (2002). More recently, Dilmé and Garrett (2019) consider a dynamic model of deterrence in which the principal faces a transition cost to commence monitoring. A common feature of equilibria in these games is that, without commitment, it is impossible to induce the agent to comply fully, which stands in contrast to our results with persistent information.

Our model of compliance as a two-state Markov process is based on Board and Meyer-ter-Vehn (2013), which studies the reputation problem of a firm that has to make costly investments towards quality improvements, where quality becomes publicly observable at random times. We embed this framework into a principal-agent problem, allowing the principal to choose the times at which to inspect the state of compliance and to control the agent's payoff by fines. A similar model has been studied by Kim (2015) in the context of environmental control, comparing specific classes of inspection policies without limited liability. Most closely related is the contribution by Varas et al. (2020), which also studies inspections in a principal-agent model that incorporates the reputation-for-quality framework. In their model, the agent is motivated by the desire to generate a positive reputation and inspections make the agent's current type public. Additionally, inspections serve an information-acquisition role for the principal. The authors find random inspections to be most effective for incentive provision, but deterministic inspections may nevertheless be optimal when inspections have the purpose of revealing socially valuable information. In contrast, in our model, the agent voluntarily discloses the state of compliance, and therefore inspections do not serve the purpose of diminishing public uncertainty. Instead, the reason for the optimality of non-random inspections in our paper is the principal's lack of commitment power.[3]

Finally, our model is also related to the classic machine maintenance problem in operations research. The machine maintenance problem is a statistical decision problem in which a machine "fails" at random times which can be observed only through inspection (see Osaki, 2002, for an overview). Similar models have also been applied in the accounting literature to study the optimal timing of audits (Kaplan, 1969; Carey and Guest, 2000; Hughes, 1977). All of these models are non-strategic, however.

2 Model

There are an agent and a principal. Time $t \in [0, \infty)$ is continuous. The principal and the agent are risk-neutral and discount future payoffs at a common rate $r > 0$. The principal requires the agent to comply with an exogenously given "standard" (representing a set of rules, for example on quality or conduct). Compliance is represented by a binary indicator variable $\theta_t \in \{0, 1\}$ which fluctuates over time with random transitions following a two-state Markov process. At each instant, the agent chooses effort level $\eta_t \in [0, 1]$ at instantaneous cost of $c\eta_t \, dt$ with $c > 0$. Effort affects transition rates of the state of compliance as follows.

[3] Varas et al. (2020) explain the occurrence of periodic, non-random inspections of aircraft as mandated by the FAA to the importance of the information revealed in safety checks. Our findings suggest that a complementary explanation could be that the principal must maintain a reputation for vigilance due to a lack of commitment.

State dynamics. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let the marked point process $z = \{z_t\}_{t \geq 0}$ represent the arrival of random shocks, where $z_t = 0$ except at isolated times $t_0 < t_1 < \cdots$. Shocks arrive at constant rate $\lambda > 0$. At each random time t_j, the value of the shock z_{t_j} is independently and uniformly distributed on $[0, 1]$. Let $\{\mathcal{F}_t\}$ be the natural filtration generated by z. The process $\theta = \{\theta_t\}_{t \geq 0}$ evolves according to a two-state Markov process that depends on the shock process and the agent's effort choices. Immediately after the arrival of a shock at time t_j, we have $\theta_{t_j} = 1$ if $\alpha\eta_{t_j} \geq z_{t_j}$, and $\theta_{t_j} = 0$ if $\alpha\eta_{t_j} < z_{t_j}$. The state θ_t is constant between shocks. The control parameter $\alpha \in (0, 1)$ measures the agent's control of the state conditional on the arrival of a shock; $\alpha < 1$ represents the possibility that the agent cannot always maintain compliance despite best efforts. At any time t, the probability of $\theta_{t+\mathrm{d}t} = 1$ is therefore

$$\mathrm{Prob}(\theta_{t+\mathrm{d}t} = 1 | \theta_t) = \begin{cases} 1 - \lambda\,\mathrm{d}t + \alpha\eta_t\lambda\,\mathrm{d}t & \text{if } \theta_t = 1, \\ \alpha\eta_t\lambda\,\mathrm{d}t & \text{if } \theta_t = 0. \end{cases}$$

Inspections and fines. The principal has an interest in the agent's compliance. At each instant t, the principal's flow payoff is $\theta_t R$, where $R > 0$. The agent privately observes θ_t at each time $t \geq 0$. The principal cannot observe the agent's effort, and compliance is observable for her only through costly inspections.[4] At any time $t \geq 0$, the principal can inspect the state at lump-sum cost $\kappa > 0$. In addition to inspection decisions, the principal can punish the agent through fines. Fines may be understood literally, as compulsory monetary payments, or they may be interpreted as remedial actions that negatively impact the agent in some other way. We assume that the principal does not benefit directly from fining the agent. This assumption is innocuous if fines are interpreted as remedial actions. When interpreting fines as monetary payments, this assumption prevents rent-seeking incentives for the principal, who could use fines as a means to transfer surplus. In the context of public institutions, this represents a benevolent view of the regulator that uses transfers with the intention to correct market failures.

In addition to effort and reporting decisions, we allow both the principal and the agent to *exit*, which permanently ends the relationship and results in a continuation value of zero for the principal, and a continuation payoff of $-B$ for the agent. For the principal, this implies a constraint on the severity of fines she can impose. We assume that the exogenously given bound B is larger than $\frac{c(r+\lambda)}{\alpha\lambda r}$. Otherwise, the maximal punishment is insufficient to incentivize effort even if θ_t were public at all times. The players' option to exit reflects the idea that they can limit their liability by dissolving the relationship: a lender can withdraw a loan, a firm or bank can shut down, etc.

Timing. The timing at each $t \geq 0$ is as follows.[5] First, the agent chooses effort level η_t. Subsequently, nature determines whether a shock arrives and, conditional on the arrival of a shock and the effort level, draws a new state. The agent then observes the realized state and sends a report $\hat{\theta}_t \in \{0, 1\}$ to the principal. The principal, in turn, makes an inspection decision and, conditional on the inspection outcome, chooses a fine incurred immediately by the agent. Denote by N_t^I the number of inspections and by F_t the cumulative fines up to and including time t.

[4] Thus, to ensure that the principal cannot infer the state without inspection, we assume that she does not observe her flow payoff. Alternatively, one could model the principal's payoff difference as stemming from the possibility of a detrimental event that arrives only during the low state, ends the game, and creates expected cost of R.

[5] We outline the sequentiality at a given instant to establish some intuition about the order of events over time. Formally, this order is captured by continuity properties of the respective action and state paths.

Histories and strategies. A history at time t is a collection of paths

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

where

$$(\eta_s, \theta_s, \hat{\theta}_s, N_s^I, F_s) \in [0,1] \times \{0,1\} \times \{0,1\} \times \mathbb{N}_0 \times \mathbb{R}_+.$$

Throughout, we denote strict histories for which the realization at time t is excluded by h_{t-}. Let H_t be the set of all time-t histories and H_{t-} the set of all strict histories. Let $H = \bigcup_{t\geq 0} H_t$ and $H_- = \bigcup_{t\geq 0} H_{t-}$.

The agent's strategy specifies efforts and reports as functions of histories. A strategy for the agent is defined as a pair $(e, \rho) = (\{e_t, \rho_t\}_{t\geq 0})$ with

$$e_t : H_{t-} \to [0,1], \quad \rho_t : H_{t-} \times \{0,1\} \to \{0,1\},$$

where $e_t(h_{t-})$ is the agent's effort at time t and $\rho_t(h_{t-}, \theta_t)$ is the agent's report at time t after history h_{t-} when the state at time t is θ_t. To capture the principal's uncertainty about the agent's effort choices and the true state of compliance, consider a partition $\mathcal{H}_t^P$ of the history set H_t which comprises all subsets of H_t whose elements are indistinguishable to the principal. Define the partition $\mathcal{H}_{t-}^P$ similarly for strict histories at t. To allow for randomized inspections, we equip the principal with a (private) random signal π, defined on a sufficiently rich probability space with state space Π. A strategy for the principal is defined as a pair $(n, f) = (\{n_t, f_t\}_{t\geq 0})$ of mappings

$$n_t : \Pi \times H_{t-} \times \{0,1\} \to \{0,1\}, \quad f_t : H_{t-} \times \{0,1\}^3 \to \mathbb{R}_+,$$

which are constant on every $H_{t-}^P \in \mathcal{H}_{t-}^P$ for each $t \geq 0$, where f_t is required to be weakly increasing over time. Here, $n_t(\pi, h_{t-}, \hat{\theta}_t)$ is equal to 1 if an inspection is performed at time t and equal to 0 otherwise. By $f_t(h_{t-}, \theta_t, \hat{\theta}_t, \mathrm{d}N_t^I)$ we denote the cumulative fine imposed by the principal at time t. We abuse notation slightly and write $f_t(h_t)$ instead of $f_t(h_{t-}, \theta_t, \hat{\theta}_t, \mathrm{d}N_t^I)$ whenever there is no danger of confusion. The exit decision for each player at any history is a binary variable indicating whether this player decides to exit or not. For the ease of exposition, we do not introduce additional notation for these choices; they translate into lower bounds on the expected payoffs of the players in the equilibrium definition below. The strategies above are to be understood as conditional on no player having exited previously. Actions to be chosen after one player exited are irrelevant.

Equilibrium. In continuous-time games with observable actions, strategies may not produce well-defined action paths, and —in stochastic environments— agents' behavior may be non-measurable. We adopt the approach by Kamada and Rao (2018) and impose restrictions on strategies to ensure well-defined action paths. We refer the interested reader to Appendix A. We do not impose restrictions on strategies that rule out non-measurable behavior. Instead, our equilibrium definition below requires that strategies lead to measurable actions along the equilibrium path. Histories away from the equilibrium path may lead to non-measurability. 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. Therefore, potential non-measurabilities 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, 2018). 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 payoff for the principal at time t is

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

Similarly, the discounted net present payoff for the agent at time t is given by

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

Given a strategy profile, the principal and the agent form expectations about h based on their past observations. For strategies that induce measurable action processes on path, we denote the expected payoff for the agent and the principal at t by $U_t = \mathbb{E}_{t-}^A[u_t]$ and $V_t = \mathbb{E}_{t-}^P[v_t]$ respectively, where the expectation is with respect to shock process $\{z_s\}_{s\in[0,\infty)}$ and the randomization device π, and it is conditional on the information that is available to the agent and the principal, respectively.[6]

We define a combination of strategies $((e,\rho),(n,f))$, together with processes $\{V_t, U_t\}_{t\geq 0}$, to be a *perfect Bayesian equilibrium* if the following holds.

1. At no time $t \geq 0$ is there a deviation for the principal that yields a payoff strictly higher than V_t.

2. At no time $t \geq 0$ is there a deviation for the agent that yields a payoff strictly higher than U_t.

3. At all histories at any time t, we have $V_t \geq 0$, and $U_t \geq -B$.

4. Along the equilibrium path, V_t and U_t are equal to the conditional expectations given above. Away from the equilibrium path, V_t and U_t are equal to the conditional expectations whenever these are well-defined.

We say that the agent's strategy is *truthful* if $\rho(h_{t-},\theta_t) = \theta_t$ at all histories h_{t-} along the equilibrium path. Further, we call the agent's strategy *maximally compliant* if $e_t(h_{t-}) = 1$ after any history along the equilibrium path. Note that with maximum effort by the agent, the probability of attaining or remaining in 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.

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 or not at any history.[7] Henceforth, we refer to inspections as random whenever they are non-predictable for the agent.

3 Principal-optimal equilibrium

Our main result is the characterization of a truthful and maximally compliant principal-optimal equilibrium. We assume at this point that the principal's gain from compliance R is large relative to the inspection cost κ so that the benefit in a maximally compliant equilibrium outweighs the necessary

[6]For the agent, the expectation is with respect to the natural filtration generated by the process $\{\eta_s, \theta_s, \hat{\theta}_s, N_s^I, F_s\}_{s\in[0,t)}$ for his effort choice, and with respect to the natural filtration generated by $\{\hat{\theta}_s, N_s^I, F_s\}_{s\in[0,t)} \cup \{\eta_s, \theta_s\}_{s\in[0,t]}$ for his report. For the principal, the expectation is with respect to the natural filtration generated by the process $\{N_s^I, F_s, \theta_\tau : \mathrm{d}N_\tau = 1\}_{s\in[0,t)} \cup \{\hat{\theta}_s, \theta_\tau : \mathrm{d}N_\tau = 1\}_{s\in[0,t]}$ for her inspection decision, and with respect to the natural filtration generated by $\{\hat{\theta}_s, N_s^I, F_s\}_{s\in[0,t)} \cup \{\eta_s, \theta_s\}_{s\in[0,t]}$ for the cumulative fine.

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

inspection costs. An explicit lower bound on R follows from the equilibrium presented in Theorem 1 below.

The equilibrium path 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 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, that is, when $\hat{\theta}_t = 1$, τ_t increases linearly with slope 1, and τ_t drops to 0 whenever $\hat{\theta}_t = 0$ or $\mathrm{d}N_t^I = 1$. At each time t with $\tau_t = T^*$, an inspection is performed. Second, while the agent reports non-compliance, he pays a constant flow fine and is not inspected. Additionally, when reporting a change from compliance to non-compliance, the agent has to pay a lump-sum fine. While in compliance, the agent's expected payoff at the beginning of each inspection cycle is u^{1*}. In phases of non-compliance, the agent's expected payoff is constant at lower level $u^{0*} = u^{1*} - \frac{c}{\lambda\alpha}$, to incentivize effort. The details of the equilibrium are presented in the following theorem.

Theorem 1. *For R sufficiently large,*[8] *there is a principal-optimal equilibrium with inspection cycle length T^* and initial expected payoff pair (u^{0*}, u^{1*}) such that*

- *If $\hat{\theta}_t = 1$, the agent pays no fine and an inspection is performed whenever the clock τ_t reaches T^*.*
- *If $\hat{\theta}_t = 0$, the agent pays a constant flow fine $f^* = -ru^{0*}$.*
- *If $\hat{\theta}_t = 0$ and $\hat{\theta}_{t-} = 1$, the agent pays lump-sum transition fine $P(\tau_t) = u^{0*} - U^0_{\tau_t}$.*
- *Whenever an inspection reveals non-compliance, the agent has to pay B immediately.*

The values of T^*, u^{1*}, u^{0*}, and U^0_τ are

$$T^* = \sup\left\{T > 0 \,\middle|\, 0 = (B - c/r)\left(1 - e^{-rT}\right)\lambda\alpha - ce^{\lambda T}(e^{rT} - \alpha) + c(1-\alpha)\right\}, \tag{3}$$

$$u^{1*} = -B + e^{(r+\lambda)T^*}\frac{c}{\lambda\alpha} < 0, \quad u^{0*} = u^{1*} - \frac{c}{\lambda\alpha}, \tag{4}$$

$$U^0_\tau = \ e^{r\tau}u^{0*} - e^{r\tau}\left(e^{\lambda\tau} - 1\right)\frac{c}{\lambda} + (e^{r\tau} - 1)\frac{c}{r}. \tag{5}$$

The proof is relegated to the appendix. We defer a discussion of intermediate results and illustrations toward the equilibrium construction to Subsections 3.2–3.4 and proceed here with a description of equilibrium properties and comparative statics.

Figure 1 illustrates how the equilibrium unfolds 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^* (the first vertical line). During compliance, the agent's expected total payoff evolves according to $U_t = U^1_{\tau_t}$, which is equal to u^{1*} initially and at the inspection time. At the second vertical line, a breach in compliance occurs. In a first step, the agent's utility drops to the current level of $U^0_{\tau_t}$, the dashed blue line; at the same time, the agent pays the transition fine $P(\tau_t)$, so that his continuation utility increases by that amount to u^{0*}. While in non-compliance, the agent pays a flow fine so that the promised payoff remains constant at u^{0*} until the next transition to compliance (the third vertical line). At this transition, the agent's utility jumps up to u^{1*} and the evolution takes the same course as at $t = 0$ and $t = T^*$. The agent's utility during times of

[8] Sufficiently large R ensures that the principal's expected benefit from compliance, $\frac{r+\lambda\alpha}{r(r+\lambda)}R$, exceeds the required inspection costs, which are given in Equation (6) below. A sufficient condition in terms of primitives is $R > \kappa\frac{r+\lambda}{\log((B\alpha/c-1/r)\lambda)}$.

