Dynamic Monitoring under Resource Constraints*

Eilon Solan† and Chang Zhao‡

June 26, 2021

Abstract

Often monitoring agencies (principal) do not have enough resources to monitor all agents, and violations are unavoidable. Questions arise regarding the structure of the monitoring scheme that minimizes the rate of violations. In dynamic monitoring problems, the principal can use the past behaviour of agents to determine her monitoring policy. In this paper, we identify the optimal dynamic monitoring scheme when the principal has a commitment power, and show that in this scheme agents first "compete" in a tournament, where the one who is monitored more frequently wins. The winner of the tournament then enjoys lower monitoring intensity, and violates more in the long run.

1 Introduction

Dynamic inspection problems are abundant: managers inspect their workers, regulatory agencies inspect firms, and tax authorities inspect taxpayers. More often than not, regulatory agencies operate under resource constraint: they lack the ability to inspect all agents at all periods.

When the punishment for a revealed violation is sufficiently high, then even if the per-period inspection probability is small, agents will refrain from violating the regulations.

*We thank Kfir Eliaz, Itzhak Gilboa, Yuval Heller, Johannes Hörner, Qingmin Liu, George Mailath, Bernhard von Stengel, and Shmuel Zamir for their helpful suggestions, and seminar audiences at Columbia University, Hebrew University of Jerusalem, LSE, Manchester University, Ohio State University, Tel Aviv University, and University of Pennsylvania. The research of the authors was supported by Global Integrity and the ISF grants #323/13 and #217/17. This document is an output from a project funded by UK Aid from the UK Government for the benefits of developing countries. The views expressed are not necessarily those of UK government's official policies.

†School of mathematical sciences, Tel-Aviv University. Email address: eilons@post.tau.ac.il

‡Institute for Social and Economic Research, Nanjing Audit University. Email address: zhaochangtd@hotmail.com

However, the punishment is usually bounded from above by legal constraints (Harrington, 1988), ethical considerations (Becker, 1968), limited liability (Braithwaite, 1981), preservation of marginal deterrence (Stigler, 1970), or ex post inelasticity of enforcement resources (Bond and Hagerty, 2010). When resources are limited and punishment is not too high, violations are unavoidable, and it is important to identify an optimal inspection policy that minimizes the frequency of violations.

In contrast to the significance of the issue, the theoretical work on the identification of the optimal monitoring policy in dynamic settings is scarce, and existing works have focused on stationary policies or policies that are stationary in some state variables. Landsberger and Meilijson (1982) were the first who allowed the inspection rule to be history dependent: each agent is assigned to one of two states, depending on his action at the last time he was inspected. Greenberg (1984) improves upon this scheme by adding to the model a third state, called the penalty state, and shows that the three-state scheme is optimal for infinitely patient players. The policies that have been studied in the literature ignore a significant portion of the behavior of the agents in past periods, and, as it turns out, are far from optimal when players discount their payoffs. Thus, the question posed by Greenberg (1984) "what is the optimal auditing scheme, in the sense that given the fixed discount rates, no other auditing mechanism will result with a smaller number of individuals who will choose to cheat" remains open.

In line with previous studies, we assume that the inspection agency has a commitment power. We contribute to the literature by providing a complete characterization for the optimal (non-stationary) monitoring policy for a fixed discount factor in the presence of two agents, among which at most one can be inspected in each period. Our results suggest that compared with stationary policies, inspection agencies can use their budget more efficiently to increase compliance among the agents.

At a technical level, to analyze the problem, we write down the recursive equation that describes the optimization problem in terms of the continuation payoffs. Because the resource constraint creates negative correlation between the inspection probability of the agents, this equation is not a Bellman equation, and its solution requires new tools.[1] We characterize the optimal policy and the value of the problem, and provide an iterative algorithm to calculate

[1]For single-agent problems, the recursive approach has been widely used in dynamic contracting problems (early works include Spear and Srivastava, 1987, and Thomas and Worrall, 1990), while we extend it to multi-agent problems. Having multiple agents complicates the analysis for two reasons. First, the recursive characterization of an agent's value function now depends not only on his past performance, but also on other agents' past performance, and hence identifying its solution is more difficult. Second, the recursive formula is not a standard Bellman equation, and the existence of a minimal solution as well as whether an iterative approach can approximate this solution do not follow from standard arguments.

the optimal policy.

It turns out that under the optimal monitoring policy, agents first "compete" in a tournament, in which the agent who is monitored more frequently "wins". The winner of the tournament then enjoys lower monitoring intensity and occasionally violates, while the other agent always adheres. A novel feature of this scheme is that, to win the tournament and to violate in the long run, an agent need not only behave well, but also have good luck: his compliance has to be observed by the principal sufficiently more often than that of the other agent.

Formally, the optimal monitoring scheme consists of two phases. In Phase 1, agents compete in a tournament that can be described as a "token game", where the amount of tokens possessed by an agent corresponds to his expected payoff if adhering in the current period. Figure 1 provides one possible realization of the token game.

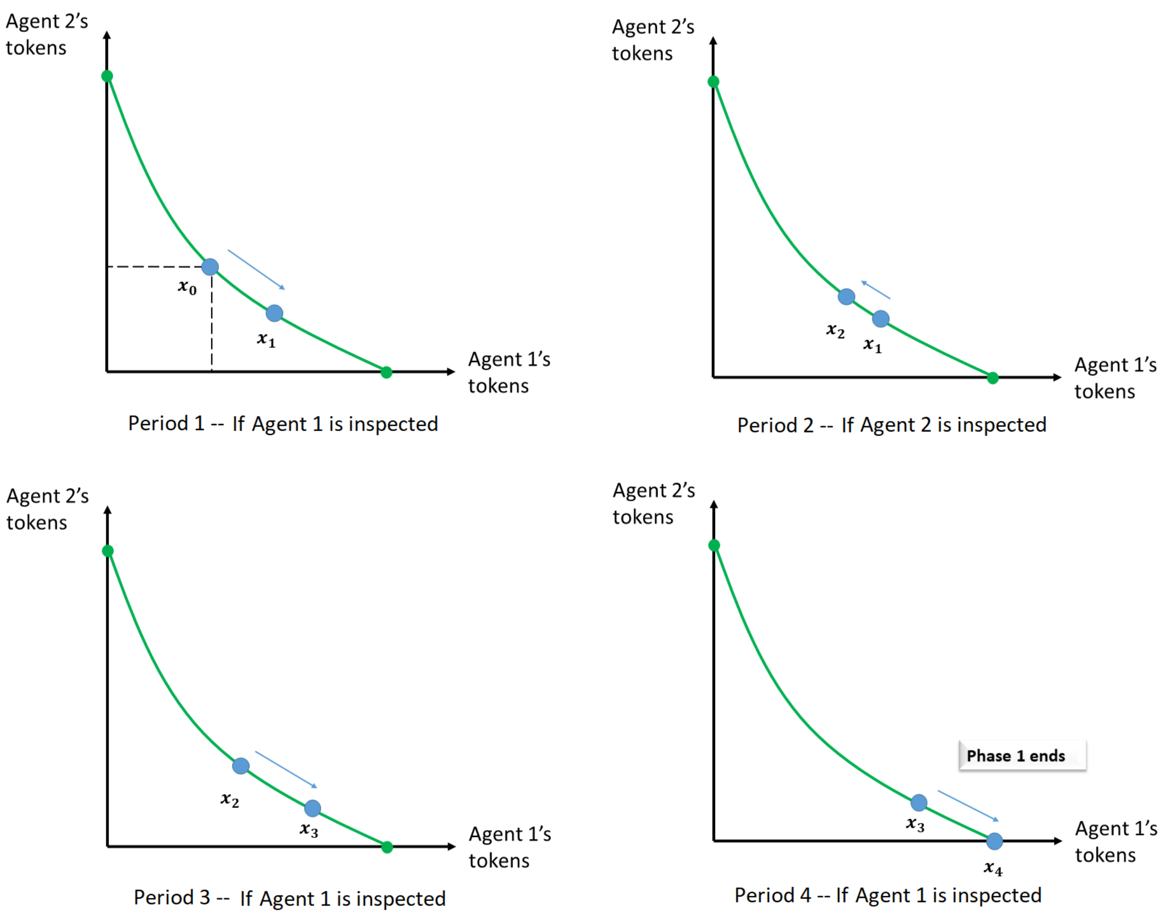

Figure 1: The dynamics in Phase 1 — an example.

Both agents start the game with the same amount of tokens and the principal starts by monitoring each agent with probability 0.5. If an agent is found violating, he loses the tournament right away. Otherwise, the agent who is inspected and found adhering gains tokens, while the agent who is not inspected loses tokens. In the next period, the principal adjusts the inspection probabilities according to the number of tokens each agent holds: the monitoring intensity for an agent is negatively related to the number of tokens he possesses. Then, again, the agent who is inspected and found adhering gains tokens, and the other one loses tokens.[2] This process continues until one of the agents loses all his tokens and hence loses the tournament. Phase 2 starts as soon as a winner is selected. In Phase 2, the winner of the tournament faces lower monitoring intensity and he occasionally violates; whereas the other agent is inspected with high probability and he always adheres. The length of Phase 1 is random, yet it ends almost surely. The updating of the tokens and the monitoring intensities are designed in a specific history-dependent way to ensure that both agents adhere throughout Phase 1.[3]

The model we analyze is stylized, with only two agents, no noise on the principal's observations, and zero inspection cost within the monitoring constraint. While at a conceptual level, our result suggests that when violations are unavoidable, the structure of the two-phase mechanism, a tournament between the agents followed by low monitoring intensities for the winner of the tournament, takes advantage of the tension between the agents, and improves their compliance.

Our leading example is in organizational contexts: A principal has limited attention to spend on monitoring her employees, who have incentives to shirk on their tasks (or violate certain rules, e.g., corruption). While a large literature has been focusing on the optimal design of compensation, there are scenarios like governmental organizations, where employees' salaries are inflexible, and monetary incentives are absent or very low-powered (see e.g., Burgess and Ratto, 2003). Our results imply that in such cases, having the agents "compete" for a position in which they will be treated more favorably can serve as a means to discipline their early performance. In particular, occasional shirking of some senior employees could be a way of the government to motivate junior employees: the government creates a tournament among junior employees, and those who win the tournament get the benefit to enjoy some perks.

[2]Note that the agent with fewer tokens is monitored with a higher intensity, and hence he is more likely to gain tokens in the next period.

[3]As players become more patient, the expected length of Phase 1 increases to infinity and the discounted loss of the principal goes to zero.

Another example concerns inspection agencies, like the health service department, which inspects restaurants for hygiene levels; or environmental protection agencies, which inspect firms that produce air or water pollution. In the environmental regulation for polluting firms, for instance, it has been well documented that strict monetary penalty is absent (see e.g., Harrington, 1988, and Harford, 1991), the compliance rate is substantially less than full,[4] and hence a more efficient monitoring scheme is in need.

Our results have several practical implications. First, to minimize her loss, the principal should allow the agents to violate in pre-specified stages (the "rewarding stages"), which are strategically allocated to reduce the overall number of violations. In the rewarding stages, no inspection takes place. The reason is that a positive (even if small) probability of inspection in a rewarding stage hurts the agent without benefiting the principal. This in turn weakens the agent's incentive to adhere in previous periods. In practice, it is observed that greater compliance typically leads to less enforcement,[5] but there is no evidence that the regulatory attention periodically drops to 0. Our result suggests that compliance can be significantly improved by employing an inspection scheme with occasional consent to violations.

A second insight is that the rewards to an agent should be delayed as much as possible, with the exception of histories where the other agent's payoff drops down to zero. The backloading of rewards, in fact, is a common feature of dynamic environments without transfers. To elaborate on this point, notice that in a given period, inspection probabilities and future rewards are complimentary in deterring an agent from violating: an agent who violates faces the risk of being detected and thereby punished and losing all his potential gains. Therefore, a lower inspection probability is needed to discipline an agent with a larger continuation payoff. When an agent is eligible for a certain amount of rewards, delaying the reward from today to some later period t implies that a higher reward is at stake for the agent in periods before t, and hence lower inspection intensities are needed to discipline him in these periods. This "frees up" inspection resources, which can be used to inspect the other agent, and reduces overall violations.

Third, the updating of an agent's continuation payoff depends not only on his own actions, but also on luck: whether he was inspected and his behavior observed. Indeed, the agent's

[4]One well-known early study conducted by the White House Council on Environmental Quality estimated 65% of regulated sources may be in violation of air pollution emission limits (Russell, 1990, p.255). In the United Kingdom, published compliance rates with many key water quality standards are sometimes as low as 50% (Heyes, 2000). More recently, the data published by US Environmental Protection Agency (EPA) shows that in 2009, only 58% of major facilities fully comply with the Clean Air Act (EPA, 2010).

[5]See, e.g., Dubin and Wilde (1988) and Gray and Deily (1996).

continuation payoff if he is inspected and found adhering (denoted v_i^A) determines whether the agent adheres or violates in the current period; whereas the agent's continuation payoff if he is not monitored (denoted v_i^{NI}) is the same regardless of his current period action, and it does not affect the agent's incentive. Since the (weighted) average of v_i^A and v_i^{NI} must be consistent with previous promised rewards, the principal is better off reducing the latter and increasing the former. Consequently, the continuation payoff of an agent who is inspected more often in early periods is high, and he will violate more in the future. As a result, an agent's probability of being inspected decreases following an inspection that reveals that the agent adhered.

Forth, even though both agents behave identically along the first phase, eventually the agent who is inspected more often in early periods gets all the reward, while the other agent never gets the chance to violate. That is, even though both agents are eligible for a certain amount of rewards in Phase 1, the crediting of the rewards is not deterministic: It takes the form of a "lottery" that gives only the winner the privilege to violate.