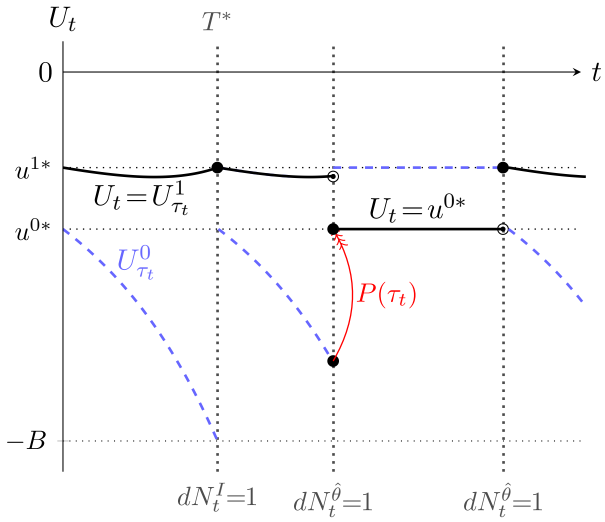

Figure 1: The evolution of an example path realization starting in the compliant state. Solid curves depict the agent's expected payoff in the current state, dashed curves depict the transitional payoff, to which the agent's payoff jumps when the state changes.

compliance, $U^1_{\tau_t}$, is lowest between two inspections. At the beginning, when the next inspection is still far in the future, the agent's incentives to delay the report of a breach in compliance is strongest so that the transition fine imposed must be moderate in comparison to the fine imposed for later reports of a transition. As time in compliance progresses, the transition fine becomes more severe, which decreases his expected payoff. Approaching the next inspection time, however, it becomes increasingly likely that no breach in compliance occurs previously, so that the persistent payoff increases again, reaching u^{1*} at the time of inspection.

Theorem 1 demonstrates that despite the principal's lack of commitment, maximal compliance and voluntary disclosure are attainable in equilibrium. This result stands in contrast to previous studies on enforcement without commitment (see, e.g., Reinganum and Wilde, 1985; Ben-Porath and Kahneman, 2003). The existence of such an equilibrium can be traced back to two sources. First, the long-term interaction between the principal and the agent generates enough value for the principal to make it worthwhile for her to bear the cost of consistently monitoring the agent. Second, the persistence of the agent's state of compliance supplies an important informational link that enables the principal to provide effective incentives for compliance based on observations at isolated inspection times alone. Indeed, our comparative statics results (see Proposition 2) show that without persistence, it is impossible to sustain a maximally compliant equilibrium at finite inspection costs.

A crucial feature of this equilibrium is that inspections are predictable from the perspective of the agent. The advantage of non-random inspections is that each inspection provides a verifiable signal to the agent which provides evidence of continued oversight. This demonstrated vigilance thus serves to shape the agent's perception that there is a high probability of detection if he was to deviate. The importance of perceived risk of detection is empirically evidenced by Makkai and Braithwaite (1994) which finds that CEOs of small organizations have greater regulatory compliance in their organizations if they perceive a high probability of detection. 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 non-beneficial for the principal.[9]

[9]See Proposition 3 for the formal result and further discussion.

This observation has important implications for audit design in practice. Of particular relevance is the recent announcement of the Public Company Accounting Oversight Board (PCAOB) that it will be making increased use of randomized inspections.[10] The PCAOB was founded in the early 2000s to re-establish public trust in oversight of the financial sector after a series of fraudulent activities within the then self-regulated audit profession, most notably the Enron scandal. Considering the context and motivation of its creation, the shift towards increased use of random audits, combined with a weakening of reporting standards,[11] may have the potential to undermine its ability to maintain effective oversight in the long run.

The equilibrium of Theorem 1 naturally features penalty reductions for early disclosures of non-compliance, which is consistent with voluntary disclosure schemes that are commonly used in practice. The U.S. environmental protection agency (EPA) for example employs a self-reporting program called "Incentives for Self-Policing," that requires participating firms to maintain an internal monitoring system and to voluntarily disclose any violations that are detected in this way. Similar to the way the agent is incentivized in the above equilibrium, firms who disclose violations early are rewarded by a reduction in penalties and a suspension of inspections until compliance is restored. Similar self-policing mechanisms are also employed by the U.S. Department of Defense for fraudulent activity among contractors and within the EU to ensure adherence to its data protection regulation (EU, 2016) and export regulation for dual-use goods (European Commission, 2019).

In our model, a penalty reduction is needed to prevent the agent from delaying a report of an incidence of non-compliance in the hope that he can remain undetected and regain compliance in time for the next inspection. To induce truthful disclosure, the penalty reduction must be decreasing over time such that, at each time between inspections, the agent's expected loss from paying an increased penalty when reporting non-compliance with delay is larger than the potential gain from avoiding the fine in case compliance recovers before the next inspection. The basic principle that the agent is monitored only in a more favorable state to him is familiar from static models of costly state-verification (Townsend, 1979). In our model, due to the coexistence of hidden information with hidden action, flow fines during non-compliance are necessary to create the incentive for effort, making the state of compliance more favorable to him.

Theorem 1 also provides insights into how exactly 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 her overall monitoring costs. On its website, the EPA points out that the advantage of these incentives lies in "making formal EPA investigations and enforcement actions unnecessary."[12] In the theoretical literature, the observation that voluntary disclosure reduces monitoring costs dates back to Kaplow and Shavell (1994), which introduces 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 non-compliance.[13] Our results add a theoretical argument to the empirical findings of Toffel and Short (2011) that monitoring

[10] Maurer,M., December 7, 2020, US watchdog will be selecting audits for inspection more randomly, *Wall Street Journal* https://www.wsj.com/articles/u-s-watchdog-will-be-selecting-audits-for-inspection-more-randomly-11607387903

[11] https://pcaobus.org/news-events/speeches/speech-detail/statement-regarding-the-pcaob-s-revised-research-and-standard-setting-agendas-reducing-credibility-accountability-and-confidence-in-the-financial-reporting-process

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

[13] See Varas et al. (2020) for a model without reports. Our results confirm the conjecture in that paper that voluntary disclosure can avoid unnecessary inspections (see Varas et al., 2020, p. 2921). Further, we discuss in Section 4 that the trade-off between random and deterministic inspections identified in that paper is resolved in favor of randomness when self-reporting and fines are possible.

has to be an essential component for the functioning of voluntary-disclosure schemes. Our results also emphasize the importance to align these schemes with the agent's dynamic incentives both for effort and disclosure.

The remainder of the paper is organized as follows. The next subsection provides comparative static results and additional discussion of equilibrium properties. The construction of the equilibrium presented in Theorem 1 is outlined in Subsections 3.2-3.4 and proceeds in three steps. First, we consider the agent's problem for arbitrary principal strategies and derive necessary and sufficient conditions for truthful reporting and maximal compliance to constitute a best response. These conditions are fully characterized in terms of the evolution of the agent's promised utilities conditional on each state. Second, we turn to the principal's problem and show that sequential rationality implies that it is without loss to use inspection schedules that are predictable. Furthermore, we show that if a predictable inspection strategy by the principal induces the agent to be truthful and maximally compliant and generates a positive payoff for the principal at all times, then this inspection strategy can be sustained in equilibrium. Thus, the principal-optimal equilibrium can be established by finding the cost-minimizing predictable strategy for the principal subject to inducing truthful reporting and maximal compliance. In the third step, we solve this mechanism design problem recursively, using dynamic programming techniques for piecewise deterministic processes (Davis, 1993). Equivalence is shown in Subsections 3.2 and 3.3. The mechanism design problem is set up and solved in Subsection 3.4. Readers who prefer to skip the technical details may jump to Section 4.

3.1 Comparative Statics

We now provide comparative statics of the principal-optimal equilibrium in Theorem 1. Specifically, we are interested in how variations in the parameters affect the length of inspection cycles and the overall inspection costs. Note that due to voluntary disclosure of non-compliance, the principal inspects only while the agent reports compliance, and thus shorter inspection cycles do not necessarily translate to higher monitoring costs.

We consider variations in four parameters: the penalty threshold B; the agent's cost of effort c; the parameter α, which measures the agent's level of control over compliance conditional on a shock; and the arrival rate of shocks λ, which measures the variability of compliance. In general, intuition would suggest that whenever a change in parameters relaxes the obedience constraint for the agent, so that it becomes easier to incentivize effort, then this should reduce the required monitoring intensity (i.e., increase the length of inspection cycles) and costs of monitoring. The next result shows that with regard to monitoring intensity, this is only partially true.

Proposition 1. *Consider a given set of parameters for which there exists an open neighborhood on which the equilibrium in Theorem 1 exists.*[14] *Holding all other parameters fixed, the length of the inspection cycle* T^* *is*

- *increasing in the penalty threshold* B*,*
- *decreasing in the effort cost* c*,*
- *increasing in control parameter* α*, and*

[14] That is, $B > \frac{c(r+\lambda)}{r\lambda\alpha}$ and R large enough.

- *increasing in the arrival rate for low λ and decreasing for high λ, with* $\lim_{\lambda \downarrow \frac{cr}{Br\alpha - c}} T^*(\lambda) = \lim_{\lambda \uparrow \infty} T^*(\lambda) = 0$.

It is straightforward to see that increasing the punishment bound B reduces the need for monitoring. A larger maximum fine for the agent provides a stronger deterrent, and thus decreases the required frequency of inspections. The basic mechanism is the same as in Becker's classical model of enforcement (Becker, 1968). The monitoring intensity also decreases when the cost of effort c decreases, or when the agent's control over compliance α increases. Either change in parameters relaxes the obedience constraint as effort becomes less costly or more effective, and thus allows the principal to incentivize effort with fewer inspections.

In contrast to the other parameters, a change in the variability λ has a non-monotone effect on monitoring intensity. To understand why, note that a change in variability has two opposing effects. On one hand, when λ is low, so that the state of compliance is highly persistent, it is very unlikely at any given instance for the state to change. This implies a low marginal benefit from effort, and therefore the principal must offer high-powered incentives to induce the agent to work. However, incentives for effort can only be generated by threats of penalty at times of inspections. Since the size of the penalty is bounded by the agent's limited liability, the only way to increase the power of incentives is for the principal to shorten inspection cycles. When λ grows large, then the state of compliance becomes extremely fragile, with a high frequency of transitions, so that delayed inspections have less incentive power as the link between current effort and future compliance weakens.

In a banking context, where the parameter λ has a natural interpretation as a bank's risk exposure, our non-monotone comparative statics on λ are consistent with the empirical findings in Delis and Staikouras (2011), which observes an inverse-u-shaped relationship between bank risk and the frequency of financial audits. The paper views risk exposure of banks as a choice variable and estimates how it responds to audit frequency abstracting from strategic considerations in the regulator's choice of monitoring policy. Our results point to an alternative explanation for the non-monotonic relationship between risk and audit frequency with reversed causality. In our interpretation, the non-monotonic relationship is the result of the regulator's choice of inspection intensity based on its risk assessment of the bank.

Next, we study the effect of changes in the parameters on the equilibrium monitoring costs. Consider the expected discounted inspection costs in the case $\theta_0 = 1$[15] as a function of the inspection cycle T^*:

$$C^1_{EQ} = \mathbb{E}\left[\int_0^\infty e^{-rt}\kappa \, \mathrm{d}N^I_t \mid \theta_0 = 1\right] = \frac{r+\lambda\alpha}{r(r+\lambda)} \cdot (r+\lambda(1-\alpha)) \frac{e^{-(r+\lambda(1-\alpha))T^*}}{1-e^{-(r+\lambda(1-\alpha))T^*}} \cdot \kappa. \tag{6}$$

The first fraction captures the relative likelihood of the good state. Future inspection costs are effectively discounted at rate $r + \lambda(1-\alpha)$ to account for the possibility that the state may deteriorate prior to inspection. Intuitively, inspection costs decrease as T^* increases. When α or λ change, there is an additional effect on costs as these parameters influence the underlying stochastic process and thus the inspection costs caused by any fixed cycle length T.

Proposition 2. *Consider a given set of parameters for which there exists an open neighborhood on which the equilibrium in Theorem 1 exists. Holding all other parameters fixed, the following holds for the discounted total inspection costs C^1_{EQ}.*

- *C^1_{EQ} is decreasing in the penalty threshold B.*

[15] The case $\theta_0 = 0$ is analogous. We have $C^0_{EQ} = \frac{\lambda\alpha}{r+\lambda\alpha} C^1_{EQ}$. For a detailed derivation of the total cost, see the proof of Proposition 2 in the appendix.

- *C^1_{EQ} is increasing in the effort cost c.*
- *For the agent's control α, the exist $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$ and $\bar{\alpha} \in [\underline{\alpha}, 1]$ such that C^1_{EQ} decreases in α if $\alpha < \underline{\alpha}$ and increases if $\alpha > \bar{\alpha}$.*
- *For the arrival rate λ, there exists $\underline{\lambda} > \frac{cr}{Br\alpha - c}$ such that C^1_{EQ} decreases in λ if $\lambda < \underline{\lambda}$ and C^1_{EQ} goes to infinity for $\lambda \to \infty$.*

The fractions in the third and fourth items represent our standing feasibility assumption $B > \frac{c(r+\lambda)}{\alpha r \lambda}$. For B and c the results follow immediately from Proposition 1 as these parameters have no effect on the inspection costs given any fixed cycle length T. As in Becker (1968), enforcement is more effective and thus less costly if the maximal punishment increases.

The non-monotone effect of α on monitoring cost provides some insight into the benefit from voluntary-disclosure schemes. The non-monotonicity is related to the value of voluntary disclosure in reducing inspection costs: when α is high and the agent is compliant most of the time, voluntary disclosure has only little advantage. With low levels of α, however, the savings from avoiding inspections during periods of non-compliance are significant. The non-monotonicity arises, because α affects inspection costs in two opposing ways. On one hand, α increases the informativeness of inspections, which increases the length of inspection cycles and in this way lowers monitoring costs. On the other hand, the agent tends to remain in compliance longer as α increases, and thus must be inspected more often on average. The second effect dominates for low values of α and the first effect dominates for high values, so that the overall effect of α on inspection costs is non-monotone.

Similarly, the variability parameter λ affects inspection costs directly through transitions in compliance and indirectly through the length of inspection cycles T^*. The direct effect of an increase in λ is a reduction in inspection costs because any inspection cycle of fixed length is increasingly often interrupted by a breach of compliance, so that inspections are less likely to be carried out. For low variability λ, an increase results in a longer equilibrium inspection cycle, and thus the overall effect of an increase in λ on inspection costs must be negative. For high variability λ, a further increase causes the cycle length to decrease, so that the two effects go in opposite directions. The last item in the result shows that the cost-increasing effect dominates for large values of λ. Note that for fixed $T > 0$ the total cost C^1_{EQ} decreases to 0 as λ grows arbitrarily large. Proposition 2 shows that T^* approaches zero fast enough so that the inspection costs explode in the limit. This cost increase arises because inspections are scheduled periodically, so that the agent has a strictly positive time to aim for a recovery in compliance prior to the next inspection. This deviation is more attractive for high values of λ.

As λ becomes arbitrarily high, the probability of reaching any time in the strict future without a prior transition vanishes. For inspections to effectively deter misreports, cycle length T^* has to shrink to 0 fast enough. The intuition why this leads to arbitrarily large inspection costs despite the increasing variability of the compliance state is the following. For the agent who considers shirking for an instant, the effective discount rate until the next inspection is $r + \lambda$. For example, effort at $t = 0$ affects the current state with probability $\lambda\,\mathrm{d}t$. The state at $t = 0$, in turn, determines the state at the next inspection only with probability $e^{-\lambda T^*(\lambda)}$, that is, if no other transition occurs before T^*. This implies that, for inspections to be effective, $\lambda T^*(\lambda)$ cannot vanish too quickly so that $\lambda e^{-(r+\lambda)T^*(\lambda)}$ remain strictly positive. However, the effective discount rate on path —i.e., the relevant rate to evaluate the principal's cost— is only $r + \lambda(1 - \alpha)$. Since the limit of C^1_{EQ} as λ goes to ∞ is proportional to $\lim \lambda e^{-(r+(1-\alpha)\lambda)T^*(\lambda)}$, it must be infinite for $\alpha < 1$ given that $\lambda e^{-(r+\lambda)T^*(\lambda)}$ remains positive in the limit. As the comparative statics uncovered, the high cost of compliance in the case of high λ stems from the agent's opportunity to regain compliance with high probability whenever the next inspection is not imminent, and making inspections imminent (choosing T^* close enough to 0) explodes the costs. This suggests that randomization over inspection times may be valuable as it permits the principal to threaten an instant inspection at all times without having to perform inspections at all times. We confirm this in Section 4 where we show that random inspection schedules dominate predictable inspections if randomization is feasible and discuss possible sources of commitment power or incentives to inspect randomly.

3.2 The agent's problem: incentive compatibility

We now characterize the incentive-compatibility conditions for the agent as constraints on the evolution of his promised utilities. Fix an arbitrary principal strategy (n, f) and let U_t be the agent's expected discounted continuation payoff from t onward, assuming he exerts full effort and reports truthfully throughout. 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:

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

By construction, the process $\{W_t\}_{t \geq 0}$ is a martingale (Davis, 1993, p. 20). Note that there are three types of random 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 the sake of consistency, we also 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 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 (non-predictable) stochastic process. Alternatively, we can think of the compensator as a generalization of the cumulative hazard function, and consequently of $\mathrm{d}\nu^a / \mathrm{d}t$ as the hazard rate of transitions in N_t^a (if it exists). 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,

$$q_t(\eta_t) = \theta_{t-}\lambda(1 - \alpha\eta_t) + (1 - \theta_{t-})\lambda\alpha\eta_t. \tag{7}$$

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

Lemma 1. *There exist predictable processes Δ^θ, $\Delta^{\hat{\theta}}$, Δ^I such that the evolution of the agent's expected utility is given by*

$$\mathrm{d}U_t = rU_t \, \mathrm{d}t + \mathrm{d}F_t + c\eta_t \, \mathrm{d}t + \sum_{a \in \{\theta, \hat{\theta}, I\}} \Delta_t^a (\mathrm{d}N_t^a - \mathrm{d}\nu_t^a). \tag{8}$$

The formal proofs of this and all remaining results are relegated to the appendix. The processes Δ^θ, $\Delta^{\hat{\theta}}$ and Δ^I have an intuitive interpretation: Δ_t^θ represents the jump in utility that results from a change in compliance at time t. Similarly, $\Delta_t^{\hat{\theta}}$ is the jump in utility that results from a change in reported compliance, and Δ_t^I represents the jump in utility that results from an inspection at time t.