Related literature: Our paper is closely related to the literature on dynamic auditing. Auditing rules in which individuals' past compliance history affects the future inspection probability are sometimes referred to as *conditional future audit rules* in this literature, and their theoretical analysis is rather scarce. Landsberger and Meilijson (1982) are the first who allow the inspection scheme to be history dependent, and show that such scheme outperforms any static mechanism. Greenberg (1984) improves upon this scheme by adding a third state into the model, and showing that the three-state scheme is optimal for infinitely patient players. Harrington (1988), Harford and Harrington (1991), and Harford (1991) adapt these simple auditing mechanisms to environmental control problems, thereby explain the phenomenon of high compliance in the absence of strict enforcement. Previous works on dynamic auditing have focused either on simple stationary policies, or policies that are optimal only when players are infinitely patient. We contribute to this literature by identifying the optimal (non-stationary) monitoring scheme for a fixed discount factor.

Our paper is also related to the literature on dynamic contracting without transfers (see, e.g., Fong, 2009, Li et al., 2017, Guo and Hörner, 2018, and Lipnowski and Ramos, 2018). All these works focus on single-agent problems, and their conclusions can be applied to multi-agent problems in which it is optimal to treat each agent separately.[6] In these works, under the optimal policy, the agent is assigned a "score", which is updated according to

[6] This is the case when, for instance, the principal faces no resource constraint and the cost for monitoring is linear in the number of inspections.

his past performance. Once the score reaches either a lower bound or an upper bound, the continuation policy adopts the form of either indefinite punishment or indefinite reward.

In contrast to previous works, we study the game with multiple agents and under limited monitoring resources. A tension between the agents, which is absent in the single-agent problem, is created. In our optimal scheme, the future payoff of an agent depends not only on his own performance, but also on the performance of the other agent. In contrast, in the single-agent problem, an agent's payoff depends only on his own behavior.

Lazear (2006), and Eeckhout et al. (2010) also study optimal monitoring under limited resources, but these works focus on static models. The feature of competition between agents has some similarities to Carrasco et al. (2019), which studies a problem where two privately informed players repeatedly take a joint action, and the optimal dynamic incentives imply that one of the agents becomes a dictator in the long run. In a recent paper, Varas et al. (2020) consider a dynamic model of inspections that involves a single agent, and identify the optimal dynamic monitoring scheme.

Another related literature concerns dynamic allocations. de Clippel et al. (2021) studies a model with one principal and several agents. Each agent generates a new idea in each period, which is of either good or bad quality, and decides whether to propose this idea to the principal. An agent always prefers his idea to be implemented, regardless of the quality. Like in our model, the principal has limited attention, hence she cannot evaluate all proposals. The goal is to design a mechanism without transfers that achieves the first-best outcome, where agents only report good quality ideas. In contrast, in our paper we focus on the case where the first-best outcome cannot be achieved.

Allowing monetary transfers, Board (2011) and Andrews and Barron (2016) study problems where a principal (firm) dynamically allocates business among several agents (suppliers). Acknowledging the difficulty of fully characterizing the optimal allocation rule, Board (2011) derives a number of economically interesting results. In particular, it is optimal for the principal to be "loyal" to a subset of "insider" agents, with whom she has previously traded. One major difference between the selection of insiders in Board (2011) and the selection of the lucky agent in our paper is that, in the two-agent version of Board (2011), the tournament ends after the first period; whereas in our model, the tournament lasts for multiple periods and its length is randomly determined.[7] Andrews and Barron (2016) relax Board's liquidity

[7]In Board (2011), the principal can always lower an agent's payoff to zero by not trading with him. Therefore, after the first period, rewarding the "insider" while yielding the "outsider" a continuation payoff zero is optimal. This stands in sharp contrast with our model, where lowering an agent's continuation payoff to zero requires a high monitoring intensity, which may not be feasible given the constraint on the other agent's payoff. Therefore, a key step in the analysis is the study of the optimal way to discipline an agent

constraint on the agents, and introduce imperfect monitoring to the moral hazard problem. They also focus on the first-best outcome and characterize a dynamic allocation rule that attains it.

The paper is organized as follows. The model is presented in Section 2 and the main results are described in Section 3. Comments and extensions appear in Section 4. Proofs are relegated to the Appendix.

2 Model

2.1 Participants and their action sets

We study a repeated inspection game between a *principal* and two *agents*. At every period, the principal inspects (at most) one of the agents and each agent decides whether to Adhere or to Violate.

The action set of agent i, $i = 1, 2$, is $\mathscr{A}_i := \{A, V\}$, and the action set of the principal is $\mathscr{A}_0 := \{I_1, I_2, \emptyset\}$, where $\emptyset$ stands for no inspection, and I_1 (resp. I_2) for inspecting Agent 1 (resp. Agent 2). Throughout the paper, whenever a variable refers to the principal we add the subscript 0, and whenever it refers to agent i we add the subscript i.

2.2 The stage payoff

The gain of each agent from adhering is normalized to 0, and his gain from an undetected violation is normalized to 1. The agent's loss from a detected violation is denoted by $c > 0$. The value of c is exogenously given, say, by legal constraints.[8] Consequently, the stage payoff function of agent i, denoted u_i, is given by:

$$u_i(a) = \begin{cases} 0 & \text{if } a_i = A, \\ -c & \text{if } a_i = V \text{ and } a_0 = I_i, \\ 1 & \text{if } a_i = V \text{ and } a_0 \neq I_i, \end{cases} \tag{1}$$

where $a = (a_0, a_1, a_2)$ is the vector of actions played at the current stage. The principal loses 1 for each violation, detected or undetected. Therefore, the principal's stage *loss* function,

when the other agent is subject to a certain amount of reward.

[8] In Section 4.3 we relax this assumption and allow the principal to choose a punishment level within a certain range. We will argue that the principal is better off always choosing the most severe punishment and hence our results remain intact.

denoted u_0, is given by:

$$u_0(a) = \begin{cases} 2 & \text{if } a_1 = a_2 = V, \\ 1 & \text{if } a_i = A \text{ and } a_j = V \text{ for } i \neq j, \\ 0 & \text{if } a_1 = a_2 = A. \end{cases} \tag{2}$$

The form of the principal's loss function implies that we focus on the principal's incentive to deter violations, rather than on her incentive to collect the penalties. There are several interpretations for this assumption. First, in the absence of monetary penalties, the punishment can only take the form of, say, public reprimand, which hurts the agent but does not benefit the principal. Second, even if the punishment takes the form of monetary penalties, it may not be comparable with the damage caused by the violation behaviour, as they are in different dimensions. For instance, in the environmental control problem, the damage caused by a marine oil spill is difficult to evaluate with monetary payment.

Eq. (2) also implies that inspections are costless for the principal, and the only restriction is on the number of inspections that can be conducted simultaneously. This assumption fits scenarios where the principal has a fixed number of inspectors at her disposal, and the hiring of additional inspectors is impractical either because of budget constraint or lack of positions.[9]

In the one-shot game, an agent will adhere (resp. violate) as long as the probability to be inspected is larger (resp. smaller) than $\frac{1}{1+c}$. Therefore, if $c > 1$, the principal can deter both agents from violating by inspecting each agent with probability $\frac{1}{2}$. This observation implies that when the fine for violation is sufficiently large, the principal can completely deter violations. In the rest of the paper we will focus on the nontrivial case $c < 1$.

Assumption 1. $0 < c < 1$.

When c is less than 1, in the one-shot game an agent will adhere only if the inspection probability at the current period is larger than $\frac{1}{1+c} > \frac{1}{2}$. The best the principal can do in the one-shot game is to deter one agent from violating, while allowing the other agent to violate. This asymmetric strategy outperforms the symmetric strategy that inspects each agent with probability $\frac{1}{2}$, and does not deter any of the agents from violating.

[9]In a model where the principal faces no resource constraint and there is a fixed cost for each inspection, it is optimal to treat each agent independently, and the problem is reduced to a single-agent problem.

2.3 Monitoring structure and histories

In the repeated game, at the end of each period, the two agents observe the principal's action. If the principal monitors one of the agents at that period, the action of the monitored agent is publicly observed. This is equivalent to the situation in which the players observe at the end of each period a public signal y, drawn from the set $Y = \{V_1, A_1, V_2, A_2, \emptyset\}$, with the interpretation that the signal $\emptyset$ means that no agent was inspected, and the signal A_i (resp. V_i) means that agent i was inspected and found adhering (resp. violating). In Remark 3 we discuss the private monitoring case, where the actions of the principal and of the inspected agent are not observed by any uninspected agent.

We assume that the players have at their disposal a public correlation device, which outputs at the beginning of every period t a random signal ζ^t that is uniformly distributed on the interval $[0, 1]$ and independent of past history.

In the repeated game, the only public information available in period t is the $(t-1)$-period history of public signals (which contains the elements of Y and the outcomes of the correlation device), denoted h^t. The set of finite-length *public histories* is denoted by $\mathscr{H}$.

Remark 1. *The public correlation device is crucial in some parts of the analysis. On the equilibrium path, however, the use of the correlation device is minimal (see Remark 2 below) and its impact becomes negligible when players are patient.*

2.4 Strategies

A strategy of the principal is a function from the set of public histories to the set of her mixed actions. Denoting by $\Delta\mathscr{A}_0$ the set of mixed actions of the principal, a strategy of the principal is a function

$$\sigma_0 : \mathscr{H} \to \Delta\mathscr{A}_0.$$

Denote by $\mathscr{B}_0$ the set of strategies of the principal in the infinite repeated game.

It is assumed that the principal announces her entire inspection strategy at the beginning of the game and she is able to commit to it. The commitment power enables the principal to achieve a higher payoff compared with the case where such a commitment is absent, and therefore the principal is always better off committing to a strategy, if possible.

Since the principal has a commitment power, her entire strategy is known to the agents from the outset of the game. A *public strategy* for an agent is a strategy that assigns a mixed action to every strategy of the principal and every *public history.* Without loss of generality, we assume that the agents use only public strategies.[10] Formally, for $i = 1, 2$, a strategy of agent i is a function

$$\sigma_i : \mathscr{B}_0 \times \mathscr{H} \to \Delta\mathscr{A}_i,$$

where $\Delta\mathscr{A}_i$ is the set of mixed actions of agent i.

2.5 Payoffs

A *play* is an infinite sequence of action profiles $\mathbf{a} \equiv (a^1, a^2, a^3 ...) \in \mathscr{A}^{\mathbb{N}}$, where $\mathscr{A} = \mathscr{A}_0 \times \mathscr{A}_1 \times \mathscr{A}_2$ is the set of all action profiles.

Every vector of players' strategies $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ induces a probability distribution $\mathbf{P}_\sigma$ over the set of infinite plays. We denote by $\mathbf{E}_\sigma$ the expectation operator that corresponds to this probability distribution. Agent i's discounted payoff under strategy vector σ is

$$v_i(\sigma) := \mathbf{E}_\sigma \left[\sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t) \right], \tag{3}$$

where δ is the common discount factor. The principal's discounted loss $v_0(\sigma)$ is similarly defined.

Note that the agents know their stage-game payoff, while the principal does not. This is because the principal knows only the action of the inspected agent, and she is therefore unaware of the damage inflicted on her by the uninspected agent.

2.6 Equilibrium

Denote the subgame following the announcement of an inspection strategy σ_0 by $\Gamma(\sigma_0)$. We study *agent perfect public equilibria* (agent PPE) of the subgame $\Gamma(\sigma_0)$.

Given a player i, for each history h^t and each strategy σ_i, player i's *continuation strategy* given history h^t, denoted $\sigma_i|_{h^t}$, is defined by

$$\sigma_0|_{h^t}(h^\tau) := \sigma_0(h^t h^\tau), \quad \forall h^\tau \in \mathscr{H}.$$

The pair of strategies (σ_1, σ_2) is an agent PPE of the subgame $\Gamma(\sigma_0)$ if for every public history $h^t \in \mathscr{H}$, the pair of strategies $(\sigma_1|_{h^t}, \sigma_2|_{h^t})$ is an agent Nash equilibrium of $\Gamma(\sigma_0|_{h^t})$.

[10]Conditional on the public history, each agent's private information (that is, his uninspected past actions) is independent from the private information of the other agent. Therefore, for any equilibrium strategy in which agents use their private information, we can find a public-strategy equilibrium which yields the same outcome.

We will often abuse the terminology and call the triplet $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ a PPE if (σ_1, σ_2) is an agent PPE of $\Gamma(\sigma_0)$. Denote by $\mathscr{E}$ the set of all PPEs $\sigma = (\sigma_0, \sigma_1, \sigma_2)$. Our goal is to identify a PPE σ^* that yields the principal the minimum loss:

$$\sigma^* \in \operatorname*{argmin}_{\sigma \in \mathscr{E}} v_0(\sigma). \tag{4}$$

When the agents are very impatient, the principal will not be able to deter more violations than in the one-shot game. We therefore focus only on relatively patient agents. In particular, we assume that $\delta \geq \frac{1-c^2}{1-c^2+c}$ (see Section 4.5 for a detailed discussion).

Assumption 2. $\delta \geq \frac{1-c^2}{1-c^2+c}$.

3 Main Results

3.1 No violation-free inspection strategy

Our first observation is that regardless of the inspection strategy adopted, at least one agent violates infinitely often.

Proposition 1. *For every PPE, with probability 1 the number of violations along the play is infinite.*

Proof. Fix a PPE $\sigma = (\sigma_0, \sigma_1, \sigma_2)$. We first prove that at least one agent violates on the equilibrium path. Since $c < 1$, one of the agents is inspected with probability $p < \frac{1}{1+c}$ in the first period. It follows that $p \cdot (-c) + (1-p) \cdot 1 > 0$. Therefore, this agent can guarantee a positive expected payoff, by violating in the first period and adhering in all subsequent periods. This implies that in every equilibrium, at least one player obtains a positive payoff by violating.

If (σ_1, σ_2) is an agent PPE of $\Gamma(\sigma_0)$, then $(\sigma_1, \sigma_2)|_{h^t}$ is an agent PPE of $\Gamma(\sigma_0|_{h^t})$. This observation, together with the previous result, imply that in any PPE, after every history, at least one of the agents obtains a positive payoff. Proposition 1 follows. □

The practical implication of Proposition 1 is that, the goal of the principal is not to deter violations completely, but rather to manage violations optimally; that is, to make the violations as rare as possible.

3.2 Reducing the set of PPEs

In this section, we show that to solve the principal's optimization problem (4), it is sufficient to consider a subset of all PPEs. Specifically, we will show that we can restrict attention to PPEs that satisfy the following conditions:
(i) The agents play pure strategies and whenever an agent is indifferent between adhering and violating, he adheres.
(ii) On the equilibrium path the principal never inspects an agent who violates. Equivalently, the loss of the principal equals the sum of payoffs of the agents.
(iii) Whenever an agent is inspected and found violating, he is punished most severely by being inspected with probability 1 in all future periods.

Formally, for every strategy σ_i of player $i \in \{0, 1, 2\}$ and every action $a_i \in \mathscr{A}_i$, denote by $\sigma_i(a_i|h^t)$ the conditional probability that player i plays action a_i at history h^t, where the initial history is h^0.

Proposition 2. *Let σ be a PPE. Then there exists a PPE, denoted $\widehat{\sigma} = (\widehat{\sigma}_0, \widehat{\sigma}_1, \widehat{\sigma}_2)$, which satisfies $v_0(\widehat{\sigma}) = v_0(\sigma)$ and the following conditions:*
(i) $\widehat{\sigma}_1$ and $\widehat{\sigma}_2$ are pure strategies. If an agent i is indifferent between adhering and violating at some history h^t, then $\widehat{\sigma}_i(A|h^t) = 1$.
(ii) $v_0(\widehat{\sigma}) = v_1(\widehat{\sigma}) + v_2(\widehat{\sigma})$.
(iii) For every finite history h^t that occurs with positive probability, $v_i(\widehat{\sigma}|_{(h^t,V_i)}) = 0$ and $\widehat{\sigma}_0(I_i|(h^t, V_i, h^\tau)) = 1$ for every public history $h^\tau \in \mathscr{H}$.

Proof. See Appendix A.1. □

Proposition 2 is technically useful, as it restricts the set of PPEs that we should consider. Moreover, the proposition has significant practical implications regarding the structure of the optimal inspection scheme. First, whenever an agent is supposed to violate in equilibrium, he knows that he will not be inspected. Therefore, the principal will never detect violations, and any detected violation implies that the inspected agent deviated. This further implies that, an agent who is inspected and found violating is punished in the most severe way.

Part (i) of Proposition 2 holds because the correlation device can be used to mimic lotteries performed by the agents. To illustrate Part (ii), recall that by Eqs. (1)–(2), every gain of 1 unit by an agent is the loss of 1 unit of the principal, and in addition, any violation that is inspected reduces the gain of the violating agent without reducing the loss of the principal. Hence, for every strategy profile σ, the principal's loss is no less than the sum of payoffs of the two agents: $v_0(\sigma) \geq v_1(\sigma) + v_2(\sigma)$, with strict inequality if some violations are observed by the principal with positive probability. According to Proposition 1, in every PPE, violations occur infinitely often. Inspecting an agent with positive probability in a stage in which the agent's equilibrium action is Violate hurts the agent without benefiting the principal. Thus, as stated in Part (ii), there is an alternative inspection strategy that yields the principal the same loss and does not inspect the agents in those stages in which they violate. For PPEs that satisfy Part (ii), violations are not inspected on the equilibrium path, and therefore any detected violation implies that the inspected agent deviated. This allows us to reduce further the set of PPEs that should be considered in the minimization problem (4) by assuming that an inspected agent who is found violating is punished in the most severe way.

Denote by $\mathscr{E}_*$ the set of all PPE profiles that satisfy Proposition 2. Under every PPE $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ in $\mathscr{E}_*$, an agent violates if and only if he is inspected with probability 0, hence in the subgame $\Gamma(\sigma_0)$, the agent PPE in $\mathscr{E}_*$ is uniquely determined. This observation is summarized in the following proposition.

Proposition 3. *Suppose that $\sigma = (\sigma_0, \sigma_1, \sigma_2) \in \mathscr{E}_*$. Then in the subgame $\Gamma(\sigma_0)$, the strategy pair (σ_1, σ_2) is the unique agent PPE in $\mathscr{E}_*$.*

We next provide an example for an inspection scheme σ_0 and its corresponding agent PPE.

Example 3.1. Consider the inspection strategy σ_0, where the principal always inspects Agent 1 and never inspects Agent 2. Agent 1's best response, denoted σ_1, is to always adhere, while Agent 2's best response, denoted σ_2, is to always violate. The reader can verify that the strategy vector $\sigma^{P,1} := (\sigma_0, \sigma_1, \sigma_2)$ is in fact a PPE in $\mathscr{E}_*$, which yields Agent 1 a payoff 0 and Agent 2 a payoff $\frac{1}{1-\delta}$. The analogous PPE with the roles of the two agents exchanged, denoted $\sigma^{P,2}$, yields Agent 1 a payoff $\frac{1}{1-\delta}$ and Agent 2 a payoff 0.

3.3 Simplifying the incentive compatibility constraint

We turn to analyze the means to deter violations in the repeated game. As argued before, in the one-shot game, an agent violates as long as the inspection probability for him is less than $\frac{1}{1+c}$. In contrast, in the repeated game, taking into account the effect of his current-period action on future payoffs, an agent who is inspected with a probability less than $\frac{1}{1+c}$ may well adhere.

In a certain period where Agent 1 is inspected with probability $p < \frac{1}{1+c}$, if Agent 1 violates, he obtains

$$v_1(V_1) = \big((1-p) - c \cdot p\big) + \delta \cdot \Big(p \cdot v_1^{V_1} + (1-p) v_1^{NI}\Big), \tag{5}$$

where $v_1^{V_1}$ (resp v_1^{NI}) is Agent 1's continuation payoff if he is inspected an found violating (resp. if he is not inspected). The first term of the RHS of Eq. (5) represents Agent 1's first period gain from violation, and the second term represents his future payoffs, depending on whether Agent 1 is inspected or not in the current period. If Agent 1 adheres, he obtains

$$v_1(A_1) = 0 + \delta \cdot \Big(p \cdot v_1^{A_1} + (1-p) v_1^{NI}\Big),$$

where $v_1^{A_1}$ is Agent 1's continuation payoff if he is inspected an found adhering.

By adhering rather than violating, Agent 1 loses in the current period, but he gains in future periods if his adhering behaviour is observed by the principal. If the latter effect is sufficiently high, that is, the agent's continuation payoff if being found adhering is sufficiently higher than his continuation payoff if being found violating (the latter value is zero for PPEs in $\mathscr{E}_*$), then the agent is better off adhering. As the next proposition shows, the following function $f : (0, 1] \to \mathbb{R}_+$ measures this difference (see Figure 2):

$$f(p) := \begin{cases} \frac{1-(1+c)p}{\delta p} & \text{if } 0 < p < \frac{1}{1+c}, \\ 0 & \text{if } \frac{1}{1+c} \leq p \leq 1. \end{cases} \tag{6}$$

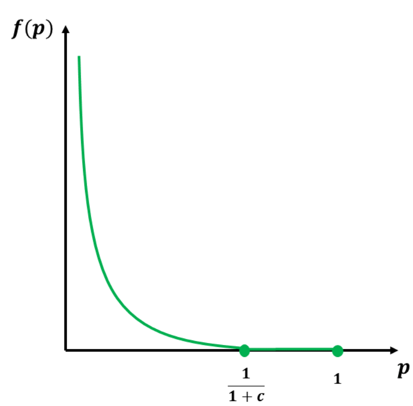

Figure 2: The function $f(p)$.

Proposition 4. *Suppose that σ is a PPE in $\mathscr{E}_*$. Then for every history h^t that occurs with positive probability, agent i adheres if and only if $v_i(\sigma|_{(h^t,A_i)}) \geq f\big(\sigma_0(I_i|h^t)\big)$.*

Proof. See Appendix A.2. □

Proposition 4 asserts that, an agent who is inspected with probability p adheres if and only if his continuation payoff upon *being inspected and found adhering* is at least $f(p)$. Thus, $f(p)$ is the minimal rewards to an agent who is inspected with probability p so that he adheres.

This result has two implications. First, the fact that f is non-increasing reflects the property that in a given period, inspection probabilities and future rewards are complimentary in deterring an agent from violating. Indeed, the higher the probability of inspection, the less attractive it is to violate, hence a lower compensation is needed to ensure that the agent adheres. On the other hand, since an agent who violates faces the risk of losing all his potential gains, when an agent faces a larger amount of rewards, a lower inspection probability is needed to discipline him. A second implication is that, the optimal response of an agent in a given period depends only on his continuation payoff if he is inspected in that period. In contrast, the agent's continuation payoff *if he is not inspected*, which is the same regardless of his current-period action, does not affect the agent's incentive compatibility constraint.

3.4 Disciplining one agent

In this section, as a benchmark, we present the optimal way to discipline one agent while ensuring the payoff of the other agent is 0. This result will be used in the next section for the construction of the optimal inspection strategy.

To ensure that Agent 2's payoff is 0, he should be inspected in each period with probability at least $\frac{1}{1+c}$. Consequently, Agent 1 is inspected with probability at most $\frac{c}{1+c}$ in each period. By Proposition 4, an agent who is inspected with probability $\frac{c}{1+c}$ adheres only when his continuation payoff in the eventuality that he is inspected and found adhering is at least $f(\frac{c}{1+c})$. By Assumption 2, $f(\frac{c}{1+c}) \leq \frac{1}{1-\delta}$, where $\frac{1}{1-\delta}$ is the maximal payoff to an agent in the game. This guarantees that an agent who is inspected with probability as low as $\frac{c}{1+c}$ is willing to adhere, so long as he is allowed to violate sufficiently often in future periods.[11] The next proposition shows that in this case, the optimal PPE yields Agent 1 a payoff $\frac{1-c^2}{1+c-\delta}$, which is lower than $\frac{1}{1-\delta}$.

[11] If, instead, Agent 1 is impatient and $f(\frac{c}{1+c}) > \frac{1}{1-\delta}$, then even the highest promise for the future reward cannot deter Agent 1 from violating, and in any PPE in which Agent 2's payoff is 0, Agent 1 will violate in all stages.

Proposition 5. *Among all PPEs in $\mathscr{E}_*$ that yield Agent 2 the payoff 0, the minimal payoff to Agent 1 is $\frac{1-c^2}{1+c-\delta}$.*

Proof. See Appendix A.3. □

Given that Agent 2 obtains 0, the optimal way to discipline Agent 1 is not unique. Here we identify a *periodic rewarding scheme* that contains cycles of length d. There exist other inspection strategies that induce the same payoff. One of them, which has a stationary flavor, is described in Appendix A.4.

Roughly speaking, the *periodic rewarding scheme* is composed of cycles of length d. In the first $d-1$ periods of each cycle, Agent 1 is inspected with probability $\frac{c}{1+c}$. If ever he is caught violating, Agent 1 is punished and will be inspected in every future period with probability 1. While if he is not found violating during the $d-1$ periods, Agent 1 is "rewarded" by not being inspected in the dth period of the cycle, and he will violate in that period. If d is not too large, the optimal responses of Agent 1 is to adhere in the first $d-1$ periods of each cycle and violate only in the dth period. While keeping Agent 1's incentive compatibility constraints (that is, Agent 1 is willing to adhere in the first $d-1$ periods in each cycle), the principal chooses d as large as possible, so as to minimize the rate of violations.

Remark 2. *Denote the optimal choice of the cycle length by d^*. If d^* is not an integer, for example, $d^* = 8.2$ when $c = 0.7$ and $\delta = 0.9$, then the length of the cycle is random and is determined at the end of the 7th period of the cycle: with some probability the 8th period is the rewarding period and the cycle has length 8, and with the remaining probability the 9th period is the rewarding period and the cycle has length 9. When players are rather patient, d^* is large, and the impact of the public correlation device becomes negligible. As will be shown in the next section, when using the cyclic rewarding scheme to construct the optimal inspection mechanism, this is the only role of the public correlation device.*

Naturally, by reducing the length of the rewarding cycles in $\sigma(d)$ from the optimal one (hence increasing the rewarding intensity), Agent 1's incentive compatibility constraints remain kept. Therefore, while keeping Agent 2's payoff at 0, every $z \in [\frac{1-c^2}{1+c-\delta}, \frac{1}{1-\delta}]$ can be supported as an Agent 1's PPE payoff. This is the content of the next proposition.

Proposition 6. *For every $z \in [\frac{1-c^2}{1+c-\delta}, \frac{1}{1-\delta}]$, there exists a PPE that yields Agent 1 and Agent 2 the payoffs z and 0, respectively.*

Proof. See Appendix A.5. □

3.5 Optimal PPE

3.5.1 Optimal PPE payoff

We are now ready to present the optimal PPE in the presence of two agents. To this end, it is crucial to identify the boundary of the PPE payoff set. For every $v_1 \in [0, \frac{1}{1-\delta}]$, denote by $g(v_1)$ the minimal payoff that Agent 2 gets, where the minimum is over all PPE outcomes in $\mathscr{E}_*$ that yield Agent 1 the payoff v_1.[12] By standard continuity arguments, $g(v_1)$ is attained by some PPE in $\mathscr{E}_*$.[13] Figure 3 provides a graphical illustration of g. As we will show in Appendix A.6, g is non-negative, non-increasing, and continuous. Moreover, because of the use of the correlation device, g is convex. By the symmetry between the agents, this function further satisfies $g\big(g(x)\big) = x$ for every $x \in (0, \frac{1-c^2}{1+c-\delta})$. By Proposition 5, g intersects with the x-axis (resp. y-axis) at $(\frac{1-c^2}{1+c-\delta}, 0)$ (resp. $(0, \frac{1-c^2}{1+c-\delta})$).