To characterize when truthful reports and maximal compliance is a best response for the agent, fix a principal-strategy and take a strict history at any time t. Now, define the agent's continuation utility from truthful reporting and maximum effort as

$$(9)\qquad \begin{aligned} U_t^0 &= \mathbb{E}_{t-}^A[U_t|\theta_t = 0], \\ U_t^1 &= \mathbb{E}_{t-}^A[U_t|\theta_t = 1], \end{aligned}$$

when h_{t-} is followed by the realization $\theta_t = 0$ or $\theta_t = 1$, respectively. 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 the history h_{t-} is such that $\theta_{t-} = 1$, and the *transitional payoff* in case $\theta_{t-} = 0$, and vice versa for U_t^0.

The next lemma provides a complete characterization of the agent's incentive-compatibility constraints in terms of three objects: the most recent state θ_{t-}, and the utilities U_t^0 and U_t^1. For $U_t^{\theta_{t-}}$ to represent the agent's expected utility consistently, it has to satisfy a promise-keeping constraint. Furthermore, the principal's strategy has to provide incentives for the agent to truthfully reveal the state of compliance (honesty constraint), incentives for exerting full effort (obedience constraint) and it must be sufficiently lenient to deter the agent from withdrawing (participation constraint). By standard arguments, it is optimal for the principal to enforce the most severe punishment after a verified misreport to induce the agent to reveal the state truthfully. Thus, the agent's payoff after a false report that was detected by an inspection is $-B$.

Lemma 2. *A principal's strategy that generates the process $\{U_t^1, U_t^0\}_{t\geq 0}$ of utilities induces maximal compliance and truthful reporting if and only if there exists a predictable process $\{\mathrm{d}\mu_t \geq 0\}_{t\geq 0}$ such that for $i = \theta_{t-}$ and $j = 1 - \theta_{t-}$ we have*

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

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

(O) $U_t^1 - U_t^0 \geq c/\lambda\alpha,$

(P) $U_t^0, U_t^1 \in [-B, 0],$

at all $t \geq 0$ with $\mathrm{d}N_t^a = 0$ for each $a = \theta, \hat{\theta}, I$.

Condition (Pk) is the promise-keeping constraint which is the expectation of Equation (8) in Lemma 1 conditional on no intervention at time t. Condition (O) is the obedience constraint that ensures that maximal compliance is a best response for the agent. This condition has an intuitive interpretation. The marginal cost of effort is c. The marginal benefit from effort is $\lambda\alpha(U_t^1 - U_t^0)$, where the factor $\lambda\alpha$ is the marginal rate of arrival of a shock that leads to compliance. The utility gain from the high state is $U_t^1 - U_t^0$. Thus, (O) states that for maximum effort to be optimal for the agent, the marginal benefit must exceed the marginal cost. Condition (H) is the honesty constraint that ensures that the agent reports truthfully. This constraint says that transitional utility cannot increase too quickly. We have $\mathrm{d}\mu_t > 0$ when the honesty constraint (H) is slack. As a positive value for $\mathrm{d}\mu_t$ implies a more rapidly declining transitional utility, we shall refer to $\mathrm{d}\mu_t$ as the principal's *threat* to the agent. The variable $\mathrm{d}\nu_t^I > 0$ can be interpreted as the rate of inspections. When an inspection reveals a misreport, the continuation value for the agent is $-B$. Had he reported truthfully instead, he would have received transitional utility U_t^j. Thus, $-B - U_t^j$ is the utility the agent loses if an inspection reveals a false report.

We now present a heuristic derivation of the honesty constraint (H). For illustration, we focus on the case in which the state changes from high to low and there is no inspection. Suppose a decline in θ_t

occurs at time $t \geq 0$. The agent is willing to report the decline without delay only if he cannot gain from delaying the change in report. In particular, this means that on a small interval $[t, t + \mathrm{d}t)$, the value of reporting the low state must exceed the value of misreporting the high state, i.e.,

$$U_t^0 \geq \int_t^{t+\mathrm{d}t} e^{-(r+\alpha\lambda)(s-t)}(\alpha\lambda U_s^1 \, \mathrm{d}s - \mathrm{d}F_s - c\, \mathrm{d}s) + e^{-(r+\alpha\lambda)\mathrm{d}t} U_{t+\mathrm{d}t}^0.$$

The integral is the instantaneous gain from reporting high instead of low, followed by a truthful report of the low state at time $t + \mathrm{d}t$ if no change happened in the meantime. Taking a first-order approximation, we obtain, after a few rearrangements,

$$U_t^0 \geq \alpha\lambda U_t^1 \, \mathrm{d}t - \mathrm{d}F_t - c\, \mathrm{d}t + (1 - r\, \mathrm{d}t - \alpha\lambda\, \mathrm{d}t) U_{t+\mathrm{d}t}^0.$$

If we further substitute the approximation $\mathrm{d}U_t^0 := U_{t+\mathrm{d}t}^0 - U_t^0$ and ignore higher-order terms, then this necessary condition for a truthful report is equivalent to

$$\mathrm{d}U_t^0 \leq rU_t^0 \, \mathrm{d}t - \alpha\lambda(U_t^1 - U_t^0)\, \mathrm{d}t + \mathrm{d}F_t + c\, \mathrm{d}t.$$

This inequality is precisely condition (H) for the high state without inspection ($\mathrm{d}\nu_t^I = 0$). Note that while this heuristic derivation generates a necessary condition, the general result in Lemma 2 is also sufficient and captures the possibility of random arrivals of inspections.

3.3 The principal's problem: sequential rationality and predictability

It is sequentially rational for the principal to carry out inspections only if failing to do so results in some form of punishment. To provide such punishments for the principal, inspections must be at least partially predictable for the agent. In fact, the following result shows that, without commitment power, the principal cannot gain from *any* randomness in the timing of inspections.

Proposition 3. *For any truthful and maximally compliant equilibrium, there exists a principal-strategy such that truthful reporting and maximal compliance is a best response for the agent and*

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

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

The idea behind the proof is to take the principal's random equilibrium strategy and consider the realization in which inspections take place at the earliest possible time. The outcome generated by this strategy can then be replicated by a predictable inspection strategy. To see this, consider any maximally compliant equilibrium in which the principal randomizes over inspection dates. For a mixed strategy to be optimal for the principal, she must be indifferent along any path of play that is consistent with her mixed strategy. Call an inspection process *most vigilant* if, after any history, it generates the earliest possible inspection date that is consistent with the underlying mixed strategy. From the agent's perspective, there is zero probability that the principal performs inspections any earlier than given by the most vigilant inspection process. Moreover, the agent's incentive-compatibility conditions must be satisfied between inspections. Therefore, we can replace the principal's mixed strategy with a strategy in which inspections are predictable and determined by the most vigilant inspection process. We then show that we can modify the principal's strategy in this way without violating the incentive-compatibility conditions for the agent so that, by indifference, the new principal strategy generates the same payoff as the original equilibrium.

We now combine the previous proposition with the next result, which shows that the predictability of inspections is indeed the only restriction implied by sequential rationality.

Proposition 4. *Let $((n, f), (e, \rho))$ be a strategy profile and suppose the following holds.*

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

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

(iii) The expected payoff for the principal along any history generated by $((n, f), (e, \rho))$ is non-negative.

(iv) Every action path generated by the strategy profile $((n, f), (e, \rho))$ is measurable.

Then there exists a perfect Bayesian equilibrium $((n^*, f^*), (e^*, \rho^*))$ which generates the same distribution over action paths as $((n, f), (e, \rho))$.

This proposition says that for any strategy combination leading to well-defined action paths, for which the principal's inspections are predictable for the agent, the principal obtains a non-negative payoff and the agent's best response is truthful and maximally compliant, there exists a perfect Bayesian equilibrium that induces the same outcome. Intuitively, 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 we construct, the agent immediately stops working and exits if the principal deviates by not inspecting as expected. The agent's exit represents, in reduced form, a possibly much richer continuation play in which the agent believes that a principal has become non-vigilant, in the sense that she no longer inspects in a way that would allow her to satisfy the agent's obedience constraint. The agent would thus begin to shirk, and the principal would retaliate by setting large fines which eventually force the agent to exit. The agent's immediate exit is the worst case in terms of payoffs for both the principal and the agent and thus the equilibria supported in this way generate a higher payoff for the principal than other, more favorable off-path continuation plays.

Propositions 3 and 4 in combination imply that for any equilibrium that involves randomized inspections, we can find another equilibrium in which inspections are non-random and the expected payoff for the principal is the same. Therefore, to characterize principal-optimal equilibria, it is sufficient to find a strategy for the principal with non-random inspections that induces truthfulness and maximum effort and minimizes the principal's monitoring costs.

3.4 Derivation of the principal-optimal predictable strategy

Based on the results of the previous subsection, we can transform our equilibrium optimization problem into a mechanism design problem in which inspections must be predictable for the agent. A mechanism is characterized by a strategy for the principal. The mechanism is incentive compatible if the paths of promised utilities resulting from truthful reporting and maximal compliance satisfy the conditions in Lemma 2. Standard results then allow us to formulate the optimization problem in recursive form, with the promised utilities as state variables. We derive the optimal predictable strategy using a recursive approach due to Davis (1993). This method involves restricting the principal to perform a fixed number of inspections and then solving for the optimal strategy recursively as the number of inspections grows large. To do this, we first restrict attention to histories for which the state is always in compliance. We then show how the principal's strategy can be adjusted to optimally respond to reports of non-compliance. Here, we provide a heuristic derivation, the formal arguments are contained in the proof of Theorem 1.

Between inspections, the evolution of promised utilities during continued periods of compliance can be characterized by a pair of first-order differential equations. To see this, note that when there are no inspections, the principal's choice of $\mathrm{d}F_t$ cannot depend on the true state (conditional on the report $\hat{\theta}_t = 1$). Moreover, the principal and the agent are risk-neutral, and therefore it is without loss to shift all fines into the future until after the next inspection, that is, we can set $\mathrm{d}F_t = 0$ for all $t \geq 0$ strictly before the next inspection. Assume additionally that constraint (H) binds between inspections, i.e., $\mathrm{d}\mu_t = 0$ whenever $\mathrm{d}N^I_t = 0$. We later verify that the principal cannot lower inspection costs through threats between inspections. Hence, if we start at $t = 0$ with initial payoffs U^0_0 and U^1_0, the trajectories of the transitional payoff U^0_t and the persistent payoff U^1_t up until the first inspection are pinned down by the constraints (Pk) and (H) in Lemma 2. This pair of coupled first-order differential equations has the following closed-form solution:

$$U^0_t = e^{rt}(U^0_0 - \alpha(e^{\lambda t} - 1)(U^1_0 - U^0_0)) + c(e^{rt} - 1)/r, \tag{10}$$

$$U^1_t = e^{rt}(U^1_0 + (1 - \alpha)(e^{\lambda t} - 1)(U^1_0 - U^0_0)) + c(e^{rt} - 1)/r. \tag{11}$$

We now iterate over the number of inspections. As the promised utilities are determined by (10) and (11), the optimization problem reduces to the choice of initial values $(u^0, u^1) = (U^0_0, U^1_0)$ and inspection time T subject to the remaining constraints (O): $U^1_t - U^0_t \geq \frac{c}{\alpha\lambda}$ and (P): $U^i_t \in [-B, 0]$ for $i \in \{0, 1\}$ and all $t \in [0, T]$.

Note that maximal compliance is not achievable if the number of inspections is bounded. To ensure that a solution exists for problem step $k \in \mathbb{N}$, (when the number of inspections cannot exceed k), we set the principal's objective to ensure compliance for as long as possible, which is equivalent to maximizing the time between inspections. Consider first the case with no inspection ($k = 0$). At any time t, the principal has no possibility to distinguish the states, so that $U^0_t = U^1_t$ and effort can never be incentivized. As a consequence, if the principal has one inspection ($k = 1$), effort is achievable at most until the time of this inspection. The principal's goal is to perform this inspection as late as possible such that the payoff pair (U^0_t, U^1_t) fulfills conditions (O) and (P) up until this inspection time. The trajectories in (10) and (11) can be combined to obtain

$$U^1_t - U^0_t = (U^1_0 - U^0_0)e^{(r+\lambda)t},$$

which shows that the obedience constraint (O) is fulfilled for all $t \geq 0$ if it holds at $t = 0$.

Given any initial values (u^0, u^1) satisfying $u^1 - u^0 \geq \frac{c}{\alpha\lambda}$, the optimal inspection time is the largest value $T > 0$ such that U^i_t satisfy (P) for all $t \in [0, T]$. That is, the minimum of the two boundary hitting times

$$T^0(u^0, u^1) = \inf\{t > 0 : U^0_t \leq -B\} \qquad \text{and} \qquad T^1(u^0, u^1) = \inf\{t > 0 : U^1_t \geq 0\}.$$

The principal chooses (u^0, u^1) to maximize $\min\{T^0, T^1\}$. Note that, by (10), U^0_t is increasing in u^0 while by (11), U^1_t is decreasing in u^0, for all $t \geq 0$. Thus, an increase in u^0 increases both T^0 and T^1 so

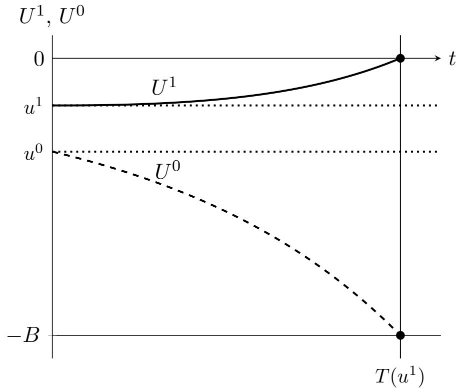

Figure 2: The evolution of promised utilities over time conditional on continued compliance with a single inspection. Persistent utility is shown as solid line, transitional utility is dashed.

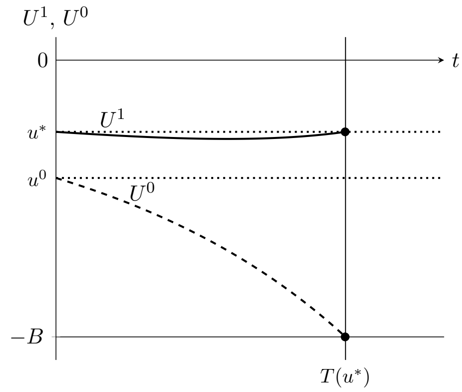

Figure 3: The evolution of promised utilities over time conditional on continued compliance with repeated inspections. Persistent utility is shown as solid line, transitional utility is dashed.

that it is optimal to set u^0 as large as possible, i.e., $u^0 = u^1 - \frac{c}{\alpha\lambda}$. The problem in step $k = 1$ is thus reduced to finding the optimal initial utility u^1. Again, from (10) and (11) we observe that U_t^0 and U_t^1 are both increasing in u^1 for fixed value of $u^1 - u^0 = \frac{c}{\alpha\lambda}$. Thus, T^0 is increasing and T^1 decreasing in u^1. The minimum of the two hitting times, $T(u^1) \equiv \min\left\{T^0(u^1 - \frac{c}{\alpha\lambda}, u^1), T^1(u^1 - \frac{c}{\alpha\lambda}, u^1)\right\}$, is then maximized by choosing u^1 so that each utility path hits the respective boundary at the same time. Figure 2 illustrates this. For any other choice of initial value u_1, the hitting time is lower. For higher values u^1, the upper boundary is reached earlier, for lower values, the lower boundary is reached earlier.

Next, consider the case $k = 2$, in which the principal can perform two inspections. If the first of the two inspection times was chosen in the same way as in the previous step with both utility levels at the boundary, then U_t^1 would stay constant at 0 after the first inspection. No further incentives could be created despite another inspection being left. Hence, the principal-optimal initial value after the first inspection is strictly lower than in the case where only one inspection is available, so that the trajectory of U_t^0 reaches the lower boundary $-B$ before the trajectory of U_t^1 reaches the upper boundary 0. A lower level of u^1 forces the principal to inspect earlier, but she retains the option to fine the agent in the future, which is necessary for the following inspections to be valuable.