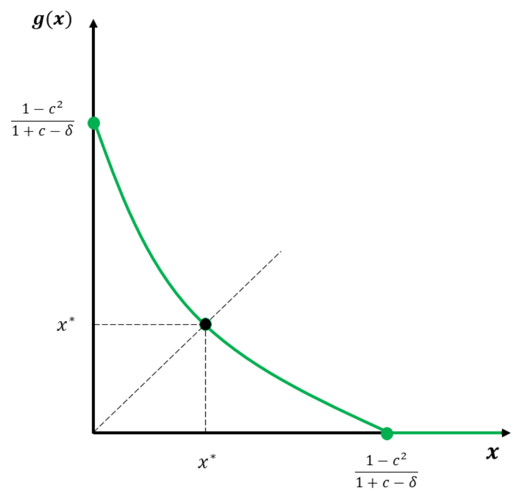

Figure 3: The function g.

Since we consider PPEs in $\mathscr{E}_*$, when Agent 1's payoff is x and Agent 2's payoff is $g(x)$, the principal's loss is $x + g(x)$. Consequently, the optimal payoff of the principal is the minimum

[12]The minimal payoff of Agent 1 in a PPE is 0 and is obtained by the strategy profile $\sigma^{P,1}$ defined in Example 3.1. The maximal payoff of Agent 1 in a PPE is $\frac{1}{1-\delta}$ and is obtained by the strategy profile $\sigma^{P,2}$ defined in Example 3.1. Using the correlation device to choose one of these two PPEs at the outset of the game, we deduce that the set of possible PPE payoffs of Agent 1 is $[0, \frac{1}{1-\delta}]$.

[13]Because a limit of equilibria in $\mathscr{E}_*$ is an equilibrium in $\mathscr{E}_*$, the minimum in the definition of $g(x)$ is attained. Indeed, it is standard (though a little bit tedious because of the presence of the correlation device) to show that any sequence $\{\sigma^k\}_{k\in\mathbb{N}}$ of PPEs has an accumulation point $\widehat{\sigma}$. Since the payoffs are discounted, $\widehat{\sigma}$ is a PPE. We argue that $\widehat{\sigma}$ is in $\mathscr{E}_*$. By continuity, $\widehat{\sigma}$ satisfies Parts (ii) and (iii) of Proposition 2. To see that $\widehat{\sigma}$ satisfies Part (i), note that if $\widehat{\sigma}_i(V|h) = 1$ for some agent i, then for every k sufficiently large we have $\sigma_i^k(V|h) = 1$, and by Proposition 2(ii) it implies that $\sigma_0^k(I_i|h) = 0$, and hence $\widehat{\sigma}_0(I_i|h) = 0$. Therefore, agent i strictly prefers violating at the history h under $\widehat{\sigma}$.

of $x + g(x)$. By the properties of g (see Figure 3), the minimum is attained at the point x^*, where $g(x^*) = x^*$. The next Theorem provides a recursive characterization of g, which uniquely determines this function.[14]

Theorem 1. *The function g is the minimal function that satisfies*

$$g(x) = \delta \cdot w\left(1 - w^{-1}\left(\frac{x}{\delta}\right)\right) \text{ for every } x \in \left(0, \frac{1-c^2}{1+c-\delta}\right), \tag{7}$$

where

$$w(p) := \begin{cases} p \cdot f(p) + (1-p) \cdot g\big(f(1-p)\big) & \text{if } 0 < p \leq 1, \\ \frac{1}{\delta} + g(0) & \text{if } p = 0. \end{cases} \tag{8}$$

Proof. See Appendix A.7. □

As will be shown in the next section, the description of the optimal PPE heavily relies on Theorem 1, hence we explain the derivation of Eqs. (7) and (8). The recursive representation of g in Eq. (7) is equivalent to the following characterization. Given that Agent 1 obtains a payoff x, the minimal PPE payoff of Agent 2 is attained by:

$$\begin{aligned} & g(x) = \min_{p_1} \delta \cdot \left[p_1 \cdot v_2^{A_1} + (1-p_1) \cdot v_2^{A_2}\right] \\ s.t. \quad & x = \delta \cdot \left[p_1 \cdot v_1^{A_1} + (1-p_1) \cdot v_1^{A_2}\right] \\ & (v_1^{A_1}, v_2^{A_1}) = \Big(f(p_1), g\big(f(p_1)\big)\Big), \\ & (v_1^{A_2}, v_2^{A_2}) = \Big(g\big(f(1-p_1)\big), f(1-p_1)\Big), \end{aligned} \tag{9}$$

where $(v_1^{A_1}, v_2^{A_1})$ (resp. $(v_1^{A_2}, v_2^{A_2})$) is the continuation payoff vector of the agents if Agent 1 (resp. Agent 2) is inspected in the first period and found adhering.

The term in the minimization is the expected continuation payoff, discounted by one period. This representation is equivalent to saying that to implement the payoff vector $\big(x, g(x)\big)$, Agent 2 adheres in the first period. The first condition ensures that Agent 1's current-period payoff agrees with his continuation payoffs and Agent 1 also adheres in the first period. The second and the third conditions on the continuation payoff vectors imply that: (i) the incentive compatibility constraints of the two agents are binding – the continuation payoff of an agent who is inspected and found adhering is the minimal amount that makes

[14] The recursive characterization of g in Theorem 1 is restricted to the interval $(0, \frac{1-c^2}{1+c-\delta})$. Outside this interval, $g(0) = \frac{1-c^2}{1+c-\delta}$, and $g(z) = 0$ for all $z \in [\frac{1-c^2}{1+c-\delta}, \frac{1}{1-\delta}]$. The function g is hence well defined on the whole interval $[0, \frac{1}{1-\delta}]$.

him adhere at the current period; and (ii) regardless of which agent is inspected in the current period, the continuation payoff vector lies on the boundary of the PPE payoff set – the continuation payoff of the agent who is *not inspected* is set to be as low as possible.

In fact, the real number p_1 that satisfies the last condition is uniquely determined. The function g is then the minimal solution to the following two equalities:

$$x = \delta \cdot \Big(p_1 \cdot f(p_1) + (1 - p_1) \cdot g\big(f(1 - p_1)\big)\Big) = \delta \cdot w(p_1), \tag{10}$$

and

$$g(x) = \delta \cdot \Big(p_1 \cdot g\big(f(p_1)\big) + (1 - p_1) \cdot f(1 - p_1)\Big) = \delta \cdot w(1 - p_1), \tag{11}$$

which is equivalent to the recursive characterization in Theorem 1. Intuitively, $\delta w(p)$ is the payoff of an agent who adheres when facing the inspection probability p.

Note that the recursive characterization in Eqs. (7) and (8) is not a standard Bellman equation, since w is a composite function of g. Therefore, the existence of a minimal solution is not straightforward. We show that the minimal solution to Eq. (7) does exist, and it can be approximated by the following iterative algorithm: let $g_0 := 0$; for every natural number k, define w_k using Eq. (8) with $g = g_k$, and define g_{k+1} using Eq. (7) with $w = w_k$. As k increases, the function g_k converges to the minimal solution to Eq. (7). As an example, let $c = 0.7$ and $\delta = 0.9$. Figures 4 and 5 show the simulation result for the functions $(g_k)_{k=1}^4$, and g. The optimal PPE payoff is attained at $x^* = 0.19$, which yields the principal a loss $x^* + g(x^*) = 0.38$.

Since Theorem 1 is crucial for the construction of the optimal PPE, we provide a sketch of its proof.

Sketch of proof for Theorem 1.

Step 1: Suppose $x \in (0, \frac{1-c^2}{1+c-\delta})$. We first show that to implement the payoff vector $\big(x, g(x)\big)$, both agents adhere in the first period (Lemma 5). Given that Agent 1 obtains x, delaying his violation to periods after some t rather than letting Agent 1 violate immediately, implies that a higher reward is at stake for Agent 1 in periods before t, and hence lower inspection intensities are needed to deter Agent 1 from violating in these periods. This, roughly speaking, frees up early inspection resources, which can be used to inspect Agent 2, and reduces the number of his violations. Since $x = g\big(g(x)\big)$, by reversing the roles of the two agents, a similar argument can be made to show that Agent 2 also adhere in the first period.

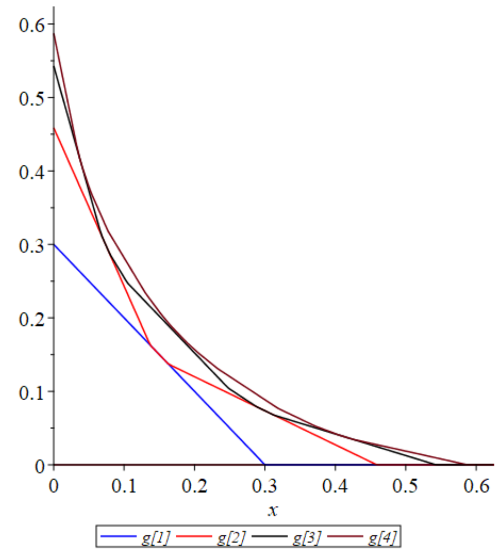

Figure 4: The first four iterations of g_k when $c = 0.7$ and $\delta = 0.9$.

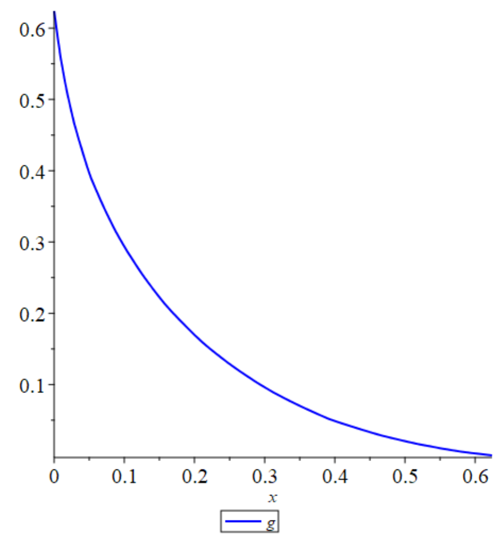

Figure 5: The function g when $c = 0.7$ and $\delta = 0.9$.

The function g then must satisfy the following characterization:

$$\begin{aligned} g(x) = & \min_{p_1, v_1^{A_1}, v_2^{A_1}, v_1^{A_2}, v_2^{A_2}} \delta \cdot \left[p_1 \cdot v_2^{A_1} + (1 - p_1) \cdot v_2^{A_2} \right] \\ s.t. \quad & v_1^{A_1} \geq f(p_1), \\ & v_2^{A_2} \geq f(1 - p_1), \\ & (v_1^{A_1}, v_2^{A_1}) \in \mathscr{E}_*, \\ & (v_1^{A_2}, v_2^{A_2}) \in \mathscr{E}_*, \\ & x = \delta \cdot \left[p_1 \cdot v_1^{A_1} + (1 - p_1) \cdot v_1^{A_2} \right]. \end{aligned} \tag{12}$$

The first and second conditions are incentive compatibility constraints, and ensure that both agents are better off adhering in the first period. The third and forth conditions guarantee that the continuation payoffs of the agents is a PPE outcome in $\mathscr{E}_*$. The last condition ensures that Agent 1's current-period payoff agrees with his continuation payoffs. The value $g(x)$ minimizes Agent 2's payoff under these conditions.

Step 2: We further show that to solve the minimization problem (12), both incentive compatibility constrains are binding, and hence $v_1^{A_1} = f(p_1)$ and $v_2^{A_2} = f(1 - p_1)$. That is, the continuation payoff of an agent who is inspected and found adhering is the minimal amount that makes him adhere at the current period. This result is not straightforward: Increasing the continuation payoff $v_1^{A_1}$ of Agent 1 above $f(p_1)$ leads to a decrease in the continuation payoff $v_2^{A_1}$ of Agent 2, hence it is not clear whether such a change benefits or hurts the principal.

We prove this result by contradiction. Suppose, to the contrary, that $v_1^{A_1} > f(p_1)$. This implies that while keeping agents' continuation payoffs at history A_1 to be $(v_1^{A_1}, v_2^{A_1})$, even if Agent 1 faces a slightly lower inspection intensity, he is still willing to adhere in the first period. Therefore, the principal can shift some monitoring intensities from Agent 1 to Agent 2, without violating Agent 1's incentive compatibility constraint. We show that because of the convexity of g and since $(v_1^{A_1}, v_2^{A_1}) \neq (v_1^{A_2}, v_2^{A_2})$, as Agent 2 is inspected with a higher probability, the continuation payoffs of the agents at history A_2 can be made lower than the weighted average of $(v_1^{A_1}, v_2^{A_1})$ and $(v_1^{A_2}, v_2^{A_2})$ (see Lemma 9 for a detailed discussion). Therefore, shifting monitoring intensities from Agent 1 to Agent 2 benefits the principal. This leads to a contradiction to the assumption that $\big(v_1 = x, v_2 = g(x)\big)$ is on the boundary of the PPE payoff set. An analogous argument applies to the case $v_2^{A_2} > f(1-p_1)$.

Step 3: We then show that, conditional on Agent 1 (resp. Agent 2) being inspected in the first period and obtaining a continuation payoff $f(p_1)$ (resp. $f(1-p_1)$), the other agent's continuation payoff is set to be as low as possible. In other words, regardless of which agent is inspected in the first period, the continuation payoff vector lies on the boundary of the PPE payoff set. The part $v_2^{A_1} = g(v_1^{A_1})$ is intuitive: Since our goal is to minimize Agent 2's expected payoff while keeping Agent 1's payoff to be x, and since Agent 2's continuation payoff if he is not monitored, $v_2^{A_1}$, does not affect his current period incentive, the value $v_2^{A_1}$ is set to be as low as possible. The part $v_1^{A_2} = g(v_2^{A_2})$ is less straightforward. This is because Agent 1's continuation payoff if he is not monitored, $v_1^{A_2}$, affects Agent 1's expected payoff in period 1, which is pre-determined as x. But recall that the agents are symmetric and $x = g\big(g(x)\big)$. That is, given that Agent 2 obtains $g(x)$, the quantity x is the minimal payoff that Agent 1 can get among equilibria in $\mathscr{E}_*$. This observation implies that, the part $v_1^{A_2} = g(v_2^{A_2})$ follows from a similar argument as before, by reversing the roles of the agents.

The minimization problem can therefore be rewritten as Eq. (9), and the recursive characterization of g in Theorem 1 follows. □

3.5.2 Optimal PPE strategy

Our next goal is to provide an explicit characterization of the PPE that implements the optimal payoff vector.

Theorem 2. *The following strategy profile σ^* is a PPE in $\mathscr{E}_*$ and is optimal for the principal.*

Phase 1:

P1.1 *In the first period the principal inspects each agent i with probability $p_i^1 := \frac{1}{2}$. Both agents adhere.*

P1.2 *For every $t \geq 2$, if no violation is detected before period t, then in period t the agent who is inspected in period $t-1$, denoted i, is inspected with probability $p_i^t := w^{-1}\left(\frac{f(p_i^{t-1})}{\delta}\right) < p_i^{t-1}$, and the other agent is inspected with the remaining probability $1 - p_i^t$. Both agents adhere.*

P1.3 *Repeat stage P1.2 until the first period t in which the inspection probability for one of the agents, denoted agent i^*, drops below $f^{-1}(\frac{1-c^2}{1+c-\delta})$ and in addition this agent happens to be inspected again in the current period. Phase 2 starts when this event happens.*

Denote by $p_{i^*}^n$ the probability by which agent i^* is inspected in the period in which Phase 1 ends.

Punishment*. If an agent is found violating in Phase 1, he will be inspected with probability 1 in all future periods, he will adhere in all future periods, and the other agent will violate in all future periods.*

Phase 2: *The principal implements the cyclic rewarding scheme described in Section 3.4 that yields agent i^* the payoff $f(p_{i^*}^n) \geq \frac{1-c^2}{1+c-\delta}$. The other agent is inspected with probability $\frac{1}{1+c}$ and adheres in every period.*

Proof. See Appendix A.8. □

In Phase 2, there are multiple ways to implement the desired payoffs, and the cyclic rewarding scheme we characterized in Section 3.4 is only one of them. Phase 1, in contrast, is more delicate, and its description follows from the recursive characterization of function g in Theorem 1 (we discuss the uniqueness of this structure in Section 4.1). To better understand the structure of Phase 1, we can describe it as a "token game" between the two agents, with the tokens represent an agent's expected payoff if he adheres in the current period.

In the first period, the principal inspects each agent with probability $\frac{1}{2}$, and the agents hold the same amount of tokens x^*, where x^* is the point at which $g(x^*) = x^*$. Depending on the realization of the principal's first-period action, the agent who is inspected and found adhering gains tokens and his amount of tokens increases to $f(\frac{1}{2}) > x^*$; while the agent who is not inspected loses tokens and his amount of tokens decreases to $g(f(\frac{1}{2})) < x^*$. In the second period, the principal adjusts the inspection probabilities according to the updated amount of tokens held by each agent: the agent with more tokens is inspected with a lower probability

$p_i^2 < \frac{1}{2}$, and the other agent is inspected with a higher probability. Again, depending on the realization of the principal's second-period action, the agent who is inspected and found adhering gains tokens, and the other agent loses tokens (see Proposition 10 in Appendix A.8).

The exact number of tokens that are added to and subtracted from each player is derived from the recursive characterization of g in Theorem 1: If agent i is inspected at period t and found adhering, the amount of tokens he holds increases to $f(p_i^t)$, and the amount of tokens the other agent holds decreases to $g\big(f(p_i^t)\big)$. The inspection intensity in each period is then uniquely determined by the value that makes the continuation payoffs (the amount of future tokens) agree with previous commitment. Consequently, in each period of Phase 1, each agent's incentive compatibility constraint is binding, and the agents' payoff vector lies on the lower boundary of the PPE payoff set.

The process continues until one of the agents loses all his tokens. This agent loses the tournament and Phase 1 ends. The length of Phase 1 is random, yet it ends almost surely.[15] In the case where $c = 0.7$ and $\delta = 0.9$, the distribution of Phase 1's length in $10,000$ trials is shown in Figure 6, with an average of 7.86. Due to parameter values, in this case the lowest possible length of Phase 1 is 3.

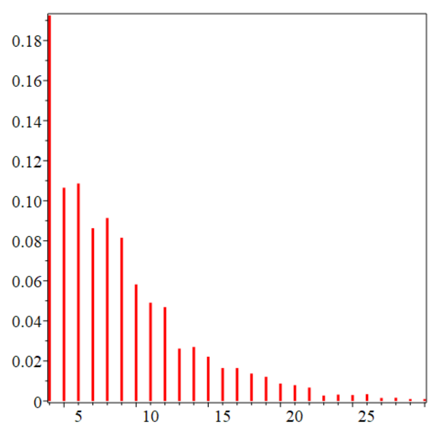

Figure 6: Distribution of Phase 1's length when $c = 0.7$ and $\delta = 0.9$.

[15]Indeed, if there were a public finite history h^t after which Phase 1 ends with probability q smaller than 1, then, conditioned that the history h^t occurs, the probability that Phase 1 ends before stage $t+k$ converges to q as k goes to infinity. Therefore, for every $\epsilon > 0$ there is a positive integer k such that conditioned that the history h^t occurs, the probability that Phase 1 ends after stage $t+k$ is smaller than ϵ. This in turn implies that there is a finite public history of length $t+k$ which is an extension of h^t such that at this history the expected continuation payoff of both agents is smaller than ϵ, contradicting Proposition 1 and the fact that σ^* is a PPE.

Remark 3. *The PPE σ^* described in Theorem 2 can be turned into an agent sequential equilibrium in the game where monitoring is private and the uninspected agent does not learn the outcome of the inspection. This is because (i) under σ_0^* the principal is never idle, hence in the subgame $\Gamma(\sigma_0^*)$ an agent who is not inspected can infer that the other agent is inspected, and (ii) under σ^* the agents play in a deterministic way, hence in the subgame $\Gamma(\sigma_0^*)$ each agent can also infer the action played by the other agent.*

4 Discussion

4.1 Uniqueness of the optimal scheme

In Theorem 1 we characterized the lower bound on the set of PPE outcomes, and identified the payoff vector that yields the minimum violations. In Theorem 2 we provided one inspection strategy that implements the optimal payoff. The question we address in this section is, whether the PPE characterized in Theorem 2 is unique in implementing the desired payoff. We will argue that the structure of Phase 1 is unique under mild assumptions.

Since g is convex and symmetric, the minimum of the principal's loss is obtained at the point x^* where the two agent's payoffs are the same ($x^* = g(x^*)$). This implies that under the optimal scheme the two agents are treated identically in the first period and each is inspected with probability 0.5 (see Appendix A.8 for a detailed argument). Appendix A.7.1 further shows that as long as both agents' payoffs are positive (Phase 1), the dynamics characterized in Theorem 2 is uniquely optimal among inspection schemes that (i) do not use public correlation device, and (ii) do not assign a positive probability to no inspection.[16]

Once an agent's payoff reaches zero and Phase 2 starts, to implement the desired payoff of the other agent, there are many different approaches. The cyclic scheme characterized in Section 3.4 is one option. Another implementation of Phase 2 that has a stationary flavor is described in Appendix A.4. A third implementation, which is not included in the paper, involves delaying the reward in Phase 2 as long as possible.

[16] Even without these assumptions, we tend to believe that the optimal dynamics in Phase 1 is unique. Nevertheless, to formally prove this result, a key step is to show that g is *strictly* convex on $[0, g(0)]$. Even though from numerical simulations this seems to hold, but a formal proof is too involved so we omit it in our paper.

4.2 Noise in detecting violations

In our model we assume that if an agent who violates is inspected, the violation is detected with probability 1. In fact, the structure of the optimal policy is robust to a small noise in monitoring.

Suppose that monitoring fails to detect violation with probability ϵ, yet adherence is properly detected. In this case, the expected cost of violation (for the agent) decreases. Consequently, to induce an agent who is inspected with probability p to adhere in a certain period, the minimum reward if the agent is found adhering increases from (6) to

$$f_\epsilon(p) := \begin{cases} \frac{1-(1+c)p(1-\epsilon)}{\delta p(1-\epsilon)} & \text{if } 0 < p < \frac{1}{1+c}, \\ 0 & \text{if } \frac{1}{1+c} \leq p \leq 1. \end{cases} \tag{13}$$

An argument similar to the proof of Proposition 2 holds, and agents always adhere on the equilibrium path. Therefore, the noise in monitoring affects only the amount of reward claimed by the inspected agent. As long as ϵ is small and an analogous version of Assumption 2 holds, the qualitative structure of the optimal policy remains unchanged (and the proof is analogous to the proof of Theorem 1).

4.3 Flexible punishment levels

In this paper we assume that the punishment c on a detected violation is fixed. Suppose, instead, that whenever a violation is detected, the principal can choose a punishment c in some given range $[\overline{c}, \underline{c}]$, where $\overline{c} > \underline{c} \geq 0$. In this case, Proposition 2 still applies, violation is never detected in equilibrium, and the principal always sets $c = \overline{c}$. The reason is that a detected violation signifies a deviation, and by inflicting the maximal possible penalty $\overline{c}$, the inspection makes deviations less profitable.

This result contrasts some literature in repeated punishment, where a repeated offender is being punished more severely compared with a first time offender. The use of varying punishment levels is typically driven by some form of incomplete information. For instance, Polinsky and Rubinfeld (1991) assume that agents have different characteristic (or offense propensity), which are not observed by the enforcement authority, and Mungan (2010) assumes that repeated violators may gain experience, which lead them to be detected with a lower probability in their subsequent violations.

4.4 Dropping the commitment power of the principal

To support the optimal strategy profile σ^* as an equilibrium, the principal's commitment power is essential. Indeed, a deviation by the principal in a given period does not affect her payoff in that period, since her payoff is determined by the agents' actions, but it affects the agents' continuation play. One way in which the principal could gain by deviating from σ_0^* is by re-allocating monitoring probabilities across agents in the tournament phase (Phase 1) so that it never ends.[17] Since in Phase 1 both agents adhere, this will lower the loss of the principal to 0.

Every PPE in the game where the principal has no commitment power constitutes a PPE in our model (with commitment). Consequently, a principal with no commitment power is not able to obtain a payoff better than the optimal payoff (with commitment) that is identified in Section 3.

4.5 Small δ that violates Assumption 2

In our analysis we assume that the agents are not too impatient, namely $\delta \geq \frac{1-c^2}{1-c^2+c}$ (or equivalently, $f(\frac{c}{1+c}) \leq \frac{1}{1-\delta}$). We now discuss the complementary case.

If the agents are very impatient, explicitly, if $\delta < \frac{1-c}{2-c}$, we have $f(\frac{1}{2}) > \frac{1}{1-\delta}$ and the loss from not violating of an agent who is inspected with probability at most $\frac{1}{2}$ cannot be compensated by future violations. Hence the principal can do no better than using the myopic strategy. Thus, in this case the repetition of the stage game does not benefit the principal.

If $\frac{1-c}{2-c} \leq \delta < \frac{1-c^2}{1-c^2+c}$, the principal can do better than the myopic strategy, but the inspection scheme characterized in Theorem 2 is no longer optimal. This is because now there exists $\widehat{p} \in (\frac{c}{1+c}, \frac{1}{2}]$ such that an agent who is inspected with probability less than $\widehat{p}$ cannot be deterred from violating, even if this agent is promised that in the eventuality that he is inspected and found adhering, he will be inspected with probability 0 in all subsequent periods. Consequently, the principal must reward the agents already along Phase 1. The detailed analysis is tedious and uninspiring, hence we leave it out of the paper.

[17] For instance, suppose in period 5 the principal should inspect Agent 1 with probability 0.2 and Agent 2 with probability 0.8. Suppose also that if Agent 1 is inspected, Phase 1 ends, otherwise, Phase 1 continues. The principal, who has an incentive to prolong Phase 1, can benefit from inspecting Agent 2 with probability 1, instead of 0.8.

References

Andrews, I. and Barron, D. (2016). The allocation of future business: Dynamic relational contracts with multiple agents. *American Economic Review*, 106(9):2742–59.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The Economic Dimensions of Crime*, pages 13–68. Springer.

Board, S. (2011). Relational contracts and the value of loyalty. *American Economic Review*, 101(7):3349–67.

Bond, P. and Hagerty, K. (2010). Preventing crime waves. *American Economic Journal: Microeconomics*, 2(3):138–59.

Braithwaite, J. (1981). The limits of economism in controlling harmful corporate conduct. *Law and Society Review*, 16:481.

Burgess, S. and Ratto, M. (2003). The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy*, 19(2):285–300.

Carrasco, V., Fuchs, W., and Fukuda, S. (2019). From equals to despots: The dynamics of repeated decision making in partnerships with private information. *Journal of Economic Theory*, 182:402–432.

de Clippel, G., Eliaz, K., Fershtman, D., and Rozen, K. (2021). On selecting the right agent. *Theoretical Economics*, 16(2):381–402.