For $k > 2$, we proceed in a similar fashion. Denote by $u^1(k)$ the optimal initial value of the trajectory of U_t^1 when the maximal number of inspections is k. By a similar argument as above, the inspection time T^k is determined by the time at which the trajectory of U_t^0 reaches $-B$. Iterating over the number of inspections k, we find that the optimal initial value $u^1(k)$ decreases as k increases. The more inspections are available to the principal, the more she will reduce the agents' persistent payoff in order to retain the option to fine him in the future. This implies that the inspection time T^k decreases. As k grows large, $u^1(k)$ converges to a unique limit u^{1*} and T^k converges to a unique limit T^*, the length of each inspection cycle. The optimal mechanism ensuring maximal compliance has the property that the trajectory of the persistent utility U_t^1 is u-shaped, and it returns to the initial value u^{1*} at the time of each inspection (see Figure 3). The limit values u^{1*} and T^* are given by the solution to (10) and (11) with boundaries

$(U_0^0, U_0^1) = (u^{0*}, u^{1*})$ and $(U_{T^*}^0, U_{T^*}^1) = (-B, u^{1*})$:

$$U_t^0 = e^{rt}u^{1*} - e^{(r+\lambda)t}\frac{c}{\lambda} - e^{rt}\frac{1-\alpha}{\alpha}\frac{c}{\lambda} + (e^{rt}-1)\frac{c}{r}, \tag{12}$$

$$U_t^1 = e^{rt}u^{1*} + e^{(r+\lambda)t}\frac{1-\alpha}{\alpha}\frac{c}{\lambda} - e^{rt}\frac{1-\alpha}{\alpha}\frac{c}{\lambda} + (e^{rt}-1)\frac{c}{r}. \tag{13}$$

We provide an implicit characterization of T^* and u^{1*} in Theorem 1 below. Finally, for the case of a breach of compliance ($\hat{\theta}_t = 0$) at time t, the promised utilities jump to $U_t^1 = u^{1*}$ and $U_t^0 = u^{1*} - \frac{c}{\lambda\alpha}$. Using constant flow fines, the promised utilities can be held constant at these levels while $\hat{\theta}_t = 0$. In this way, upon another transition, the promised utilities are already at their optimal initial values.

4 Overcoming the commitment problem

Our previous results show that without commitment the principal cannot benefit from inspecting at random. We now show that when the principal can commit, then stochastic inspection procedures reduce her monitoring cost. We discuss two possible ways for the principal to implement stochastic inspection strategies even without commitment. First, the principal can rely on the help of independent third parties. Second, when the principal is responsible for a large pool of agents, then randomization can be implemented without commitment by inspecting a sample of firms and making inspection reports publicly available.

4.1 Stochastic inspection mechanisms

We begin by considering a simple stationary stochastic mechanism: whenever the agent reports non-compliance, he pays a fixed flow fine

$$f^* = (r+\lambda)\frac{c}{\lambda\alpha}, \tag{14}$$

and is never inspected. On the other hand, whenever the agent reports compliance, he pays no fines, but the principal inspects at random at a stationary Poisson rate

$$m^* = \frac{cr(r+\lambda)}{Br\lambda\alpha - c(r+\lambda)}. \tag{15}$$

Finally, if the agent is inspected and found to have made a false report, he receives the maximum penalty. There are no other transfers. It is straightforward to verify that in this mechanism, when the agent exerts maximum effort and reports truthfully, then his promised utilities are stationary, and given by

$$\bar{U}^1 = -\frac{c}{r\alpha}\ , \quad \bar{U}^0 = -\frac{c}{r\alpha} - \frac{c}{\lambda\alpha}, \tag{16}$$

where $\bar{U}^1$ is the persistent utility when the agent reports compliance and $\bar{U}^0$ is the persistent utility when the agent reports non-compliance. It is immediate to see that for either $\theta \in \{0, 1\}$, when we substitute the promised utilities, m^*, and f^* into the honesty constraint (H) and obedience constraint (O), then each holds with equality. Thus, for this mechanism, the honesty and the obedience conditions bind at all times. The next result shows that this mechanism is optimal among all stationary stochastic mechanisms in the sense that the monitoring costs for the principal are lower than in any other mechanism generating a stationary promised utility in each state of compliance.

Theorem 2. *The following characterizes an optimal stationary stochastic mechanism:*

1. *If $\hat{\theta}_t = 1$, inspections are conducted at a constant rate m^* given by* (15)*. There is no fine after an inspection if it reveals the high state, and a maximum penalty $B + \bar{U}^0$ if it reveals the low state.*
2. *If $\hat{\theta}_t = 0$, the agent pays a constant flow fine f^* given by* (14) *and no inspections are performed.*
3. *The agent's promised utilities $\bar{U}^1$ and $\bar{U}^0$ are given by* (16)*.*

Similar to the non-commitment case, the agent is fined a fixed flow fine in phases of non-compliance, and inspections are performed only when he reports compliance.

Proof. To see that the mechanism described above is optimal, consider an alternative stationary stochastic mechanism that delivers some given promised utility u. From the promise-keeping constraint (Pk) and the honesty constraint (H), it is straightforward to obtain that the constant rate $m_0(u)$ of inspection that keeps the promised utility stationary at the level $u \in [-B + \frac{c}{\lambda\alpha}, -\frac{c}{r\alpha}]$ is

$$m_0(u) = \frac{r(c - \alpha\lambda u)}{\alpha\lambda(B + u) - c}.$$

The principal's expected monitoring costs in the stationary random mechanism that provides promised utility $U_t^1 = u$ throughout can be determined recursively as follows:

$$C_R^0(u) = \int_0^\infty e^{-(r+\lambda)t}\lambda\left(\alpha C_R^1(u) + (1-\alpha)C_R^0(u)\right)\mathrm{d}t \qquad \text{and}$$
$$C_R^1(u) = \int_0^\infty e^{-(r+\lambda)t}\left(m^0(u)\kappa + \lambda(\alpha C_R^1(u) + (1-\alpha)C_R^0(u)\right)\mathrm{d}t,$$

where $C_R^0(u)$ denotes the expected costs while in non-compliance and $C_R^1(u)$ the expected costs while in compliance. Solving the above equations for the cost during compliance gives

$$C_R^1(u) = \frac{r + \lambda\alpha}{r}\frac{m_0(u)}{r+\lambda} = \frac{r+\lambda\alpha}{r}\frac{r(c-\alpha\lambda u)}{(r+\lambda)(\alpha\lambda(B+u)-c)}. \tag{17}$$

It is easy to see that $C_R(u)$ 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 in Theorem 2, it follows that this mechanism is indeed the optimal stationary mechanism. □

The next result shows that the principal's monitoring cost with predictable inspections is generally higher than with random inspections:

Proposition 5. *The monitoring cost for the principal in the optimal stationary stochastic mechanism is lower than in the principal-optimal equilibrium.*

Proof. We can express the total cost of compliance in the principal-optimal equilibrium recursively as follows:

$$C_{EQ}^0 = \int_0^\infty e^{-(r+\lambda)t} \lambda \left(\alpha C_{EQ}^1 + (1-\alpha) C_{EQ}^0\right) \mathrm{d}t \qquad \text{and}$$

$$C_{EQ}^1 = \int_0^\infty e^{-(r+\lambda)t} \lambda \left(\alpha \tilde{C}^1(\tau_t) + (1-\alpha) C_{EQ}^0\right) \mathrm{d}t + \sum_{k=1}^\infty e^{-(r+\lambda)kT^*} \kappa,$$

where C_{EQ}^0 denotes the expected costs while in non-compliance and

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

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

$$C_{EQ}^1 \geq \underline{C} = \frac{r + \lambda\alpha}{r} \frac{e^{-(r+\lambda)T^*}}{1 - e^{-(r+\lambda)T^*}} \kappa. \tag{18}$$

To see that the costs from random inspection (17) are lower, use Equation (4) to write T^* in (18) as a function of u^{1*}, giving

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

Now it is immediate to check that $\underline{C}\left(-\frac{c}{r\alpha}\right) = C_R\left(-\frac{c}{r\alpha}\right)$ and $\underline{C}'(u) < C_R'(u)$ for $u < -\frac{c}{r\alpha}$ and $B > \frac{c(r+\lambda)}{r\lambda\alpha}$, where the latter is precisely our restriction on the bound B that ensures that the punishment is sufficient to induce effort. Since $-\frac{c}{r\alpha}$ is an upper bound on u, it follows that $C_R(u^{1*}) < \underline{C}(u^{1*})$. It follows from (18) that $C_{EQ}^1 > C_R^1$. □

Random inspections dominate predictable inspection procedures for two reasons. One reason is that due to noise and delay, a deterministic inspection regime is less effective in providing incentives for effort than a stochastic monitoring scheme, where the threat of punishment is instantaneous (Varas et al., 2020). Moreover, predictable inspections generate additional costs when the principal must provide incentives for voluntary disclosure. When inspections are predictable, the agent must be deterred from trying to hide a breach in compliance in the hope of recovering before the arrival of the next inspection. To generate incentives for truthtelling, the principal-optimal equilibrium of Theorem 1 requires a transition fine that must increase exponentially to prevent the agent from concealing such a breach in compliance. The risk of having to pay the transition fine, however, decreases the agent's overall payoff in equilibrium. The reduction in his promised utility has the side effect of decreasing the maximum loss that the principal can impose, which makes inspections overall less powerful. Thus, with predictable inspections, truthful reporting requires a higher monitoring intensity. More formally, it is easy to see that the initial promised utility with predictable inspection is smaller than the utility in the optimal stationary mechanism with random inspections, i.e., $u^{1*} < \bar{U}^1$. Moreover, as $C_{EQ}(\cdot)$ is decreasing in u, we have $C_R(\bar{U}^1) < C_{EQ}(\bar{U}^1) < C_{EQ}(u^{1*})$. The first inequality represents the higher cost of deterministic inspections for any payoff level, and the second inequality represents the cost increase that stems from the dynamic voluntary disclosure incentives.

Our finding that random inspections dominate deterministic inspections for incentive provision is consistent with Varas et al. (2020) for the case without voluntary disclosure. The authors show that (partially) predictable inspections can be optimal when the principal derives direct value from knowledge of the true state, i.e. when her flow-payoff is convex is her posterior belief. In our model, the principal uses fines to induce honest self-disclosure by the agent. Along the equilibrium path, the principal thus knows the true state of compliance. 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 et al. (2020) identify a trade-off according to which incentive provision recommends randomization while the social value of information makes predictable inspections more profitable. Our analysis shows that when the current state is known to the agent and monetary incentives are feasible, self-reporting can resolve this trade-off in favor of randomization.

4.2 Implementation of stochastic inspection mechanisms

While Proposition 3 shows that the principal without commitment power in our baseline model cannot benefit from random inspections, the result does not rule out randomization altogether. Proposition 3 merely says that the indifference between conducting the inspection and postponing it, which is required in any mixed strategy, precludes the principal from benefiting from randomization.

4.2.1 Extraneous randomization

In our main result in Theorem 1, we assumed implicitly that the principal and the agent do not have access to a public randomization device. Randomization devices are commonly encountered in models of relational contracts, and repeated games more generally, where they serve predominantly to ensure that equilibrium payoffs form a convex set. Here, introducing a randomization device would serve a slightly different purpose. By conditioning on the random realizations of a sufficiently rich, publicly observable stochastic process, the principal could inspect at random, while allowing the agent to detect deviations by the principal. Despite the theoretical appeal, however, implementing a random strategy with the help of a public randomization device is uncommon and it can be a rather involved process in practice.

Traditionally, theorists have proposed three different ways to obtain jointly observable random signals. One method is to rely on publicly observable random events. A prominent example of such natural random events that has frequently been cited in the literature is the occurrence of "sun spots", which are literal spots on the surface of the sun that randomly change their size and position. The literature uses sunspots predominantly as a metaphor, however, to illustrate how otherwise irrelevant random events might affect equilibrium outcomes if they influence expectations. Another way to obtain publicly observable random information is through a procedure known as "jointly controlled lottery" in which players have access to one or more randomization devices that they operate together. Examples include joint coin tosses or simultaneous message exchange (i.e., a cheap-talk game, such as rock-paper-scissors). In practice, joint lotteries are used for generating "keys" that are needed for example for cryptocurrencies or electronic travel documents.

For the setting we are envisioning in this paper, in which the relationship between the principal and the agent is at arm's length,[16] it seems implausible that the players employ elaborate stochastic methods based on naturally occurring random events or jointly controlled lotteries to coordinate the

[16]The modeling of the agent's reports as a continuous process is for notational ease. The agent's reports should be interpreted as occurring at times of transition only.

principal's inspection dates. Perhaps the most compelling way in which such stochastic information can be generated is through an impartial third party. One way to achieve this is by separating the inspection planning from its execution on an institutional level. For example, inspections can be prescribed by the compliance manager but carried out by external practitioners. This eliminates the principal's incentive to skip inspections. In practice, the separation of planning and execution can be observed in the context of banking supervision in Germany: Depending on the bank's size, the European Central Bank or the supervisory agency at the federal Finance Ministry (BaFin) fulfills the supervisory function and is responsible for scheduling audits. The execution of these audits, however, is always done by the German Bundesbank (BaFin, 2016). This way, the inspection cost is incurred by a different party from the one taking the inspection decision.

There is an important difference between this suggestion and two seemingly analogous alternatives: outsourcing the entire oversight activity at one extreme and directly compensating the principal for the inspection costs at the other. Delegating to an independent oversight agency solely shifts the question of how to avoid skipping costly inspections when that action can be accounted for with an alternative realization. Compensating the principal for each performed inspection resolves the problem highlighted above only if there is precise knowledge of the cost and effort required by the principal to carry out an inspection (cost κ in our model). If the compensation for an inspection falls below this value, the incentive to skip it remains. If the compensation is too generous, this creates an incentive to inspect inefficiently often.

4.2.2 Randomization through inspection sampling and public records

The lack of detectability which hinders profitable randomization may be mitigated if the principal is responsible for monitoring a large pool of independent agents. The principal can employ a strategy that involves inspecting a fixed proportion of inspectees at all times, and publicizing the outcomes to create a verifiable signal of vigilance and monitoring effort. Indeed, many public regulators are responsible for overseeing large pools of companies and their oversight activity can be verified through the publication of inspection outcomes. For example, the EPA's database "Enforcement and Compliance History Online"[17] collects over 44,000 inspected facilities within the 12 months up to April 2021; the PCAOB publicizes approximately 100-300 inspection reports per year.[18]

Inspection sampling can be formally incorporated into our model as follows. Suppose there is a pool of $M \in \mathbb{N}$ identical and independent agents. Let $M_0 \leq M$ be the number of these agents who are currently in compliance; this implies that M_0 is Binomially distributed with parameters (α, M). For simplicity, suppose that $\alpha = 1$, so that $M_0 = M$ (for the case $\alpha < 1$, the expected number of compliant firms is $M_0 = \alpha M$ and the construction below works in a similar way). Fix a time interval of positive length $\mathrm{d}t$. Suppose that on this time interval, theorems $k = M\,\mathrm{d}t$ separate inspections at isolated dates, and at each of these inspections, a total number of $m < M$ firms is inspected. For each firm, the probability of being among the m inspected agents at a given inspection date is given by $p = \frac{m}{M}$. On the entire time interval dt, the number of inspections for this agent is binomially distributed with parameters k and p, where k is the number of trials and p the probability in each trial. Now, when $M \to \infty$, then also $k \to \infty$ proportionally, and by the Poisson limit theorem, the distribution of the number of inspections for each firm converges to the Poisson distribution with parameter $kp = m\,\mathrm{d}t$. Hence, by definition, the rate of inspection for any agent follows a Poisson process with arrival rate m, while the number of inspections

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

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

for the principal at each time t is constant.

If inspections are publicly observable or their outcomes have to be published by the principal, it is easily detectable for any agent if the principal ceases her oversight activity by reducing inspections. At the same time, each agent's own inspection times are not predictable and follow the optimal constant rate derived in Theorem 2. This example indicates that random inspection protocols are easier to sustain if the principal's vigilance is evaluated on the basis of overseeing an industry that consists of a large pool of firms. By contrast, in monopolistic or oligopolistic industries in which there is a small number of large companies, such as Boeing in the U.S. aircraft manufacturing business as mentioned in the introduction, consistent oversight with random protocols will be much harder to sustain.

5 Conclusion

The paper studies a principal-agent setting with costly inspections in which the principal incentivizes the agent to achieve and maintain compliance and disclose any incidence of non-compliance. Under relational enforcement, when the principal cannot commit to the timing of inspections, it is possible to induce maximum compliance and voluntary disclosure and the principal achieves this optimally through non-random inspections. A fully committed principal, however, would prefer to inspect at random.

The persistent and delayed effect of the agent's effort on the state of compliance makes it possible to create incentives through isolated and predictable inspections and fines. We find that an intermediate level of persistence is optimal for the principal. If the state of compliance becomes arbitrarily persistent, the agent's effort is unlikely to have an effect. If the state of compliance grows arbitrarily variable, the informational link between past effort and compliance at future inspections erodes. At both extremes, it becomes infeasible to achieve maximal compliance with finite inspection costs.

The use of predictable inspection schedules makes deviations by the regulator easy to observe for the agent. Regular inspections serve as demonstrations of the principal's vigilance and they are needed to maintain the agent's perception that non-compliance will likely be detected. Random inspection schedules, on the other hand, lower the principal's accountability, as it is difficult or impossible for the agent to verify whether the principal remains vigilant when the inspections are conducted at random. Decreased accountability, paired with pressure to reduce the regulatory burden and costs of monitoring, can then result in a failure of oversight and enforcement.

A crucial assumption maintained throughout is that the principal seeks maximal compliance, defined as full effort and truthful reporting after any history along the equilibrium path. This assumption keeps the analysis tractable. We interpret the results in this paper as a benchmark for the design of optimal policies, and as a sound theoretical approach to generating predictions of the cost of effective enforcement. For a subset of the parameter space, in particular, if the principal's reward from compliance, R, is large enough, implementing effort always is optimal. There are cases, however, in which full compliance is not socially optimal. Extending the analysis to allow for periods of non-compliance by the agent would be interesting, but it makes the underlying optimization substantially more complicated. Other aspects not considered in this paper are, for example, exogenous signals about the state, an agent who is imperfectly informed, or an imperfect monitoring technology. These and other variants might be fruitful avenues of future research.

Appendix A: Strategies and outcomes

This part of the appendix contains the formal restrictions on the players' strategy spaces to ensure that any combination of strategies leads to a unique and well-defined outcome. We adopt the approach by Kamada and Rao (2018) and require that actions are not changed 'too frequently' on any time interval. To apply this approach we first restrict the strategy spaces for the fine and effort choices. A history $h_t \in H_t$, has an *intervention for the agent* at time t if either $t = 0$, or if $t > 0$ and at least one of the following holds: (i) $\theta_t - \theta_{t-} \neq 0$, (ii) $\hat{\theta}_t - \hat{\theta}_{t-} \neq 0$, (iii) $N_t^I - N_{t-}^I \neq 0$. Similarly, there is an *intervention for the principal* if either $t = 0$, or if $t > 0$ and at least one of the properties (ii) and (iii) holds. No new information arrives in between interventions. We restrict the principal's fine strategy to reflect this, and require that it be predictable in between inspections. Formally, for any two histories $h_t and h_t'$: $f_t(h_t) \neq f_t(h_t')$ only if there exists $\tau \leq t$ such that τ is an intervention time for the principal and the truncation of the above histories at time τ, h_τ and h_τ', are distinguishable for the principal. In other words, this restriction requires the principal's fines to be specified pathwise; at each intervention, it is fully specified how fines proceed until another intervention arrives. Similarly, we restrict the agent's effort strategy to be predictable in between interventions: For any two histories h_{t-}, h_{t-}': $e_t(h_{t-}) \neq e_t(h_{t-}')$ only if there exists $\tau < t$ such that τ is an intervention time for the agent and $h_\tau \neq h_\tau'$. Based on Kamada and Rao (2018), we require all strategies to fulfil the properties *traceability* and *frictionality* as defined below. Lemma A then shows that any combination of strategies from this class yields a well-defined and unique outcome path. A history h is said to be *consistent* with the agent's strategy (e, ρ) at time t if $\rho_t(h_{t-}, \theta_t) = \hat{\theta}_t$ and $e_t(h_t) = \eta_t$. Similarly, a history h is consistent with the principal's strategy (n, f) at time t if $n_t(\pi, h_{t-}, \hat{\theta}_t) = \mathrm{d}N_t^I$ and $f_t(h_t) = \mathrm{d}F_t$.

Definition 1. *The agent's strategy (ρ, e) is* ***traceable*** *if for any time-t history h_t and any principal-action path $\{N_s^I, F_s\}_{s\geq 0}$ that coincides with h_t for all $s < t$, there is a continuation path $\{\hat{\theta}_s, \eta_s\}_{s\geq t}$ that is consistent with (ρ, e). Analogously, The principal's strategy (n, f) is traceable if for any time-t history h_t and any agent-action path $\{\hat{\theta}_s, \eta_s\}_{s\geq 0}$ that coincides with h_t for all $s < t$, there is a continuation path $\{N_s^I, F_s\}_{s\geq t}$ that is consistent with (n, f).*

Definition 2. *The agent's strategy (ρ, e) is* ***frictional*** *if for any time-t history h_t, there is conditional probability one that the report path $\{\hat{\theta}_s\}_{s\geq t}$ has only finitely many report changes on any finite interval $[t, u]$ for all paths $\{\eta_s, \hat{\theta}_s\}_{s\geq t}$ such that there is a principal-action path $\{N_s^I, F_s\}_{s\geq t}$ for which the history $\Big(h_{t-}, \{N_s^I, F_s\}_{s\geq t}, \{\eta_s, \hat{\theta}_s\}_{s\geq t}\Big)$ is consistent with the agent's strategy. Analogously, the principal's strategy (n, f) is frictional if for any time-t history h_t, there is conditional probability one that the inspection path $\{N_s\}_{s\geq t}$ has only finitely many inspections on any finite interval $[t, u]$ for all paths $\{N_s^I, F_s\}_{s\geq t}$ such that there is an action path $\{\eta_s, \hat{\theta}_s\}_{s\geq t}$ for which the history $\Big(h_{t-}, \{N_s^I, F_s\}_{s\geq t}, \{\eta_s, \hat{\theta}_s\}_{s\geq t}\Big)$ is consistent with the principal's strategy.*

Lemma A (Existence and Uniqueness of consistent Outcome Path)**.** *Given any possible history $h_{u-} = \Big\{\pi_0, z_t, \eta_t, \hat{\theta}_t, N_t^I, F_t\Big\}_{t\in[0,u)} \cup \{\eta_u\}$, any combination of strategies $((e, \rho), (n, f))$ that are traceable and frictional yields a unique consistent path $\Big(\{\eta_t\}_{t\in(u,\infty)}, \{\hat{\theta}_t, N_t^I, F_t,\}_{t\in[u,\infty)}\Big)$ almost surely.*

Proof. The proof of Lemma A proceeds in two steps. First we show uniqueness and then existence.

Step 1: Uniqueness. Fix a pair of strategies, a history up to u, and any realization of the shock process $\{z_t\}_{t\in[u,\infty)}$. Suppose there are two distinct continuation paths $x = \{\eta_t^x, \hat{\theta}_t^x, N_t^{I^x}, F_t^x\}_{t\in[u,\infty)}$

and $y = \{\eta_t^y, \hat{\theta}_t^y, N_t^{I^y}, F_t^y\}_{t\in[u,\infty)}$ that are consistent with the strategies and the shock path. Let $\underline{t} = \inf\{t \geq u : x_t \neq y_t\}$ be the first time at which the processes differ. Strategy e maps history $h^A_{t_k^A}$ into a deterministic process $\{\eta_s\}_{s\in(t_k^A,\infty)}$ only for times t_k^A at which an intervention for the agent occurs. Likewise, strategy f maps history $h_{t_k^P}$ into a deterministic process $\{F_s\}_{s\in[t_k^P,\infty)}$ for times t_k^P with an intervention for the principal. Therefore, if $\eta_s^x \neq \eta_s^y$ for $s > u$ or $F_s^x \neq F_s^y$ for $s \geq u$, then there must also be a time $t \leq s$ with an intervention at t, i.e. $\exists k \in \mathbb{N}$ s.t. $t = t_k^A$ or $t = t_k^P$. Furthermore, we must have $h_t^x \neq h_t^y$ at this intervention. With probability 1, the realization $\{z_t\}_{t\in[u,\infty)}$ has only finitely many jumps on any closed interval. Hence, by frictionality, there are at most finitely many interventions on any closed interval. Therefore, $\underline{t}$ defined above must be an intervention time and the infimum is attained, i.e., $x_{\underline{t}} \neq y_{\underline{t}}$. We therefore must have $\hat{\theta}_{\underline{t}}^x \neq \hat{\theta}_{\underline{t}}^y$ or $N_{\underline{t}}^{I^x} \neq N_{\underline{t}}^{I^y}$ and, as $\underline{t}$ is the first such time, $h_{\underline{t}-}^x = h_{\underline{t}-}^y$. As $\hat{\theta}_{\underline{t}}^x$ and $\hat{\theta}_{\underline{t}}^y$ both result from the same strategy, this, however, implies that $\hat{\theta}_{\underline{t}}^x = \hat{\theta}_{\underline{t}}^y$, leaving as only possibility that $N_{\underline{t}}^{I^x} \neq N_{\underline{t}}^{I^y}$. This contradicts consistency of both processes with the fixed strategy (as $h_{\underline{t}-}^x = h_{\underline{t}-}^y$). Hence, any pair of traceable and frictional strategies gives at most one consistent outcome.

Step 2: Existence. Existence of a consistent outcome path is shown constructively: Start with arbitrary history $h_{u-} = \left\{\pi_0, z_t, \eta_t, \hat{\theta}_t, N_t^I, F_t\right\}_{t\in[0,u)} \cup \{\eta_u\}$ and fix a realization of the shock process $\{z_t\}_{t\in[u,\infty)}$. We apply the steps below iteratively until they give an outcome path consistent with z and the strategies for $t \geq u$: Define paths $\{\eta_t^0, \hat{\theta}_t^0, N_t^{I^0}, F_t^0\}$ equal to the history up to u and such that for $t > u :\ \eta_t^0 = e_t(h_{\max_k t_k^A < u})$, and for $t \geq u$: $\hat{\theta}_t^0 = \hat{\theta}_{u-}$, $N_t^{I^0} = N_{u-}^I$ and $\mathrm{d}F_t^0 = f_t(h_{\max_k t_k^A < u})$.[19] Let $n = 1$ and $t(1) = u$.

i) By traceability, there are paths $\{\eta_t^n, \hat{\theta}_t^n\}_{t\geq 0}$ such that, for $t < t(n)$: $\{\hat{\theta}_t^n, \eta_t^n\} = \{\eta_t^{n-1}, \hat{\theta}_t^{n-1}\}$ and that $\{\eta_t^n, \hat{\theta}_t^n, N_t^{I^{n-1}}, F_t^{n-1}\}_{t\geq 0}$ is consistent with the agent's strategy and process z for $t \geq t(n)$. Set $\{\eta_t^n, \hat{\theta}_t^n\}$ equal to these processes. Similarly, traceability implies that there exist paths $\{N_t^{I^n}, F_t^n\}$ with $(N_t^{I^n}, F_t^n) = (N_t^{I^{n-1}}, F_t^{n-1})$ for $t < t(n)$ and such that $\{\eta_t^n, \hat{\theta}_t^n, N_t^{I^n}, F_t^n\}_{t\geq 0}$ is consistent with the principal's strategy on $t \geq u$. Set $\{N_t^{I^n}, F_t^n\}$ equal to these processes and continue to step (ii).

ii) If $\{\eta_t^n, \hat{\theta}_t^n, N_t^{I^n}, F_t^n\}$ is consistent with the strategies for all $t \in [u,\infty)$, stop the procedure. The proof is complete. Otherwise, redefine $n = n+1$ and set $t(n+1)$ equal to the largest time v such that there is an intervention at v and $\{\eta_t^n, \hat{\theta}_t^n, N_t^{I^n}, F_t^n\}$ is consistent with the strategies for all $t \in [u, v)$, go to step (i).

If the above procedure stops after finite n, that's because of having given a consistent process and the proof is complete. In the case in which it does not stop after finitely many iterations,

$$\lim_{n\to\infty} \{\eta_t^n, \hat{\theta}_t^n, N_t^{I^n}, F_t^n\}_{t\geq 0}$$

is consistent with the strategies on $[u,\infty)$ with probability one. To see this, note that for every n, $t(n+2) > t(n)$. Given that, with probability one, any finite interval has only finitely many interventions, $\lim_{n\to\infty} t(n) = \infty$ which implies consistency of the resulting process for all $t \in [u,\infty)$. □

[19]That is, report and inspections are held constant from u onward and fines and effort are chosen according to the strategies (depending only on the last intervention before u) for the case that no further interventions occur.

Appendix B: Proofs

B.1 Proofs of intermediate results

Proof of Lemma 1. Denote by $\mathcal{F}$ the filtration generated by the random processes θ, $\hat{\theta}$ and ν^I. Define

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

The corresponding representation in differential form is

$$dW_t = e^{-rt}(-\,\mathrm{d}F_t - c\eta_t\,\mathrm{d}t) - re^{-rt}U_t + e^{-rt}\,\mathrm{d}U_t. \tag{19}$$

The process $\{W_t\}$ is an $\mathcal{F}$-martingale by construction. By the martingale representation theorem for marked point processes (Last and Brandt, 1995, Theorem 1.13.2), there exist $\mathcal{F}$-predictable functions $\tilde{\Delta}_t^\theta$, $\tilde{\Delta}_t^{\hat{\theta}}$ and $\tilde{\Delta}_t^I$ such that

$$dW_t = \sum_{a\in\{\theta,\hat{\theta},I\}} \tilde{\Delta}_t^a(\mathrm{d}N_t^a - \mathrm{d}\nu_t^a) \tag{20}$$

Replacing $\tilde{\Delta}_t^a = e^{-rt}\Delta_t^a$ and then equating (19) and (20) yields

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

This is the representation of the evolution of promised utilities shown in the lemma. □

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

(i) $(r + q_t(1))\Delta_t^{\hat{\theta}} - \mathrm{d}\nu_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}}) \geq d\Delta_t^{\hat{\theta}}$ *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 non-truthfully, 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

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

Using Lemma 1 to replace $\mathrm{d}U_t$ yields

$$dW_t = \left(e^{-rt}(-\mathrm{d}F_t - c\eta_t\,\mathrm{d}t) - re^{-rt}U_t\,\mathrm{d}t + e^{-rt}\left(rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t + \sum_{a\in\{\theta,\hat{\theta},I\}} \Delta_t^a(\mathrm{d}N_t^a - \mathrm{d}\nu_t^a) \right) \right)$$
$$= e^{-rt}\left((1-\eta_t)c\,\mathrm{d}t + \sum_{a\in\{\theta,\hat{\theta}\}} \Delta_t^a(\mathrm{d}N_t^a - q_t(1)\,\mathrm{d}t) + \Delta_t^I(\mathrm{d}N_t^I - \mathrm{d}\nu_t^I) \right),$$

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

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

It follows from Condition (ii) that

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

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

$$\mathbb{E}_0^A[W_t] = \mathbb{E}_0^A\left[\int_0^t dW_s + W_0 \right] = \int_0^t \mathbb{E}_0^A\,[dW_s] + \mathbb{E}_0^A\,[W_0] = \int_0^t \mathbb{E}_0^A\left[\mathbb{E}_s^A[dW_s] \right] + W_0 \le W_0. \tag{21}$$

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 non-truthful 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 the following:

$$\begin{aligned} d\tilde{U}_t &= \tilde{U}_t - \tilde{U}_{t-\mathrm{dt}} \\ &= \mathrm{d}N_t^\theta(U_t - U_{t-\mathrm{dt}} - \Delta_{t-\mathrm{dt}}^{\hat{\theta}}) + \mathrm{d}N_t^I(U_t + \Delta_t^I - U_{t-\mathrm{dt}} - \Delta_{t-\mathrm{dt}}^{\hat{\theta}}) \\ &\qquad + (1 - \mathrm{d}N_t^\theta - \mathrm{d}N_t^I)(U_t + \Delta_t^{\hat{\theta}} - U_{t-\mathrm{dt}} - \Delta_{t-\mathrm{dt}}^{\hat{\theta}}) \\ &= \mathrm{d}N_t^\theta(\mathrm{d}U_t + d\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}}) + \mathrm{d}N_t^I(\mathrm{d}U_t + d\Delta_t^{\hat{\theta}} - \Delta_t^{\hat{\theta}} + \Delta_t^I) + (1 - \mathrm{d}N_t^\theta - \mathrm{d}N_t^I)(\mathrm{d}U_t + d\Delta_t^{\hat{\theta}}) \\ &= \mathrm{d}U_t + d\Delta_t^{\hat{\theta}} - \mathrm{d}N_t^\theta\Delta_t^{\hat{\theta}} + \mathrm{d}N_t^I(\Delta_t^I - \Delta_t^{\hat{\theta}}). \end{aligned} \tag{22}$$

Using again Lemma 1 to replace $\mathrm{d}U_t$, we obtain

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

It follows from the honesty constraint (i) that, in expectation, $d\Delta_t^{\hat{\theta}} \le (r + q_t(1))\Delta_t^{\hat{\theta}} - \mathrm{d}\nu_t(\Delta_t^I - \Delta_t^{\hat{\theta}})$.