Dubin, J. A. and Wilde, L. L. (1988). An empirical analysis of federal income tax auditing and compliance. *National Tax Journal*, pages 61–74.

Eeckhout, J., Persico, N., and Todd, P. E. (2010). A theory of optimal random crackdowns. *American Economic Review*, 100(3):1104–35.

EPA, U. (2010). 2009 state summary data for clean air act majors. *Fiscal Year 2009 Final Report*.

Fong, K. (2009). Evaluating skilled experts: Optimal scoring rules for surgeons. Working paper Stanford University.

Gray, W. B. and Deily, M. E. (1996). Compliance and enforcement: Air pollution regulation in the us steel industry. *Journal of Environmental Economics and Management*, 31(1):96–111.

Greenberg, J. (1984). Avoiding tax avoidance: a (repeated) game-theoretic approach. *Journal of Economic Theory*, 32(1):1–13.

Guo, Y. and Hörner, J. (2018). Dynamic allocation without money. Cowles foundation discussion paper.

Harford, J. D. (1991). Measurement error and state-dependent pollution control enforcement. *Journal of Environmental Economics and Management*, 21(1):67–81.

Harford, J. D. and Harrington, W. (1991). A reconsideration of enforcement leverage when penalties are restricted. *Journal of Public Economics*, 45(3):391–395.

Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics*, 37(1):29–53.

Heyes, A. (2000). Implementing environmental regulation: Enforcement and compliance. *Journal of Regulatory Economics*, 17(2):107–129.

Landsberger, M. and Meilijson, I. (1982). Incentive generating state dependent penalty system: the case of income tax evasion. *Journal of Public Economics*, 19(3):333–352.

Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, 121(3):1029–1061.

Li, J., Matouschek, N., and Powell, M. (2017). Power dynamics in organizations. *American Economic Journal: Microeconomics*, 9(1):217–41.

Lipnowski, E. and Ramos, J. (2018). Repeated delegation. SSRN Discussion Paper.

Mungan, M. C. (2010). Repeat offenders: If they learn, we punish them more severely. *International Review of Law and Economics*, 30(2):173–177.

Polinsky, A. M. and Rubinfeld, D. L. (1991). A model of optimal fines for repeat offenders. *Journal of Public Economics*, 46(3):291–306.

Rockafellar, R. T. and Wets, R. J.-B. (2009). *Variational Analysis*, volume 317. Springer Science & Business Media.

Russell, C. S. (1990). *Public Policies for Environmental Protection*, chapter Monitoring and Enforcement. Routledge.

Spear, S. E. and Srivastava, S. (1987). On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617.

Stigler, G. J. (1970). The optimal enforcement of laws. *Journal of Political Economy*, 78(3):526–536.

Thomas, J. and Worrall, T. (1990). Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory*, 51(2):367–390.

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

For Online Publication

A Appendix

A.1 Proof of Proposition 2

A.1.1 Proof of Proposition 2(i)

A mixed PPE is a strategy profile $(\sigma_0, \sigma_1, \sigma_2)$, where σ_0, σ_1, and σ_2 are behavior public strategies, and for every finite history h^t, the pair $(\sigma_1|_{h^t}, \sigma_2|_{h^t})$ is an agent Nash equilibrium of the subgame $\Gamma(\sigma_0|_{h^t})$. In this section we show that the principal's payoff in every mixed PPE can be attained in a pure PPE that satisfies the conditions in Proposition 2(i).

As a first step, we show that every mixed PPE σ is realization equivalent to a pure PPE σ' (that does not necessarily satisfy Proposition 2(i)). The idea is to mimic agents' mixed actions using the correlation device. In the second step, we show that there exists another PPE $\overline{\sigma}$ that yields the principal the same payoff as σ' and whenever an agent is found deviating, he is inspected with probability 1 in all future periods. This change guarantees that off-equilibrium paths satisfy the conditions in Proposition 2(i). Finally, in Step 3, we show that there exists a pure PPE $\widehat{\sigma}$ that yields the principal the same payoff as $\overline{\sigma}$ and satisfies the conditions in Proposition 2(i).

Step 1: Mimicking σ by a pure PPE σ'.

Suppose that σ is a mixed PPE. Let σ' be the strategy profile σ with the exception that the correlation device replaces the lotteries made by the agents. That is, if under σ_1 at the history h^t Agent 1 adheres with probability $0 < p < 1$, then under σ'_1 Agent 1 at h^t adheres if and only if the outcome of the correlation device is in some set C that has probability p (and is otherwise independent of the players' strategies). The strategy profiles σ and σ' are payoff equivalent. Since $(\sigma_1|_{h^t}, \sigma_2|_{h^t})$ is an agent Nash equilibrium of the subgame $\Gamma(\sigma_0|_{h^t})$ for every finite history h^t, so is the pair $(\sigma'_1|_{h^t}, \sigma'_2|_{h^t})$. It follows that strategy profile σ' is a pure PPE.

Step 2: Modifying the play off-the-equilibrium.

We now construct a strategy profile $\overline{\sigma}$ by modifying σ' according to the following rule. If agent i is inspected and found deviating, he is punished most severely: from the following period and on he will be inspected with probability 1. The agent being punished will adhere and the other agent will violate. Since $\overline{\sigma}$ differs from σ' only at histories that are off-the-equilibrium, since deviations are less profitable in $\overline{\sigma}$ than in σ', and since agents follow an agent PPE after a deviation is detected, it follows that $\overline{\sigma}$ is a PPE that yields the players the same payoffs as σ'.

Step 3: Constructing a pure PPE $\widehat{\sigma}$ that yields the principal the same payoff as $\overline{\sigma}$ and satisfies the conditions in Proposition 2(i).

Since the strategy $\overline{\sigma}$ is pure, we have to consider only histories under which an agent who is indifferent between adhering and violating chooses Violate with probability 1 under $\overline{\sigma}$. Such a history must occur on the equilibrium path, because by Step 2, off-the-equilibrium path the agent who is inspected with probability 1 strictly prefers adhering, and the other agent strictly prefers violating. We now construct a PPE under which Agent 1 is inspected with probability 0 at this history, so that violating becomes a strictly preferable action.

Let $\widehat{\sigma}$ be the strategy profile that is derived from $\overline{\sigma}$ with the following changes: Suppose that at a given history, denoted h^t, an agent (say, Agent 1) is indifferent between adhering and violating and violates with probability 1. Then (a) at history h^t, when the lottery made by the principal tells her to inspect Agent 1, she does not inspect any agent, and (b) for the continuation of the play she (and the agents) act as if Agent 1 was inspected and violated.

The modification of $\widehat{\sigma}$ at h^t increases Agent 1's payoff at h^t, and does not affect the principal's loss, nor Agent 2's payoff. Since h^t is on the equilibrium path, this modification does not affect the equilibrium property in earlier stages. Since the modification in $\widehat{\sigma}$ does not affect Agent 2's payoff, it does not change Agent 2's best response. It follows that $\widehat{\sigma}$ is a PPE, and, relative to $\overline{\sigma}$, the payoff of the agents increase while the principal's loss remains the same.

A.1.2 Proof of Proposition 2(ii)

In this section we prove Proposition 2(ii), which states that we can assume that the principal's loss is equal to the sum of payoffs of the two agents. Suppose that $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ is a PPE under which

$$v_0(\sigma) > v_1(\sigma) + v_2(\sigma), \tag{14}$$

i.e., the loss of the principal exceeds the total gain of the agents. By Eqs. (1)–(2), this can happen only if under a history on the equilibrium path, an agent violates while he is inspected with positive probability. That is, there exists a history h^k that occurs with positive probability under σ, such that $\sigma_0(I_i|h^k) > 0$ and $\sigma_i(V|h^k) = 1$ for some agent, say, Agent 1. Then, by a similar argument as Step 3 in Section A.1.1, we can construct an alternative PPE that yields the principal the same payoff as σ, and instead of inspecting Agent 1 with positive probability, the principal inspects him with probability 0 at history h^k, and uses a public correlation device at the beginning of the next period to mimic the original continuation play. As in Section A.1.1, this change does not affect Agent 2's incentives, and, since we increase Agent 1's payoff only at histories on the equilibrium path, Agent 1's incentives are not affected either.

By repeating this argument for every history h^k that occurs with positive probability in equilibrium and satisfies $\sigma_0(I_i|h^k) > 0$ and $\sigma_i(V|h^k) = 1$ for some agent i, we can construct a PPE $\widehat{\sigma} = (\widehat{\sigma}_0, \widehat{\sigma}_1, \widehat{\sigma}_2)$ under which whenever $\widehat{\sigma}_i(V|h^k) = 1$, we have $\widehat{\sigma}_0(I_i|h^k) = 0$. Proposition 2(ii) follows.

A.1.3 Proof of Proposition 2(iii)

Let σ be a PPE that satisfies $v_0(\sigma) = v_1(\sigma) + v_2(\sigma)$, which implies that under σ no violation is observed on-the-equilibrium path. If a violation is detected, then necessarily the inspected agent deviated. Let $\widehat{\sigma}$ be similar to σ, except that violations are followed by indefinite punishment, the punished agent adheres and the other agent violates. Formally, if Agent 1 is found violating at history h^k, then from period k onward the principal inspects Agent 1 with probability 1. By a similar argument as in Step 1 in Section A.1.1, $\widehat{\sigma}$ is a PPE that yields the players the same payoff as σ.

A.2 Proof of Proposition 4

The behavior of an agent at a given period depends on the probability he is inspected, as well as on the continuation payoffs following every eventuality. The next lemma, which implies Proposition A.2, provides the explicit relationship between these parameters that ensure that an agent adheres.

Lemma 1. *Let $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ be a PPE in $\mathscr{E}_*$, and let $i \in \{1, 2\}$ be an agent. For every finite history $h^k \in \mathscr{H}$, we have $\sigma_i(A|h^k) = 1$ if and only if*

$$v_i(\sigma|_{(h^k,A_i)}) - v_i(\sigma|_{(h^k,V_i)}) \geq \frac{\left(1-\sigma_0(I_i|h^k)\right)-c\cdot\sigma_0(I_i|h^k)}{\sigma_0(I_i|h^k)\cdot\delta}. \tag{15}$$

Proof. Without loss of generality let $i = 1$. Fix a finite public history h^k, and denote by $p_1 := \sigma_0(I_1|h^k)$ the probability that Agent 1 is inspected at history h^k. If Agent 1 violates, his payoff at h^k is

$$E \equiv (1 - p_1) - c \cdot p_1 + (1 - p_1)\delta v_1^{NI} + p_1\delta \cdot v_1(\sigma|_{(h^k,V_1)}), \tag{16}$$

where v_1^{NI} is Agent 1's expected continuation payoff if he is not inspected at h^k. If Agent 1 adheres, his payoff at h^k is

$$F \equiv 0 + (1 - p_1)\delta v_1^{NI} + p_1\delta \cdot v_1(\sigma|_{(h^k, A_1)}). \tag{17}$$

It is worth noting that Agent 1's expected continuation payoff at h^k if he is not inspected does not depend on the action he actually plays. Subtracting Eq. (16) from Eq. (17), we obtain

$$F - E = p_1\delta\big(v_1(\sigma|_{(h^k,A_1)}) - v_1(\sigma|_{(h^k,V_1)})\big) - \big((1 - p_1) - cp_1\big).$$

Since $\sigma \in \mathscr{E}_*$, we have $\sigma_i(A|h^k) = 1$ if and only if $F - E \geq 0$, and the claim follows. $\square$

Proposition 4 follows from Proposition 2(iii) and Lemma 1.

A.3 Proof of Proposition 5

We first describe a PPE that yields the payoff vector $\big(v_1 = \frac{1-c^2}{1+c-\delta}, v_2 = 0\big)$. We will then prove that this is the lowest PPE payoff for Agent 1, provided that Agent 2 obtains 0.

In any PPE that yields Agent 2 a payoff 0, Agent 2 is always inspected with probability $\frac{1}{1+c}$, and he always adheres. We therefore focus only on the inspection for Agent 1. Consider the following strategy profile σ:

1. In the first period, inspect Agent 1 with probability $\frac{c}{1+c}$. Agent 1 adheres.
2. If Agent 1 is inspected and found adhering, his continuation payoff is $f\left(\frac{c}{1+c}\right)$, and the players continue with a strategy profile that implements this payoff vector (will be explained later).
3. If Agent 1 is not inspected, the players forget past play and restart implementing σ.
4. Punishment: if Agent 1 is found violating, he is inspected with probability 1 in all future periods, and he will always adhere.

By Proposition 4, adhering is the best response of Agent 1 in the first period of σ. Denote $v_1 = v_1(\sigma)$ the payoff of Agent 1 under σ. Then $v_1 = \delta\cdot\big(\frac{c}{1+c} \cdot f(\frac{c}{1+c}) + \frac{1}{1+c} \cdot v_1\big)$, and therefore $v_1 = \frac{1-c^2}{1+c-\delta}$.

The rest of this section describes a cyclic rewarding scheme that yields Agent 1 a payoff $f(\frac{c}{1+c})$. Consider the following strategy profile $\sigma(d) = (\sigma_0(d), \sigma_1(d), \sigma_2(d))$, where d is a natural number:

1. In the first $d-1$ periods of the cycle, inspect Agent 1 with probability $\frac{c}{1+c}$. Agent 1 adheres.

2. If Agent 1 is not found violating in these $d-1$ periods, in the dth period of the cycle the principal inspects Agent 1 with probability 0, and Agent 1 violates. A new cycle starts from the next period.

3. Punishment: if Agent 1 is found violating, he is inspected with probability 1 in all future periods, and Agent 1 will adhere in all future periods.