When we substitute it into dW_t and simplify, using again $\tilde{U}_t = U_t + \Delta_t^{\hat{\theta}}$, we obtain

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

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 $\mathbb{E}_t^A[dW_t] \leq 0$. By the same argument as in (21), we have

$$\mathbb{E}_0^A[W_t] \leq W_0 = U_0.$$

so that the agent cannot profit from deviating. Taking the limit, we find that

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

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. □

Proof of Lemma 2*.* We show that condition (Pk) follows from Lemma 1 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 $\{U_t\}_{t\geq 0}$ for the agent, and denote by $\{U_t^1, U_t^0\}_{t\geq 0}$ the associated pair of promised utilities defined in Equation (9).
(1.) By the definition of U_t^0, U_t^1, we have

$$\text{(23)} \qquad \Delta_t^\theta + \Delta_t^{\hat{\theta}} = \begin{cases} U_t^1 - U_t^0 & \text{if } \theta_{t-} = \hat{\theta}_{t-} = 0 \\ U_t^0 - U_t^1 & \text{if } \theta_{t-} = \hat{\theta}_{t-} = 1 \end{cases}, \qquad q_t(1) = q_t(1) = \begin{cases} \alpha\lambda & \text{if } \theta_{t-} = 0 \\ (1-\alpha)\lambda & \text{if } \theta_{t-} = 1 \end{cases}.$$

Combining these two expressions, we can write more succinctly:

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

Lemma 1 implies that, conditional on the event that $\mathrm{d}N_t^\theta = \mathrm{d}N_t^{\hat{\theta}} = \mathrm{d}N_t^I = 0$, we have

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

which proves condition (Pk) in Lemma 2. Note that Δ_t^I measures the difference in utility before and after an inspection when the agent reports his type truthfully. (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

$$
\begin{aligned}
\mathrm{d}U_t^i &= (U_{t+\mathrm{d}t} + \Delta_{t+\mathrm{d}t}^{\hat{\theta}}) - (U_t + \Delta_t^{\hat{\theta}}) \\
&= rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t - q_t(1)\Delta_t^{\theta} + d\Delta_t^{\hat{\theta}} \\
&\leq rU_t\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t - q_t(1)\Delta_t^{\theta} + (r + q_t(1))\Delta_t^{\hat{\theta}} - \mathrm{d}\nu_t(\Delta_t^I - \Delta_t^{\hat{\theta}}) \\
&= r(U_t + \Delta_t^{\hat{\theta}}) - q_t(1)(\Delta_t^{\theta} - \Delta_t^{\hat{\theta}}) - \mathrm{d}\nu_t(\Delta_t^I - \Delta_t^{\hat{\theta}}) + \mathrm{d}F_t + c\,\mathrm{d}t \\
&= rU_t^i + \lambda(i - \alpha)(U_t^1 - U_t^0) - \mathrm{d}\nu_t(\Delta_t^I - \Delta_t^{\hat{\theta}}) + \mathrm{d}F_t + c\,\mathrm{d}t
\end{aligned}
\tag{24}
$$

The second line follows from Lemma 1, 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 (24) holds since

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

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

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

Substituting this last equation into Equation (24) yields

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

which is equal to Condition (H) in Lemma 2. 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. (3.) Substituting Equation (23)

into the obedience constraint (ii) we obtain for each θ_{t-}:

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

The last inequality is identical to (O) in Lemma 2. Conversely, if (ii) is violated at some t, then the inequality is reversed, so that (O) is violated. □

Proof of Proposition 3. Take any truthful maximally compliant equilibrium. The following steps present a modified inspection schedule that satisfies the properties stated in Proposition 3. Let U_t^0, U_t^1 be the equilibrium continuation payoffs of the agent in this equilibrium. As the original equilibrium is truthful and maximally compliant, U_t^0 and U_t^1 satisfy the constraints from Lemma 2. First, we argue for any inspection following a high report, it is without loss to assume that truthful (high) reports are never punished more than (low) misreports at the time of an inspection. That is, if the persistent payoff U_t^1 jumps downward 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_{t-}^1 be the payoff just prior to time t. Recall that by definition $\Delta_t^I = \bar{U}_t^1 - U_{t-}^1$. We show that, without loss, $\Delta_t^I > -B - U_t^0$.

Suppose, to the contrary, that $\Delta_t^I \leq -B - U_t^0 \leq 0$. 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 truthtelling 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 $d\hat{F}_t = \mathrm{d}F_t - \mathrm{d}\nu_t^I \Delta_t^I$, where $\mathrm{d}F_t$ denotes the fine specified in the original equilibrium. Hence, the agent's expected loss from the inspection caused by $\Delta_t^I < 0$ is paid as a fine at time t. This way, the path of persistent payoff U_t^1 remains unchanged for all $s \leq t$. Similarly, the path of transitional utility, U_s^0, remains unchanged as the continuation equilibria after a transition remain the same. As both paths U_s^1 and U_s^0 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 $j = 0$:

$$\mathrm{d}U_t^0 \leq r(U_t^0)\,\mathrm{d}t - \lambda\alpha(U_t^1 - U_t^0)\,\mathrm{d}t + \mathrm{d}\nu_t^I(B + U_t^0) + \mathrm{d}F_t + c\,\mathrm{d}t.$$

The effect of the proposed change on the right-hand-side of this constraint is $-\,\mathrm{d}\nu_t^I(B + U_t^0) - \mathrm{d}\nu_t^I \Delta_t^I$. As $\Delta_t^I \leq -B - U_t^0$, this effect is positive and the path of U_s^0 still satisfies the honesty constraint. In order to randomize at time t in the original equilibrium, the principal must have been indifferent between inspecting and continuing without, so that removing instant t from the support weakly lowers inspection costs. Now, with $\Delta_t^I > -B - U_t^0$, we prove Proposition 3. Suppose, towards 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 non-random 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. We show at the end of the proof how the arguments extend 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 $\Delta_{t^0}^I \geq 0$, introduce an additional fine at t^0 so that the new fine is given by $d\hat{F}_{t^0} = \mathrm{d}F_{t^0} + (1 - \mathrm{d}\nu_{t^0}^I)\Delta_{t^0}^I$, where $\mathrm{d}F_{t^0}$ denotes the fine in the original equilibrium. The payoff paths U_s^0 and U_s^1 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 $\Delta_{t^0}^I < 0$, increasing the inspection probability from $\mathrm{d}\nu_{t^0}^I$ to 1 decreases the persistent payoff path U_s^1 for all $s \leq t^0$ by $|\Delta_{t^0}^I|(1 - \mathrm{d}\nu_{t^0}^I)e^{-(r+(1-\alpha)\lambda)(t^0 - s)}$. 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 truthtelling by lowering the transitional payoff by the necessary amount. To this end, introduce an additional transition fine of $|\Delta_{t^0}^I|(1 - \mathrm{d}\nu_{t^0}^I)e^{-(r+(1-\alpha)\lambda)(t^0 - s)}$ to be paid at time $s \leq t^0$ if a transition to the bad state occurs. This additional fine ensures that the difference $U_s^1 - U_s^0$ 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 $U_s^0 - |\Delta_{t^0}^I|(1-\mathrm{d}\nu_{t^0}^I)e^{-(r+(1-\alpha)\lambda)(t^0-s)} \geq -B$. This term is decreasing in s, so it is sufficient to verify that $U_{t^0}^0 + \Delta_{t^0}^I(1-\mathrm{d}\nu_{t^0}^I) \geq -B$. Recall that we have shown that for any inspection time, $\Delta_{t^0}^I > -B - U_{t^0}^0$. Feasibility follows since $\mathrm{d}\nu_{t^0}^I < 1$. Last, we show how the arguments extend to the case $t^0 \notin \mathcal{T}$. First, note that if its infimum t^0 is not contained in the set $\mathcal{T}$, then for any $\delta > 0$, we can find an $\epsilon \in (0, \delta)$ such that $t^0 + \epsilon \in \mathcal{T}$. Further, by choosing δ small enough, we can ensure that the expected inspection probability $\int_{t^0}^{t^0+\delta} \mathrm{d}\nu_s^I$ becomes arbitrarily small. In the first case with $U_{t^0}^0 > -B$, there exists an $\epsilon > 0$ small enough such that $t^0 + \epsilon \in \mathcal{T}$ and also $U_{t^0+\epsilon}^0 > -B$ by right-continuity of U_t^0. In this case we can apply the argument above to schedule a predictable inspection at time $t^0 + \epsilon$. To satisfy the agent's incentive constraints, this modification is paired either with an additional fine after a high report at $t^0 + \epsilon$ or with an additional transition fine for any transition at times $s \in [t^0, t^0 + \epsilon)$, depending on the sign of $\Delta_{t^0+\epsilon}^I$. In the second case with $U_s^0 = -B$, on $[t^0, t^0 + \delta)$ for some $\delta > 0$, then by $\Delta_s^I > -B - U_s^0$, we have that $\Delta_{t^0+\epsilon}^I > 0$. In this case, we can proceed in a similar way as above and introduce an additional fine to compensate for the increase in the agent's expected payoff caused by performing the inspection with probability 1 and keep the path of persistent payoffs U_s^1 unchanged for $s \leq t^0$. However, to ensure that the obedience and honesty constraints are also satisfied on $(t^0, t^0 + \epsilon]$, the fine is increased gradually on the interval $(t^0, t^0 + \epsilon)$. Specifically, construct the fine such that the honesty constraint (H) binds (with $U_s^0 = -B$):

$$0 = -rB\,\mathrm{d}t - \lambda\alpha(U_s^1 + B)\,\mathrm{d}t + \mathrm{d}F_t + c\,\mathrm{d}t. \tag{25}$$

In the promise-keeping constraint (Pk), substituting for $\mathrm{d}F_s$ with the binding honesty constraint (25) and inserting $U_s^0 = -B$ determines the evolution of U_s^1 on $(t^0, t^0 + \epsilon)$ via the differential equation

$$\hat{u}_s' = (r + \lambda)(\hat{u}_s + B).$$

We keep the persistent utility at t^0 unchanged, so the initial condition for the ODE is $\hat{u}_{t^0} = U_{t^0}^1$, which leads to the solution

$$\hat{u}_s = U_{t^0}^1 e^{(r+\lambda)(s-t^0)} + B\left(e^{(r+\lambda)(s-t^0)} - 1\right),$$

for $s \in [t^0, t^0 + \epsilon)$. To ensure, that this trajectory of persistent utility is feasible, we verify that the fine $\mathrm{d}Fs$ is positive and that the solution $\hat{u}_{t^0+\epsilon}$ does not exceed $U_{t^0+\epsilon}^1 + \Delta_{t^0+\epsilon}^I$ from the original equilibrium. The latter is necessary to reach $U_{t^0+\epsilon}^1 + \Delta_{t^0+\epsilon}^I$ as the continuation payoff after inspection at $t^0 + \epsilon$. For the fine, (25) with $U_s^1 = \hat{u}_s$ gives

$$\frac{\mathrm{d}F_s}{\mathrm{d}t} = -c + rB + \lambda\alpha(\hat{u}_s + B) = -c + rB + \lambda\alpha(U_{t^0}^1 + B)e^{(r+\lambda)(s-t^0)}.$$

This term is decreasing in s and therefore smallest at $s = t^0$, where it is positive if

$$(r + \lambda\alpha)B + \lambda\alpha U_{t^0}^1 \geq c.$$

For the original equilibrium to satisfy the obedience constraint we must have $U_{t^0}^1 \geq -B + \frac{c}{\lambda\alpha}$, so that the above inequality must be satisfied and the fines are positive. To check that $\hat{u}_{t^0+\epsilon}$ constructed above does not lie above $U_{t^0+\epsilon}^1 + \Delta_{t^0+\epsilon}^I$ from the original equilibrium, note that the inspections in the original equilibrium had no effect on the honesty constraint (H) as, by assumption, we are in the case $U_s^0 = -B$. Therefore, as the original equilibrium satisfied the honesty constraints, the evolution of $\hat{u}_s$, which was constructed by making the honesty constraint binding, must lie weakly below the original U_s^1

and therefore $\hat{u}_{t^0+\epsilon} \le U^1_{t^0+\epsilon} + \Delta^I_{t^0+\epsilon}$ since $\Delta^I_{t^0+\epsilon}$ is positive by $\Delta^I_{t^0+\epsilon} > -B - U^0_{t^0} = 0$. Hence, the newly constructed equilibrium includes a fine at inspection time $t^0 + \epsilon$ of $\hat{U}^1_{t^0+\epsilon} - (U^1_{t^0+\epsilon} + \Delta^I_{t^0+\epsilon})$ so that the persistent utility increases to the one from the original continuation equilibrium after inspection at time $t^0 + \epsilon$. 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. □

Proof of Proposition 4. We show that any predictable principal strategy that generates a positive value for the principal at each t can be implemented in equilibrium. First, note that for any history, there is a possible continuation equilibrium in which the agent chooses to exit the relationship with probability one. To support exit by the agent as a best response, the principal's strategy is such that whenever the agent deviates and fails to exit although he was supposed to do so, the principal implements the harshest possible fine of $-B$. We show that this bad continuation equilibrium can be leveraged to support any principal strategy as an equilibrium given that leads to predictable inspections for the agent.

Let $\{N_t, F_t\}$ be the paths induced by the strategy 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 from the result. Adapt the agent's strategy such that he exits after any history h_t with $\mathrm{d}N^I_t \neq d\tilde{N}^I_t$, that is, whenever the agent observes that the principal deviated from the original inspection strategy. Define the set $D = \{t | \, \mathrm{d}N^I_t \neq d\tilde{N}^I_t\}$ containing the dates at which the agent observes that the principal deviates from her original inspection strategy. Since the payoff for the principal from the strategy in the result is positive 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 payoff of 0 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 which leads to exit at $t = \inf D$ is incentive compatible, and the constructed equilibrium differs from the initial strategy profile in Proposition 4 at most off the equilibrium path. □

B.2 Proof of Theorem 1

The general outline of the proof is as follows. First, we consider a relaxed mechanism design problem in which the honesty constraint applies only while the agent reports compliance. We then solve for the optimal mechanism in this case using an iteration argument, assuming that fines are levied only at the time of inspections or transitions. Second, we show that fines between inspections cannot increase the principal's payoff. Third, we verify that there is no mechanism in non-Markov strategies that performs better than the optimal Markovian mechanism of the relaxed problem. Finally, we show that the solution to the relaxed problem is also achievable in the original problem.

B.2.1 Auxiliary control problem

We begin by considering the auxiliary control problem

$$\max_{\{N^I_t, F_t\}_{t\ge 0}} \mathbb{E}^P \left[\int_0^\infty e^{-rt} (\theta_t R \,\mathrm{d}t - \kappa \,\mathrm{d}N^I_t) \right] \tag{26}$$

subject to the incentive-compatibility conditions (H), (O), (P), and the following two 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 (i.e., $\mathrm{d}\mu_t = 0$).

(B) When $\hat{\theta}_t = 0$, then 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, while Condition B relaxes the incentive-compatibility restriction for the agent. 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 which 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 continue iteratively and consider the limit as the total number of available inspections k goes to infinity. More specifically, for any integer $k \geq 0$, consider the following optimization problem:

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

subject to $\lim_{t\to\infty} N_t^I \leq k$ pathwise, and to the incentive-compatibility conditions (H), (O), (P), (A) and (B) at all $t \geq 0$ at which $N_t^I < k$. Note that in order for a solution to exist, we cannot impose incentive compatibility at any t with $N_t = k$, as the obedience constraint (O) is necessarily violated when the principal cannot inspect. Denote by V_k the solution to the problem with k available inspections. It then follows from Proposition 54.18 in Davis (1993) that the value function for the auxiliary problem V is the limit of V_k, i.e., $V = \lim_{k\to\infty} V_k$. In the second part of the proof, we show that Assumption A is without loss. Indeed, due to risk-neutrality, any fines the agent has to pay between inspections can be moved while maintaining incentive compabitility. We then confirm that there is no non-Markovian mechanism that generates a larger payoff for the principal. Finally, we show that in the optimal mechanism, the honesty constraint holds in both states, so that the solution to the auxiliary problem is also a solution to our original maximization problem in which we do impose Conditions (A) and (B).

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

Claim 1. *Along the equilibrium path of a maximally compliant mechanism, we have* $U_t^1 \leq -\frac{c}{r\alpha}$.

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

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

Solving the integral yields

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

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

By Assumption A, the promise-keeping and truthtelling constraints in state $\hat{\theta}_t = 1$ yield a system of coupled first-order differential equations

$$\frac{\mathrm{d}U_t^1}{\mathrm{d}t} = rU_t^1 + \lambda(1-\alpha)(U_t^1 - U_t^0) + c,$$
$$\frac{\mathrm{d}U_t^0}{\mathrm{d}t} = rU_t^0 - \lambda\alpha(U_t^1 - U_t^0) + c,$$

which, for given initial values $U_0^1 = u^1$ and $U_0^0 = u^0$, has the unique solution

$$U_t^1 = e^{rt}(u^1 + (1-\alpha)(e^{\lambda t} - 1)(u^1 - u^0)) - c(1 - e^{rt})/r, \tag{27}$$
$$U_t^0 = e^{rt}(u^0 - \alpha(e^{\lambda t} - 1)(u^1 - u^0)) - c(1 - e^{rt})/r. \tag{28}$$

Inspection of (27) and (28) reveals 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

$$\phi_1(t,u) := e^{rt}\left(u + \left(\frac{1-\alpha}{\alpha}\right)(e^{\lambda t} - 1)\frac{c}{\lambda}\right) - c(1 - e^{rt})/r \tag{29}$$
$$\phi_0(t,u) := e^{rt}\left(u - \left(\frac{1-\alpha}{\alpha}\right)\frac{c}{\lambda} - e^{\lambda t}\frac{c}{\lambda}\right) - c(1 - e^{rt})/r. \tag{30}$$

It is easy to see that for $u < -\frac{c}{r\alpha}$, the function $\phi_1(t,u)$ is u-shaped in t, $\phi_0(t,u)$ is strictly decreasing in t and both are linearly increasing in u for all t. Define the boundary hitting times

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

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-concave.*

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

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

It is immediate that ϕ_1 and ϕ_0 are increasing in u. Therefore, an increase in u decreases $T^1(u)$ and increases $T^0(u)$ and vice versa. Therefore, T is quasi-concave, and T assumes its maximum at the point u_1^* at which T^0 and T^1 are equal, that is:

$$T^0(u) = T^1(u).$$

Hence, $T^{0\prime}(u) < 0$ and $T^{1\prime}(u) > 0$ and, consequently,

$$T'(u)\begin{cases} > 0 & \text{if } u < u_1^* \\ < 0 & \text{if } u > u_1^*. \end{cases} \tag{31}$$

□

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 $u < -\frac{c}{r\alpha}$ 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

$$\underbrace{\frac{\alpha}{1-\alpha}\left(-\frac{\lambda}{r} - \frac{\lambda u}{c}\right)}_{=:\text{LHS}} = \underbrace{\frac{e^{(r+\lambda)t} - 1}{e^{rt} - 1} - 1}_{=:\text{RHS}}. \tag{32}$$

It is easy to see that RHS is increasing and convex in t and that

$$\lim_{t\to 0} \frac{e^{(r+\lambda)t} - 1}{e^{rt} - 1} - 1 = \lim_{t\to 0} \frac{(r+\lambda)e^{(r+\lambda)t}}{re^{rt}} - 1 = \frac{\lambda}{r}.$$

LHS is clearly strictly decreasing in u, and for $u < -\frac{c}{r\alpha}$, 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 $u < -\frac{c}{r\alpha}$, there is a unique time $T_s(u)$ such that $\phi_1(T_s(u), u) = u$. Moreover, inspection of (32) reveals that this time is continuous and strictly decreasing in u. Note that for $u \to -\frac{c}{r\alpha}$, we have $T_0(u) > T_s(u) \to 0$ and for $u \to -B + \frac{c}{\lambda\alpha}$ 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. $\phi_1(T(u), u) > u$ *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 non-compliance. We show that during reports of non-compliance, the utility of the agent is held constant Define $\beta_1 = \frac{\lambda\alpha}{r+\lambda\alpha}$.

Claim 5. *Let* (V_k^0, V_k^1) *be the value functions in an optimal mechanism when there are* $k \geq 1$ *available inspections. Denote the pair of initial promised utilities in this mechanism by* $u^* = (u^{0*}, u^{1*})$. *Then*

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

Proof. Without loss, assume $\theta_0 = 1$. We establish the claim via contradiction. Suppose to the contrary that $V_k^0(U_t) < \beta_1 V_k^1(u^*)$, 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 payoff for the principal at time t in state $\theta_t = 0$ in the new mechanism is therefore

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

Since the new mechanism is identical to the original mechanism for $\theta_t = 1$, the expected payoff for the principal in the new mechanism is strictly higher 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 k. We derive the optimal initial promised utility u_k^* for each k, and we show that the sequence $\{u_k^*\}$ converges to $\bar{u}$ as $k \to \infty$. For the case in which the agent reports $\hat{\theta}_t = 0$, Claim 5 implies that without loss the expected payoff for the principal in state $\theta_t = 0$ with k available inspections can be written as $V_k^0(u_0, u_1) = \beta_1 V_k^1(u_1)$. We show that for all $k \geq 0$, the obedience constraint (O) binds at the outset.

Claim 6. *Suppose the total number of available inspections in k. Then there is an optimal policy such that at the initial pair of promised utility (u_k^0, u_k^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 $u^1 - u^0 > \frac{c}{\lambda\alpha}$. Denote by t^* the minimum 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

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

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

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

Thus, by shifting up u^0 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 $u^1 - u^0 > \frac{c}{\lambda\alpha}$ cannot be optimal. □

Without loss, we can now restrict attention to initial pairs of utility (u^0, u^1) such that $u^1 - u^0 = \frac{c}{\lambda\alpha}$. 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

$$V_k^1(u) = \max_{\substack{0\leq t\leq T(u)\\ u'\geq\phi_1(t,u)}} \int_0^t e^{-(r+\lambda(1-\alpha))s}(R + \lambda(1-\alpha)V_k^0)\ \mathrm{d}s + e^{-(r+\lambda(1-\alpha))t}\left(V_{k-1}^1(u') - \kappa\right) \tag{33}$$

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_k^* be a maximizer of V_k^1 and denote by t_k^* the associated first inspection date.

Claim 7. *Let u_{k-1}^* be a maximizer of $V_{k-1}^1(u)$ and suppose ${V_{k-1}^1}'(u) < 0$ for all $u > u_{k-1}^*$. Then, $t_k^* = T^0(u_k^*)$ and $\phi_1(t_k^*, u_k^*) > u_{k-1}^*$.*

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

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 payoff for given utility u is therefore:

$$V_k^1(u) = \int_0^{T(u)} e^{-(r+\lambda(1-\alpha))s}(R + \lambda(1-\alpha)V_k^0)\ \mathrm{d}s + e^{-(r+\lambda(1-\alpha))T(u)}(V_{k-1}^1(\phi_1(t,u)) - \kappa).$$

Define $\beta_0 = \frac{\lambda\alpha}{r+\lambda\alpha}$ and $\beta_1 = \frac{\lambda(1-\alpha)}{r+\lambda(1-\alpha)}$. Solving the integrals and performing a few simple rearrangements, the principal's payoff can be expressed more succinctly as

$$V_k^1(u) = a(u) + b(u)V_{k-1}^1(\phi_1(T(u), u)),$$

where

$$\begin{aligned} a(u) &= \frac{1 - e^{-(r+\lambda(1-\alpha))T(u)}}{1 - \beta_0\beta_1 + \beta_0\beta_1 e^{-(r+\lambda(1-\alpha))T(u)}} \frac{R}{r + \lambda(1-\alpha)} - \frac{e^{-(r+\lambda(1-\alpha))T(u)}}{1 - \beta_0\beta_1 + \beta_0\beta_1 e^{-(r+\lambda(1-\alpha))T(u)}}\kappa, \\ b(u) &= \frac{e^{-(r+\lambda(1-\alpha))T(u)}}{1 - \beta_0\beta_1 + \beta_0\beta_1 e^{-(r+\lambda(1-\alpha))T(u)}}, \end{aligned}$$

Simple calculus reveals

$$a'(u) = \frac{\left(e^{(r+\lambda-\alpha\lambda)T(u)}(r+\lambda-\alpha\lambda)^2(r+\alpha\lambda)(r\kappa(r+\lambda) + R(r+\alpha\lambda))\right)}{\left((1-\alpha)\alpha\lambda e^{(r+\lambda-\alpha\lambda)T(u)}r(r+\lambda)\right)^2}T'(u)$$

and

$$b'(u) = -\frac{e^{(r+\lambda-\alpha\lambda)T(u)}r(r+\lambda)(r+\lambda-\alpha\lambda)^2(r+\alpha\lambda)}{\left((1-\alpha)\alpha\lambda^2 + e^{(r+\lambda-\alpha\lambda)}r(r+\lambda)\right)^2}T'(u),$$

so that sign $a'(u) = -$ sign $b'(u) =$ sign $T'(u)$. From (31), it follows that

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

Step 0: Consider the case $k = 0$, so the principal cannot perform any interventions. The obedience constraint is then necessarily violated (all penalties must be enforced independently of the true state) and thus no effort by the agent can be induced. Thus, the value function for the principal is

$$\bar{V}_0 = \int_0^\infty e^{-(r+\lambda(1-\alpha))t}\lambda(1-\alpha)R \ \mathrm{d}t = \frac{\lambda(1-\alpha)}{r+\lambda(1-\alpha)}R.$$

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

$$V_1^1(u) = a(u) + b(u)\bar{V}_0$$

The marginal utility at a given utility u is

$$V_1^{1\prime}(u) = a'(u) + b'(u)\bar{V}_0.$$

We have $V_1^{1\prime}(u) < 0$ for $u > u_1^* > \bar{u}$ and $V_1^{1\prime}(u) > 0$ for $u < u_1{}^*$, thus $u_1{}^*$ maximizes V_1^1.

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

$$V_2^1(u) = a(u) + b(u)V_1^1(\phi_1(T(u), u)).$$

When $u > u_1^*$, then $V_2^1(u) < 0$, and therefore $u < u_1^*$. Because $V_2^1(u)$ is minimized when u lies at the participation boundary, and continuous in between, there must be a maximum u_2^*. The marginal utility is

$$V_2^{1\prime}(u) = a'(u) + b'(u)V_1(\phi_1(T(u), u)) + b(u)D_u\phi_1(T(u), u))V_1^{1\prime}(\phi_1(T(u), u)).$$

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

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

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

$$\begin{aligned} V_2^{1\prime}(u) &= a'(u) + b'(u)V_1(\phi_1(T(u), u)) + b(u)D_u\phi_1(T(u), u))V_1'(\phi_1(T(u), u)) \\ &< a'(u) + b'(u)V_1(\phi_1(T(u), u)) \\ &< a'(u) + b'(u)\bar{V}_0 = V_1^{1\prime}(u) \end{aligned}$$

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

Step k: *Suppose there are* k *inspections available. Our induction hypothesis is that* V_j^1 *has a maximum at* u_j^* *where* $u_j^* < u_{j-1}^*$ *for all* $j = 2, \ldots, k-1$ *and that* ${V_{k-1}^1}'(u) < {V_{k-2}^1}'(u)$ *for* $u > u_{k-2}^*$. By the same arguments as in Step 2, the principal's payoff,

$$V_k^1(u) = a(u) + b(u)V_k^1(\phi_1(T(u), u)),$$

has a maximum u_k^*. The marginal payoff at $u > u_{k-1}^*$ $(> \bar{u})$ is

$$\begin{aligned} {V_k^1}'(u) &= a'(u) + b'(u)V_{k-1}(\phi_1(T(u),u)) + b(u)D_u\phi_1(T(u),u))V_{k-1}'(\phi_1(T(u),u)) \\ &< a'(u) + b'(u)V_{k-1}^1(u) + b(u)D_u\phi_1(T(u),u))V_{k-2}'(\phi_1(T(u),u)) \\ &< a'(u) + b'(u)V_{k-2}^1(u) + b(u)D_u\phi_1(T(u),u))V_{k-2}'(\phi_1(T(u),u)) \end{aligned}$$

where the first line follows from our induction hypothesis. Therefore, $u \geq u_{k-1}^*$ implies ${V_k^1}'(u) < {V_{k-1}^1}'(u) < 0$. The induction shows that $u_k^* < u_{k-1}^*$ for all $k \geq 0$. It follows immediately from the definition of $\bar{u}$ that $u_k^* > \bar{u}$ for all k. Hence $\{u_k^*\}$ is a decreasing and bounded sequences, so that by the monotone convergence theorem, the sequence converges to a limit $\hat{u} \geq \bar{u}$. Since $\{u_k^*\}$ is convergent, it is a Cauchy sequence, so that

$$\lim_{k\to\infty} |u_k^* - u_{k-1}^*| = \lim_{k\to\infty} |u_k^* - \phi_1(T(u_k^*), u_k^*)| = 0 \Rightarrow \hat{u} = \bar{u}$$

by Claim 3.

B.2.2 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 k, the principal cannot gain from imposing fines between inspections when k inspections are left. Consider again Step k of the iteration in the previous section. By the same argument as before, we have $u^1 - u^0 = \frac{c}{\lambda\alpha}$ 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

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

The evolution of the difference in utilities is

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

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 k, and there no fines in step $k-1$ onwards, we must have $U_t^1 = \phi_1(T(u_k^*), u_k^*)$. This means that the policy of the previous section with initial promised utilities $(u_k^*, u_k^* - \frac{c}{\lambda\alpha})$ remains optimal even when fines between inspections are available.

B.2.3 General mechanisms in the relaxed problem

Parts (1.)-(2.) demonstrate that the mechanism described in the theorem is an optimal Markovian mechanism under the relaxing Assumption B. It remains to verify that no (non-Markovian) mechanism can do better. Let $V^{\theta_t}(U)$ denote the expected value 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 $V^{\theta_t}(U_0)$. 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, V_k, the value function with no more than k inspections, converges to value function V of the problem without bound on the number of inspections, and that V is the unique bounded and continuous function that solves the quasi-variational inequality

$$\begin{aligned} \mathcal{U}V^\theta(u) - rV^\theta(u) &\leq 0, \\ WV^\theta(u) - V^\theta(u) &\leq 0, \\ \left(\mathcal{U}V^\theta(u) - rV^\theta(u)\right)\left(WV^\theta(u) - V^\theta(u)\right) &= 0, \end{aligned}$$

on the state space $\left\{(\theta, u^0, u^1) \colon \theta \in \{0,1\}, (u^0, u^1) \in [-B, 0]^2, u^1 - u^0 \geq \frac{c}{\lambda\alpha}\right\}$. Here, $\mathcal{U}$ denotes the extended generator of the piecewise deterministic Markov process which is defined by the relationship[20]

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

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

$$WV^\theta = \max_{u_0, u_1} V^\theta(u_0, u_1) - \kappa.$$

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

$$G_t = \int_0^t e^{-rs}\left(R\theta_s\,\mathrm{d}s - \kappa\,\mathrm{d}N_s^I\right) + e^{-rt}V^{\theta_t}(U_t).$$

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

$$\mathbb{E}_s[G_t] - G_s = \int_s^t e^{-r(z-s)}\left(\mathcal{U}V^{\theta_z}(U_z) - rV^{\theta_z}(U_z)\right)\mathrm{d}z + \mathbb{E}_s\left[\int_s^t e^{-r(z-s)}\left(WV^{\theta_z}(U_z) - V^{\theta_z}(U_z)\right)\mathrm{d}N_z^I\right].$$

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

[20]See Davis, 1993, pp. 27-33.

B.2.4 Optimality in the original problem

We now consider the original model, in which we remove Assumption A so that the honesty constraint holds in both states. We show that during non-compliance, 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 non-compliance is $(u^0, u^1) := (\bar{u}, \bar{u} - \frac{c}{\lambda\alpha})$. 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 $\mathrm{d}F_t = -ru^0 + \alpha\lambda(u^1 - u^0)$ and $\mathrm{d}\mu_t = c + (r+\lambda)(u^1 - u^0)$. Substituting into the promise-keeping and truthtelling 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. □

B.3 Proofs of comparative static results

Proof of Proposition 1. Define

$$\Psi(T) \equiv (B - c/r)\left(1 - e^{-rT}\right) - c/(\lambda\alpha)e^{\lambda T}(e^{rT} - \alpha) + c/(\lambda\alpha)(1-\alpha), \tag{34}$$

so that $T^* = \inf\{T > 0 : \Psi(T) = 0\}$. This exists and is unique whenever our feasibility assumption $B > c\frac{r+\lambda}{r\lambda\alpha}$ is satisfied (Ψ is increasing from 0 at $T = 0$ and crosses 0 from above exactly once). The function Ψ is continuously differentiable in all parameters and in T on a neighbourhood of T^*. By the implicit function theorem we have

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

for all parameters $x \in \{B, c, \alpha, \lambda\}$, 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

$$sign\left(\frac{\partial T^*}{\partial x}\right) = sign\left(\Psi_x|_{T=T^*}\right).$$

The first two items of Proposition 1 follow immediately as Ψ is increasing in B and decreasing in c everywhere. Likewise for the third item, note that

$$\Psi_\alpha = \frac{c}{\lambda\alpha^2}\left(e^{(r+\lambda)T} - 1\right) > 0,$$

so that T^* is increasing in α. For the fourth item, describing the change of T^* in λ, consider Ψ in (34) as $\lambda \searrow \frac{cr}{Br\alpha - c}$, which is the lower bound on λ such that the feasibility assumption $B > \frac{c(r+\lambda)}{r\lambda\alpha}$ is fulfilled. $\Psi = 0$ is then equivalent to

$$(B - \frac{c}{r})(1 - e^{-rT}) - \left(B - \frac{c}{r\alpha}\right)\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right) = 0.$$

This can only be fulfilled at $T = 0$, as we have $B > \frac{c}{r\alpha} > \frac{c}{r}$ and for all $T > 0$,

$$0 > -\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right) > \frac{1}{\alpha}\left(e^{\lambda T}(e^{rT} - \alpha) - (1-\alpha)\right).$$

Hence, T^* is initially increasing in λ. Finally, consider Ψ in (34) to see that $T^*(\lambda) \overset{\lambda\to\infty}{\longrightarrow} 0$. In particular,

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

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

Proof of Proposition 2. Let C^0_{EQ} and C^1_{EQ} denote the expected discounted inspection cost when starting in state 0 or 1, respectively. For fixed inspection cycle length, T, they follow the following nested equations

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

and

$$\begin{aligned} C^1_{EQ} &= \int_0^T e^{-(r+\lambda(1-\alpha))t}\lambda(1-\alpha)C^0_{EQ}\,\mathrm{d}t + e^{-(r+\lambda(1-\alpha))T}\left(\kappa + C^1_{EQ}\right) \\ &= \left(1-e^{(r+\lambda(1-\alpha))T}\right)\frac{\lambda(1-\alpha)}{r+\lambda(1-\alpha)}C^0_{EQ} + e^{-(r+\lambda(1-\alpha))T}\left(\kappa + C^1_{EQ}\right) \end{aligned}$$