In $\sigma(d)$, Agent 1 violates once every d periods. Agent 1's payoff $v_1(\sigma(d)) = \delta^{d-1} \cdot \frac{1}{1-\delta^d}$ is decreasing in the cycle length d. Denote by d^* the solution to $v_1(\sigma(d)) = f(\frac{c}{1+c})$. If d^* is an integer, then $\sigma(d^*)$ is a PPE, which yields Agent 1 the payoff $f(\frac{c}{1+c})$, as desired. Note that under $\sigma(d)$, Agent 1 is willing to adhere in each of the first $d-1$ periods in a given cycle, because Agent 1's continuation payoff upon being inspected and found adhering is at least $f\left(\frac{c}{1+c}\right)$ (because of discounting, the continuation payoff of Agent 1 in some periods may exceed $f\left(\frac{c}{1+c}\right)$).

If d^* is not an integer, for example, $d^* = 8.2$ when $c = 0.7$ and $\delta = 0.9$, then the length of the cycle is random and determined at the end of the 7th period of the cycle: with some probability the 8th period is the rewarding period and the cycle has length 8, and with the remaining probability the 9th period is the rewarding period and the cycle has length 9. The probability is chosen to ensure that the expected payoff of Agent 1 is $f(\frac{c}{1+c})$.

We now show that any payoff lower than $\frac{1-c^2}{1+c-\delta}$ is not attainable. The inspection probability for Agent 1 in every period is at most $\frac{c}{1+c}$. Because of discounting, it is not optimal to allow the agent to violate in the first period and to deter him from violating in later periods.[18] Indeed, if under some PPE σ' Agent 1 violates in the first period, the principal would profit by skipping the first period, and starting the implementation of σ' from the second period. Therefore, an optimal scheme deters Agent 1 from violating in the first period. By Proposition 4, Agent 1 adheres in the first period when he is inspected with probability p only if his continuation payoff in the eventuality that he is inspected and found adhering is at least $f(p)$. Denote by v_1 the lowest PPE payoff for Agent 1. If Agent 1 is not inspected in the first period, then his lowest equilibrium payoff is still v_1. Therefore,

$$v_1 \geq \min_{0 \leq p \leq \frac{c}{1+c}} \delta \cdot \big(p \cdot f(p) + (1-p)v_1\big). \tag{18}$$

[18] Agent 1 adheres in at least one period. Indeed, by Assumption 2, $f(\frac{c}{1+c}) \leq \frac{1}{1-\delta}$, and Agent 1 can be deterred from violating in at least one period.

The right-hand-side of Eq. (18) is minimized at $p = \frac{c}{1+c}$ and Eq. (18) solves to $v_1 \geq \frac{1-c^2}{1+c-\delta}$.

A.4 Generating PPE payoffs $(v_1 = f(p), v_2 = f(1-p))$ with stationary rewards

In this section we identify an inspection strategy that yields Agent 1 and Agent 2 the payoffs $f(p)$ and $f(1-p)$, respectively, for every $p \in [\frac{c}{1+c}, \frac{1}{1+c}]$. This result, together with the definition of the function g, implies that for every $p \in [\frac{c}{1+c}, \frac{1}{1+c}]$,

$$g\big(f(1-p)\big) \leq f(p). \tag{19}$$

Set $q_1 = (1-\delta) \cdot f(p)$ and $q_2 = (1-\delta) \cdot f(1-p)$. Consider the following strategy profile $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ that is periodic with length 1 and uses the correlation device, which determines one out of four behaviors in each period.

- As long as no agent is found violating, at each period:
  1. With probability $q_1 \cdot q_2$ the principal inspects no agent. Both agents violate.
  2. With probability $q_1 \cdot (1-q_2)$ the principal inspects Agent 1 and Agent 2 with probabilities 0 and $1-p$, respectively. Agent 1 violates and Agent 2 adheres.
  3. With probability $(1-q_1) \cdot q_2$ the principal inspects Agent 1 and Agent 2 with probabilities p and 0, respectively. Agent 1 adheres and Agent 2 violates.
  4. With probability $(1-q_1) \cdot (1-q_2)$ the principal inspects Agent 1 and Agent 2 with probabilities p and $1-p$, respectively. Both agents adhere.
- If an agent is found violating, he is inspected with probability 1 in all future periods, he adheres in all future periods, and the other agent violates in all future periods.

We argue that the strategy profile σ is a PPE that yields Agent 1 and Agent 2 the payoffs $f(p)$ and $f(1-p)$, respectively. Since σ is stationary and since under σ Agent 1 violates in each period with probability q_1, Agent 1's payoff, denoted v_1, solves the equation $v_1 = q_1 + \delta v_1$, hence $v_1 = \frac{q_1}{1-\delta} = f(p)$. Analogously, Agent 2's payoff, denoted v_2 satisfies $v_2 = f(1-p)$.

We next focus on the histories under which no violation has been detected. Agent 1 violates whenever he is inspected with probability 0, and adheres whenever he is inspected with probability p. Proposition 4 implies that he cannot profit by deviating from σ_1. Similarly, Agent 2 cannot profit by deviating from σ_2. It follows that σ is a PPE that yields Agent 1 and Agent 2 the payoffs $f(p)$ and $f(1-p)$, respectively.

Note that with $p = \frac{1}{1+c}$ we obtain a PPE that yields Agent 1 the payoff $f(\frac{c}{1+c})$ and Agent 2 the payoff 0. In this case $q_2 = 0$.

A.5 Proof of Proposition 6

Proposition 6 follows from Proposition 5 and the next lemma, which asserts that if $(v_1 = y_1, v_2 = y_2)$, $y_1 < \frac{1}{1-\delta}$, can be supported as a PPE outcome, then $(v_1 = \gamma, v_2 = y_2)$, where $\gamma \in (y_1, \frac{1}{1-\delta}]$, can also be supported as a PPE outcome.

Lemma 2. *Suppose there exists a PPE $\sigma = (\sigma_0, \sigma_1, \sigma_2) \in \mathscr{E}_*$ such that $v_1(\sigma) = y_1$ and $v_2(\sigma) = y_2$, where $y_1 < \frac{1}{1-\delta}$. Then for every $\gamma \in (y_1, \frac{1}{1-\delta}]$ there exists another PPE $\sigma' = (\sigma'_0, \sigma'_1, \sigma'_2) \in \mathscr{E}_*$ such that $v_1(\sigma') = \gamma$ and $v_2(\sigma') = y_2$.*

Proof. Let σ'' be the strategy profile that is derived from σ with the following changes: (a) whenever the principal is supposed to inspect Agent 1, she does not inspect any agent, yet (b) for the continuation of the play she (and the agents) act as if Agent 1 was inspected and adhered, and (c) Agent 1 always violates. The reader can verify that σ'' is a PPE, $v_1(\sigma'') = \frac{1}{1-\delta}$, and $v_2(\sigma'') = y_2$. By using a proper correlation device that chooses at the outset of the game whether the players implement σ or σ'' we obtain a PPE σ' that satisfies the conditions of the lemma. $\square$

A.6 Properties of function g

Proposition 7. *(i) The function g is non-negative and non-increasing; (ii) The function g is convex; (iii) $g(0) = \frac{1-c^2}{1+c-\delta}$ and $g(\frac{1-c^2}{1+c-\delta}) = 0$; (iv) The function g is continuous on $[0, \frac{1-c^2}{1+c-\delta}]$; (v) For every $x \in (0, \frac{1-c^2}{1+c-\delta})$, we have $g\big(g(x)\big) = x$.*

Since an agent can guarantee himself a payoff 0 by always adhering, $g(x) \geq 0$. We turn to prove that g is monotonic. Suppose $x' > x$. By Lemma 2, the vector $(x', g(x))$ can be supported as a PPE payoff. By the definition of $g(\cdot)$, we have $g(x') \leq g(x)$. It follows that g is non-increasing. The convexity of the function g described in Part (ii) follows from the use of the correlation device, and it implies the continuity of g on $(0, \frac{1-c^2}{1+c-\delta}]$. We will prove later that g is also continuous at $x = 0$. Part (iii) is an easy corollary of Proposition 5. To see why Part (v) holds, we fix $x \in (0, \frac{1-c^2}{1+c-\delta})$ and suppose that $g\big(g(x)\big) < x$. Since g is non-increasing and continuous, there exists $y < g(x)$ such that $g(y) = x$. This implies that $(v_1 = x, v_2 = y)$ can be supported as a PPE outcome, contradicting the definition of $g(x)$.

We turn to prove the continuity of g at 0. By Part (iii), we have $\frac{1-c^2}{1+c-\delta} = g(0)$. Since g is non-increasing (by Part (i)) and bounded, the limit $y := \lim_{x \searrow 0} g(x)$ exists and $y \leq \frac{1-c^2}{1+c-\delta}$. Suppose that $y < \frac{1-c^2}{1+c-\delta}$. We claim that $g(y) > 0$. Indeed, if $g(y) = 0$, then the payoffs $(v_1 = y, v_2 = 0)$ can be supported as a PPE outcome, contradicting $g(0) = \frac{1-c^2}{1+c-\delta}$. Denote $x' := g(y)$. By the definition of y and since g is non-increasing, for every $k \in \mathbb{N}$, we have $g(0 + \frac{1}{k}) \leq y$. This observation, together with Lemma 2, implies that for every $k \in \mathbb{N}$ the vector $(v_1 = \frac{1}{k}, v_2 = y)$ can be supported as a PPE payoff, contradicting $g(y) = x' > 0$. Consequently, we have $y = \frac{1-c^2}{1+c-\delta}$, as desired.

A.7 Proof of Theorem 1

We first define a function $\widehat{g} : [0, \frac{1-c^2}{1+c-\delta}] \to [0, \frac{1}{1-\delta}]$ that is similar to g, but excludes the use of the correlation device in the first period. We will characterize the structure of a PPE that implements the payoff vector $\big(v_1 = x, v_2 = \widehat{g}(x)\big)$ (Proposition 8). As a corollary, we obtain the functional representation of $\widehat{g}$ (Corollary 1). We then show that in fact, $\widehat{g}$ is convex and it agrees with g (Proposition 9).

Definition 1. *For every $x \in [0, \frac{1-c^2}{1+c-\delta}]$, let $\widehat{g}(x)$ be the minimum payoff of Agent 2 over all PPEs that yield Agent 1 the payoff x and do not use the correlation device in the first period.*

By a similar argument as the one given in Section 3.5, the minimum in the definition of $\widehat{g}(x)$ is attained. Note that to attain $\widehat{g}(x)$, the correlation device may be used after period 1.

Proposition 8. *Let $x \in (0, \frac{1-c^2}{1+c-\delta})$.*
(i) The following strategy profile $\widehat{\sigma}$ is a PPE in $\mathscr{E}_*$ that yields Agent 1 the payoff x and does not use the correlation device in the first period:

(a) In period 1 the principal inspects Agent 1 with probability $p_1 := w^{-1}\left(\frac{x}{\delta}\right)$ and Agent 2 with probability $1 - p_1$, where the function w is defined in Eq. (8). Both agents adhere.

(b) If Agent 1 is inspected in the first period and found adhering, from the second period the players implement a strategy profile that yields the payoffs $\big(v_1 = f(p_1), v_2 = g(f(p_1))\big)$.

(c) If Agent 2 is inspected in the first period and found adhering, from the second period the players implement a strategy profile that yields the payoffs $\big(v_1 = g(f(1-p_1)), v_2 = f(1-p_1)\big)$

Punishment: if an agent is found violating in the first period, he is inspected with probability 1 in all future periods.
(ii) Moreover, $\widehat{\sigma}$ implements the payoffs $\big(v_1 = x, v_2 = \widehat{g}(x)\big)$.

Proof. We already argued (in Section 3.5) that $w(p)$ is strictly decreasing on $[0, \frac{1}{1+c}]$. It can be verified that $w(\frac{1}{1+c}) = 0$, and hence the function $w^{-1}(x)$ is well defined on $[0, w(0)]$. Since $w(0) = \frac{1}{\delta} + g(0)$, it can be verified that $\frac{1-c^2}{1+c-\delta} < \delta \cdot w(0)$, and hence $p_1 = w^{-1}\left(\frac{x}{\delta}\right)$, is well defined for $x \in (0, \frac{1-c^2}{1+c-\delta})$. Part (i) of Proposition 8 follows by Proposition 4 and the definition of w. Part (ii) of Proposition 8 will be proven in Subsection A.7.1. □

Figure 7 shows graphically the payoffs of the two agents and the principal's play in the first period under $\widehat{\sigma}$.

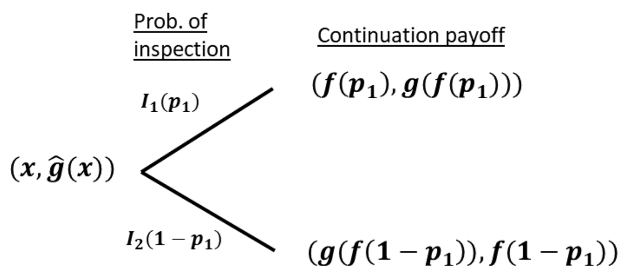

Figure 7: The first period under $\widehat{\sigma}$.

Proposition 8 implies in particular that $\widehat{g}(x) = \delta \cdot w\left(1 - w^{-1}\left(\frac{x}{\delta}\right)\right)$. Indeed, to make the expected discounted payoff agree with the continuation payoffs we have

$$x = 0 + \delta \cdot \Big(p_1 \cdot f(p_1) + (1 - p_1) \cdot g\big(f(1 - p_1)\big)\Big) = \delta \cdot w(p_1), \tag{20}$$

and

$$\widehat{g}(x) = 0 + \delta \cdot \Big(p_1 \cdot g\big(f(p_1)\big) + (1 - p_1) \cdot f(1 - p_1)\Big) = \delta \cdot w(1 - p_1). \tag{21}$$

Eq. (20) holds if and only if $p_1 = w^{-1}\left(\frac{x}{\delta}\right)$. Replacing p_1 in Eq. (21) with $w^{-1}\left(\frac{x}{\delta}\right)$, we obtain that the function $\widehat{g}$ must satisfy $\widehat{g}(x) = \delta \cdot w\left(1 - w^{-1}\left(\frac{x}{\delta}\right)\right)$. We thus proved the following characterization of the function $\widehat{g}$.