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

$$C^1_{EQ} = \frac{r+\lambda\alpha}{r(r+\lambda)}\cdot(r+\lambda(1-\alpha))\frac{e^{-(r+\lambda(1-\alpha))T}}{1-e^{-(r+\lambda(1-\alpha))T}}\cdot\kappa,$$

the expression given in the main text before Proposition 2. The results on the first two items, considering changes in B and c follow immediately from Proposition 1 as these parameters do not enter C^1_{EQ} directly and C^1_{EQ} is decreasing in T^*. For the third item, consider the total derivative of cost C^1_{EQ} w.r.t. α:

$$\begin{aligned} \frac{d}{d\alpha}C^1_{EQ} =& \frac{1}{(e^{(r+\lambda(1-\alpha))T^*}-1)^2 r(r+\lambda)} \\ & \Big[(2\alpha-1)\lambda^2 + e^{(r+\lambda(1-\alpha))T^*}\left((T^*\lambda(r+\alpha\lambda)(r+\lambda(1-\alpha)) - (2\alpha-1)\lambda^2\right) \\ & - \frac{\partial T^*(\alpha)}{\partial\alpha}\cdot(r+\lambda(1-\alpha))^2(r+\alpha\lambda)\Big]. \end{aligned} \tag{35}$$

The change in inspection cost caused by varying α contains a cost-increasing direct effect on the environment contained in the first terms of the squared bracket and a cost-decreasing indirect effect through the increase in T^*. The first effect captures the change in relative probability of high reports as well as the volatility of the state, both of which determine how often the deadline T^* is reached without previously changing to state L. To see that this effect is always positive, verify that it is 0 at $T = 0$ and increasing in T. We establish the second item of the result: there exists $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$ such that C^1_{EQ} is decreasing in α for all $\alpha < \underline{\alpha}$. Note that as $\alpha \searrow \frac{c(r+\lambda)}{Br\lambda}$, $T^*(\alpha) \searrow 0$. The squared bracket in (35) converges to $-\frac{\partial T^*}{\partial\alpha}\big|_{\alpha\searrow\frac{c(r+\lambda)}{Br\lambda}}\cdot r(r+\lambda)^2$. $\frac{\partial T^*}{\partial\alpha}$ is strictly positive for $\alpha > \frac{c(r+\lambda)}{Br\lambda}$. By continuity, the total derivative must be negative for all α smaller than some $\underline{\alpha} > \frac{c(r+\lambda)}{Br\lambda}$. Finally, we show that there exists $\bar{\alpha}$ such that C^1_{EQ} is increasing in α for all $\alpha > \bar{\alpha}$ and that $\bar{\alpha} < 1$ whenever $\frac{Br-c}{cr}$ is large enough. Consider the squared bracket in (35) as $\alpha \nearrow 1$. This is equal to

$$\lim_{\alpha\nearrow 1} -\lambda^2 \left[e^{rT^*(\alpha)} - 1\right] + \left(\lambda T^*(\alpha) - r\frac{\partial T^*}{\partial \alpha}\right) e^{rT^*(\alpha)}.$$

From the condition $\Psi|_{\alpha=1} = 0$ we get $e^{(r+\lambda)T^*(\alpha)} \overset{\alpha\to 1}{\longrightarrow} \frac{Br-c}{cr}\lambda$. Further,

$$\begin{aligned}
\frac{\partial T^*}{\partial \alpha}\bigg|_{\alpha\nearrow 1} &= \lim_{\alpha\nearrow 1}\left(-\frac{\lambda(Br-c)(1-e^{-rT}) + cr(e^{\lambda T}-1)}{\lambda\alpha(Br-c)re^{-rT} - (r+\lambda)cre^{(r+\lambda)T} + \lambda\alpha cre^{\lambda T}}\bigg|_{T=T^*(\alpha)}\right) \\
&= \frac{\lambda(Br-c)\left(1-\left(\frac{Br-c}{cr}\lambda\right)^{-\frac{r}{r+\lambda}}\right) + cr\left(\left(\frac{Br-c}{cr}\lambda\right)^{\frac{\lambda}{r+\lambda}} - 1\right)}{-\lambda(Br-c)r\left(\frac{Br-c}{cr}\lambda\right)^{-\frac{r}{r+\lambda}} + (r+\lambda)cr\left(\frac{Br-c}{cr}\lambda\right) - \lambda cr\left(\frac{Br-c}{cr}\lambda\right)^{\frac{\lambda}{r+\lambda}}} \\
&= \frac{\lambda\frac{(Br-c)}{cr}\left(\left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}} - 1\right) + \left(\left(\frac{Br-c}{cr}\lambda\right) - \left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}}\right)}{-\lambda\frac{(Br-c)}{cr}r + (r+\lambda)\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}} - \lambda\left(\frac{Br-c}{cr}\lambda\right)} \\
&= \frac{\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}} - \left(\frac{Br-c}{cr}\lambda\right)^{\frac{r}{r+\lambda}}}{(r+\lambda)\left(\left(\frac{Br-c}{cr}\lambda\right)^{\frac{2r+\lambda}{r+\lambda}} - \frac{(Br-c)}{cr}\lambda\right)}.
\end{aligned}$$

Inserting this derivative into (35) at $\alpha = 1$ and defining $\chi = \frac{(Br-c)}{cr}\lambda > 1$, we see that the deterministic inspection cost is increasing in α if and only if

$$\lambda^2 + \chi\left(-\lambda^2 + \frac{\lambda}{r+\lambda}\ln(\chi) - \frac{\chi^{\frac{2r+\lambda}{r+\lambda}} - \chi^{\frac{r}{r+\lambda}}}{\chi^{\frac{2r+\lambda}{r+\lambda}} - \chi}\right) > 0.$$

As χ grows large (for example as B increases), the fraction in the bracket approaches 1, so the second term grows arbitrarily large. Therefore, we have that for χ large enough, there exists $\bar{\alpha} < 1$ such that the deterministic cost is increasing in α for all $\alpha > \bar{\alpha}$. In the case of λ, the first result, that the cost decreases initially in λ, is shown analogously to the corresponding result in the case of α. To see that the cost becomes arbitrarily large in the limit, recall from the previous proof that $\lambda T^*(\lambda)$ grows to ∞ at lower than logarithmic rate. The total cost in the limit is given by

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

□

B.4 Proofs for Section 4

Proof of Theorem 2. The total expected inspection costs from the random mechanism with inspection rate m^* is $m^*\frac{r+\lambda\alpha}{r(r+\lambda)}\kappa$, when starting with $\theta_0 = 1$. To prove that this cost is lower than C^1_{EQ}, the cost from the predictable inspections schedule, we show that

$$\frac{1}{m^*} - \frac{e^{(r+\lambda(1-\alpha))T(c)} - 1}{r+\lambda(1-\alpha)} \geq 0. \tag{36}$$

1. For $c = \bar{c} = \frac{Br\lambda\alpha}{r+\lambda}$, we have $T(\bar{c}) = 0$ so that the inequality in (36) holds with equality. 2. Show that the LHS of (36) is decreasing in c and therefore positive for all $c < \bar{c}$.

$$\frac{1}{m^*} - \frac{e^{(r+\lambda(1-\alpha))T(c)} - 1}{r+\lambda(1-\alpha)} = \frac{B\lambda\alpha}{c(r+\lambda)} - \frac{1}{r} - \frac{e^{(r+\lambda(1-\alpha))T(c)}}{r+\lambda(1-\alpha)} + \frac{1}{r+\lambda(1-\alpha)}.$$

This is decreasing in c if

$$-\frac{B\lambda\alpha}{c^2(r+\lambda)} - \frac{\partial T(c)}{\partial c} e^{(r+\lambda(1-\alpha))T(c)} < 0.$$

By the implicit characterization of $T(c)$ in (3), we get

$$\frac{\partial T(c)}{\partial c} = \frac{\frac{B}{c^2}(1-e^{-rT})}{\frac{Br-c}{c}e^{-rT} - \frac{r+\lambda}{\lambda\alpha}e^{(r+\lambda)T} + e^{\lambda T}}.$$

Hence, we want to show that

$$\frac{\frac{B}{c^2}(1-e^{-rT})}{\frac{Br-c}{c}e^{-rT} - \frac{r+\lambda}{\lambda\alpha}e^{(r+\lambda)T} + e^{\lambda T}} e^{(r+\lambda(1-\alpha))T} > -\frac{B\lambda\alpha}{c^2(r+\lambda)}$$

$$\Leftrightarrow (r+\lambda)(e^{rT}-1)e^{-\lambda\alpha T} < -\frac{Br-c}{c}\lambda\alpha e^{-(r+\lambda)T} + (r+\lambda)e^{rT} - \lambda\alpha$$

$$\Leftrightarrow \frac{Br\lambda\alpha}{c} - \lambda\alpha + (r+\lambda)(e^{rT}-1)e^{(r+\lambda(1-\alpha))T} - (r+\lambda)e^{(2r+\lambda)T} + \lambda\alpha e^{(r+\lambda)T} < 0.$$

From condition (3), we have that

$$\frac{Br\lambda\alpha}{c} - \lambda\alpha = \frac{r}{1-e^{-rT}}\left[(e^{(r+\lambda)T}-1) - \alpha((e^{\lambda T}-1))\right].$$

Inserting this identity into the above inequality yields:

$$\frac{r}{1-e^{-rT}}\left[(e^{(r+\lambda)T}-1) - \alpha((e^{\lambda T}-1))\right] + (r+\lambda)(e^{rT}-1)e^{(r+\lambda(1-\alpha))T} - (r+\lambda)e^{(2r+\lambda)T} + \lambda\alpha e^{(r+\lambda)T} < 0$$

$$\Leftrightarrow r\left[(e^{(2r+\lambda)T} - e^{rT}) - \alpha((e^{(r+\lambda)T} - e^{rT}))\right] + (r+\lambda)(e^{rT}-1)^2 e^{(r+\lambda(1-\alpha))T}$$
$$- (r+\lambda)e^{(2r+\lambda)T}(e^{rT}-1) + \lambda\alpha e^{(r+\lambda)T}(e^{rT}-1) < 0$$

$$\Leftrightarrow r(e^{(r+\lambda)T}-1) - r\alpha(e^{\lambda T}-1) + \lambda\alpha e^{\lambda T}(e^{rT}-1) + (r+\lambda)e^{\lambda(1-\alpha)T}(e^{2rT}$$
$$- 2e^{rT} + 1) - (r+\lambda)e^{(r+\lambda)T}(e^{rT}-1) < 0$$

The factor $-r(e^{\lambda T}-1) + \lambda e^{\lambda T}(e^{rT}-1)$, multiplied by α is positive so that the LHS is smaller than

$$r(e^{(r+\lambda)T}-1) - r1(e^{\lambda T}-1) + \lambda 1 e^{\lambda T}(e^{rT}-1) + (r+\lambda)e^{\lambda(1-\alpha)T}(e^{2rT} - 2e^{rT} + 1) - (r+\lambda)e^{(r+\lambda)T}(e^{rT}-1)$$
$$= (r+\lambda)\left(-e^{\lambda T} + e^{\lambda(1-\alpha)T}(e^{2rT} - 2e^{rT} + 1) - e^{(r+\lambda)T}(e^{rT}-1) + e^{(r+\lambda)T}\right)$$
$$< (r+\lambda)\left(-e^{\lambda T} + e^{(2r+\lambda)T} - 2e^{(r+\lambda)T} + e^{\lambda T} - e^{(2r+\lambda)T} + e^{(r+\lambda)T} + e^{(r+\lambda)T}\right) = 0$$

Hence, the derivative of the LHS in (36) in c is negative and the term itself is positive for all $c < \bar{c}$. □

References

Antinolfi, G. and F. Carli (2015). Costly monitoring, dynamic incentives, and default. *Journal of Economic Theory 159*, 105 – 119.

Avenhaus, R., B. Von Stengel, and S. Zamir (2002). Inspection games. *Handbook of Game Theory with Economic Applications 3*, 1947–1987.

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

Barro, R. J. and D. B. Gordon (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of monetary economics 12*(1), 101–121.

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

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Ben-Porath, E. and M. Kahneman (2003). Communication in repeated games with costly monitoring. *Games and Economic Behavior 44*(2), 227 – 250.

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

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

Border, K. C. and J. Sobel (1987). Samurai accountant: A theory of auditing and plunder. *The Review of Economic Studies 54*(4), 525–540.

Canzoneri, M. B. (1985). Monetary policy games and the role of private information. *The American Economic Review 75*(5), 1056–1070.

Carey, P. and R. Guest (2000). Determining the Optimal External Audit Interval for Private (and Family-Controlled) Companies. *Journal of Accounting, Auditing & Finance 15*(4), 439–458.

Chang, C. (1990). The dynamic structure of optimal debt contracts. *Journal of Economic theory 52*(1), 68–86.

Davis, M. H. (1993). *Markov Models and Optimization*, Volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall.

Delis, M. D. and P. K. Staikouras (2011). Supervisory effectiveness and bank risk. *Review of Finance 15*(3), 511–543.

di Giovanni, J. 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.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The review of economic studies 51*(3), 393–414.

Dilmé, F. and D. F. Garrett (2019). Residual deterrence. *Journal of the European Economic Association 17*(5), 1654–1686.

Dye, R. A. (1986). Optimal monitoring policies in agencies. *The RAND Journal of Economics 17*(3), 339–350.

EU (2016). Directive (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

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.

Gale, D. and M. Hellwig (1985). Incentive-compatible debt contracts: The one-period problem. *The Review of Economic Studies 52*(4), 647–663.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics 10*(1), 74–91.

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

Hughes, J. (1977). Optimal Internal Audit Timing. *The Accounting Review 52*(1), 56–68.

Innes, R. (1999a). Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics 72*(3), 379–393.

Innes, R. (1999b). Self-policing and optimal law enforcement when violator remediation is valuable. *Journal of Political Economy 107*(6), 1305–1325.

Innes, R. (2001). Violator avoidance activities and self-reporting in optimal law enforcement. *Journal of Law, Economics, and Organization 17*(1), 239–256.

Kamada, Y. and N. Rao (2018). Strategies in stochastic continuous-time games. Technical report, Working Paper, Haas School of Business, University of California, Berkeley.

Kaplan, R. S. (1969). Optimal investigation strategies with imperfect information. *Journal of Accounting Research 7*(1), 32–43.

Kaplow, L. and S. Shavell (1994). Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy 102*(3), 583–606.

Kim, S.-H. (2015). Time to Come Clean? Disclosure and Inspection Policies for Green Production. *Operations Research 63*(1), 1–20.

Kitroeff, N., D. Gelles, and J. Nicas (2019). The roots of boeing's 737 max crisis: A regulator relaxes its oversight.

Krasa, S. and A. P. Villamil (2000). Optimal contracts when enforcement is a decision variable. *Econometrica 68*(1), 119–134.

Kreps, D. M. and R. Wilson (1982). Reputation and imperfect information. *Journal of economic theory 27*(2), 253–279.

Lafontaine, F. and M. E. Slade (1996). Retail contracting and costly monitoring: Theory and evidence. *European Economic Review 40*(3-5), 923–932.

Last, G. and A. Brandt (1995). *Marked Point Processes on the real line: the dynamical approach.* Springer.

Levin, J. (2003). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Ljungqvist, L. and T. J. Sargent (2018). *Recursive macroeconomic theory.* MIT press.

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

Mailath, G. J. and L. Samuelson (2006). *Repeated games and reputations: long-run relationships.* Oxford university press.

Makkai, T. and J. Braithwaite (1994). The dialectics of corporate deterrence. *Journal of research in crime and delinquency 31*(4), 347–373.

Malik, A. S. (1993). Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management 24*(3), 241–257.

Martin, R. E. (1988). Franchising and risk management. *The American Economic Review 78*, 954–968.

Monnet, C. and E. Quintin (2005). Optimal contracts in a dynamic costly state verification model. *Economic Theory 26*(4), 867–885.

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

Osaki, S. (2002). *Stochastic Models in Reliability and Maintenance.* Springer.

Pfaff, A. S. and C. W. Sanchirico (2000). Environmental self-auditing: Setting the proper incentives for discovery and correction of environmental harm. *Journal of Law, Economics, and Organization 16*(1), 189–208.

Popov, L. (2016). Stochastic costly state verification and dynamic contracts. *Journal of Economic Dynamics and Control 64*, 1–22.

Ravikumar, B. and Y. Zhang (2012). Optimal auditing and insurance in a dynamic model of tax compliance. *Theoretical Economics 7*(2), 241–282.

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

Toffel, M. W. and J. L. Short (2011). Coming clean and cleaning up: Does voluntary self-reporting indicate effective self-policing? *The Journal of Law and Economics 54*(3), 609–649.

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

Varas, F., I. Marinovic, and A. Skrzypacz (2020). Random inspections and periodic reviews: Optimal dynamic monitoring. *The Review of Economic Studies 87*(6), 2893–2937.

Wang, C. (2005). Dynamic costly state verification. *Economic Theory 25*(4), 887–916.

Webb, D. C. (1992). Two-period financial contracts with private information and costly state verification. *The Quarterly Journal of Economics 107*(3), 1113–1123.

Zhang, Y. (2009). Dynamic contracting with persistent shocks. *Journal of Economic Theory 144*(2), 635–675.