Corollary 1. *For every* $x \in (0, \frac{1-c^2}{1+c-\delta})$,

$$\widehat{g}(x) = \delta \cdot w\left(1 - w^{-1}\left(\frac{x}{\delta}\right)\right), \tag{22}$$

where

$$w(p) := \begin{cases} p \cdot f(p) + (1 - p) \cdot g\big(f(1 - p)\big) & \text{if } 0 < p \le 1, \\ \frac{1}{\delta} + g(0) & \text{if } p = 0. \end{cases} \tag{23}$$

Below we will show that $\widehat{g}(x) = g(x)$ for $x \in (0, \frac{1-c^2}{1+c-\delta})$, hence Corollary 1 provides the characterization of g stated in Theorem 1. We will also prove that this characterization provides a recursive algorithm to approximate the function g.

By Definition 1, the function $\widehat{g}$ is defined only on the interval $[0, \frac{1-c^2}{1+c-\delta}]$. We next define a *one-stage optimal operator*, $\theta : [0, \infty) \to [0, \infty)$, that agrees with $\widehat{g}$ on the interval $[0, \frac{1-c^2}{1+c-\delta}]$. Let $\theta : [0, \infty] \to [0, \infty)$ be the function

$$\theta(x) := \begin{cases} \delta \cdot w\left(1 - w^{-1}\left(\frac{x}{\delta}\right)\right) & \text{if } 0 \leq x \leq \delta w(0), \\ 0 & \text{if } x > \delta w(0). \end{cases} \tag{24}$$

As argued above, $\delta w(0) > \frac{1-c^2}{1+c-\delta}$. The next proposition states that the function θ is convex. As we will see later, this result implies that no correlation device in needed in Phase 1.

Proposition 9. *The function θ is convex on $[0, \infty)$.*

Proof. See Section A.7.2. $\square$

By Proposition 9, to implement $(v_1 = x, v_2 = g(x))$ for $x \in (0, \frac{1-c^2}{1+c-\delta})$, the correlation device is not needed in the first period. This implies that the function g agrees with $\widehat{g}$ for every $x \in (0, \frac{1-c^2}{1+c-\delta})$. Let $\mathscr{H}_e \subset \mathscr{H}$ be the set of finite histories h^t that occur with positive probability under the optimal PPE $\widehat{\sigma}$ that is described in Proposition 8 and satisfy $v_1(\widehat{\sigma}|_{h^t}) \in (0, \frac{1-c^2}{1+c-\delta})$. The agents' payoffs under every history $h^t \in \mathscr{H}_e$ can be written in the form $v_2(\widehat{\sigma}|_{h^t}) = \widehat{g}\big(v_1(\widehat{\sigma}|_{h^t})\big)$. Therefore, Proposition 8 defines a strategy profile that does not use the correlation device as long as both agents' payoffs are in the range $(0, \frac{1-c^2}{1+c-\delta})$. Note that by Proposition 7, $g(x) \in (0, \frac{1-c^2}{1+c-\delta})$ for $x \in (0, \frac{1-c^2}{1+c-\delta})$.

We complete the proof of Theorem 1 by showing, in Section A.7.3, that the minimal solution to Eq. (7) exists, and we provide an iterative algorithm to approximate it.

A.7.1 Proof of Proposition 8

As mentioned earlier, we need to prove only Part (ii) of Proposition 8. Let $\mathscr{E}_*^{nc}$ be the set of all PPEs in $\mathscr{E}_*$ where no correlation device is used in the first period.

We start by studying some properties of the function $\widehat{g}$ (Lemmas 3 and 4). Fix $x \in (0, \frac{1-c^2}{1+c-\delta})$ and denote by σ a PPE in $\mathscr{E}_*^{nc}$ that implements $(v_1 = x, v_2 = \widehat{g}(x))$. Using the properties of $\widehat{g}$, we show that in the first period of σ both agents adhere (Lemma 5) and the principal is not idle (Lemma 6). We then show that if agent i is inspected and found adhering in the first period, his continuation payoff from period 2 and on is $f(\sigma_0(I_i))$ (Lemma 9), and the other agent's continuation payoff from period 2 and on is $g\left(f(\sigma_0(I_i))\right)$ (Lemma 8).

Lemma 3. *Let $x \in (0, \frac{1-c^2}{1+c-\delta})$. For every $x' \in (x, \frac{1-c^2}{1+c-\delta}]$, the payoff vector $(v_1 = x', v_2 = \widehat{g}(x))$ can be supported as a PPE outcome in $\mathscr{E}_*^{nc}$.*

Proof. Lemma 3 differs from Lemma 2 in two respects: while in Lemma 2 the principal could use the correlation device in the first period, in Lemma 3 she cannot, and while in Lemma 2 the upper bound on the payoff x' was $\frac{1}{1-\delta}$, in Lemma 3 it is $\frac{1-c^2}{1+c-\delta}$. To prove Lemma 2, we had the principal select at the outset, using the correlation device, whether to implement the vector $(x, g(x))$ or the vector $(\frac{1}{1-\delta}, g(x))$ as PPE. Here, since the correlation device is not used, the principal can make the choice only at the beginning of the second period. Assumption 2 implies that $\frac{1-c^2}{1+c-\delta} < \frac{\delta}{1-\delta}$. Since by assumption the principal can implement the payoff vector $(x, \widehat{g}(x))$ without using the correlation device in the first period, by properly increasing the continuation payoff of Agent 1 from the second stage and on while keeping Agent 2's continuation payoff (as done in Lemma 2), the principal can implement the payoff $(x', \widehat{g}(x))$ for every $x' \in (x, \frac{1-c^2}{1+c-\delta})$.

□

Lemma 4. *(i) The function $\widehat{g}$ is non-negative and non-increasing on $[0, \frac{1-c^2}{1+c-\delta}]$; (ii) $\widehat{g}(0) = \frac{1-c^2}{1+c-\delta}$ and $\widehat{g}(\frac{1-c^2}{1+c-\delta}) = 0$.*

Proof. If an agent adheres in every period, he guarantees payoff 0. Therefore $\widehat{g}$ is non-negative. The fact that $\widehat{g}$ is non-increasing on $[0, \frac{1-c^2}{1+c-\delta}]$ follows from Lemma 3. We next show that Part (ii) of Lemma 4 holds. The inspection strategy constructed in Section 3.4 that implements the payoffs $(v_1 = 0, v_2 = \frac{1-c^2}{1+c-\delta})$ does not use the correlation device in the first period. Therefore $\widehat{g}(0) \le \frac{1-c^2}{1+c-\delta}$ and $\widehat{g}(\frac{1-c^2}{1+c-\delta}) \le 0$. Since $\widehat{g}(0) \ge g(0) = \frac{1-c^2}{1+c-\delta}$, we have $\widehat{g}(0) = \frac{1-c^2}{1+c-\delta}$. Since the function $\widehat{g}$ is non-negative, $\widehat{g}(\frac{1-c^2}{1+c-\delta}) = 0$. □

We use the properties of $\widehat{g}$ to study the structure of the PPE that implements $(v_1 = x, v_2 = \widehat{g}(x))$.

Lemma 5. *Let $x \in (0, \frac{1-c^2}{1+c-\delta})$. Suppose that $\sigma \in \mathscr{E}_*^{nc}$ is a PPE that implements $(v_1 = x, v_2 = \widehat{g}(x))$. Then under σ both agents adhere in the first period.*

Proof. Throughout the proof we restrict attention to PPEs in $\mathscr{E}_*^{nc}$. If $x < 1$ and $\widehat{g}(x) < 1$, then Lemma 5 holds since a violating agent gains 1, which exceeds each agent's total payoff. We next consider the case where at least one agent's payoff is no less than 1.

Let $\widehat{g}_{aa}(x)$ be the lowest PPE payoff of Agent 2 when Agent 1 obtains x and both agents adhere in the first period, let $\widehat{g}_{av}(x)$ (resp. $\widehat{g}_{va}(x)$) be the lowest PPE payoff of Agent 2 when Agent 1 obtains x and only Agent 1 (resp. Agent 2) adheres in the first period, and let $\widehat{g}_{vv}(x)$ be the lowest PPE payoff of Agent 2 when Agent 1 obtains x and both agents violate in the first period. If there does not exist a PPE in $\mathscr{E}_*^{nc}$ under which Agent 1 obtains x and both agents adhere in the first period, we let $\widehat{g}_{aa}(x) = \infty$. The same convention is used for $\widehat{g}_{av}$, $\widehat{g}_{va}$, and $\widehat{g}_{vv}$, and hence they are well defined over $[0, \frac{1}{1-\delta}]$. Note that $\widehat{g}(x) = \min\big(\widehat{g}_{aa}(x), \widehat{g}_{av}(x), \widehat{g}_{va}(x), \widehat{g}_{vv}(x)\big)$.

Our goal is to prove that $\widehat{g}(x) = \widehat{g}_{aa}(x)$. We will bound $\widehat{g}_{aa}(x)$ (Step 1), calculate $\widehat{g}_{av}(x)$ (Step 2), and show that $\widehat{g}_{av}(x) > \widehat{g}_{aa}(x)$ (Step 3). In Step 4 we show that $\widehat{g}_{vv}(x) \geq \widehat{g}_{aa}(x)$, and in Step 5 we verify that $\widehat{g}_{va}(x) > \widehat{g}_{aa}(x)$.

<u>Step 1</u>: Bounding $\widehat{g}_{aa}(x)$.

Consider the strategy profile $\widehat{\sigma}$ that is defined in Proposition 8. We already argued that $\widehat{\sigma}$ is a PPE. By definition, $\widehat{g}_{aa}(x) \leq v_2(\widehat{\sigma})$. By construction, $\widehat{\sigma}$ yields Agent 1 the payoff x and Agent 2 the payoff

$$\widehat{g}_{aa}(x) \leq v_2(\widehat{\sigma}) = 0 + \delta \cdot \Big(p_1 \cdot g\big(f(p_1)\big) + (1-p_1) \cdot f(1-p_1)\Big), \tag{25}$$

where $p_1 = w^{-1}(\frac{x}{\delta})$. We now argue that $p_1 \in (\frac{c}{1+c}, \frac{1}{1+c})$, and we will use this fact in Step 3. Since $x \in (0, \frac{1-c^2}{1+c-\delta})$ and by Lemma 4, we have $\widehat{g}(x) > 0$. Therefore, $v_2(\widehat{\sigma}) \geq \widehat{g}_{aa}(x) \geq \widehat{g}(x) > 0$, and the expected payoffs of the two agents under $\widehat{\sigma}$ are both positive. Since $v_1(\widehat{\sigma}) = \delta w(p_1)$ and $v_2(\widehat{\sigma}) = \delta w(1-p_1)$ (see Eqs. (20)–(21)), we have $w(p_1) > 0$ and $w(1-p_1) > 0$. Since $w(p)$ is decreasing in p and since $w(\frac{1}{1+c}) = 0$, we have $p_1 < \frac{1}{1+c}$ and $1 - p_1 < \frac{1}{1+c}$, and hence $p_1 \in (\frac{c}{1+c}, \frac{1}{1+c})$.

<u>Step 2</u>: Calculating $\widehat{g}_{av}(x)$.

Let $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ be a PPE in $\mathscr{E}_*^{nc}$ under which Agent 1 adheres in the first period and Agent 2 violates in the first period. By the definition of $\mathscr{E}_*^{nc}$, Agent 2 is inspected with probability 0 in the first period. Let $\sigma_0(I_1)$ be the inspection probability for Agent 1 in the first period. As in the proof of Proposition 2(ii), we can assume without loss of generality that $\sigma_0(I_1) = 1$.

Suppose then that Agent 1 is inspected with probability 1 in the first period. Then in the first period Agent 1 obtains zero and Agent 2 obtains 1. From the next period and on, Agent 1 obtains a payoff $\frac{x}{\delta}$, and consequently, Agent 2's payoff is at least $g\left(\frac{x}{\delta}\right)$. Since the players can implement the payoffs $\big(v_1 = \frac{x}{\delta}, v_2 = g(\frac{x}{\delta})\big)$, it follows that

$$\widehat{g}_{av}(x) = 1 + \delta \cdot g\left(\frac{x}{\delta}\right). \tag{26}$$

<u>Step 3</u>: $\widehat{g}_{aa}(x) < \widehat{g}_{av}(x)$.

By Eqs. (26), (25), and (20),

$$\begin{aligned}
\widehat{g}_{av}(x) - \widehat{g}_{aa}(x) \geq \widehat{g}_{av}(x) - v_2(\widehat{\sigma}) &= 1 + \delta \cdot g\left(\frac{x}{\delta}\right) - \delta \cdot \Big(p_1 \cdot g\big(f(p_1)\big) + (1-p_1) \cdot f(1-p_1)\Big) \\
&= 1 + \delta \cdot g\Big(p_1 \cdot f(p_1) + (1-p_1) \cdot g\big(f(1-p_1)\big)\Big) \\
&\quad - \delta \cdot \Big(p_1 \cdot g\big(f(p_1)\big) + (1-p_1) \cdot f(1-p_1)\Big) \\
&= 1 - \delta \cdot Z,
\end{aligned} \tag{27}$$

where

$$Z := p_1 \cdot g\big(f(p_1)\big) + (1-p_1) \cdot f(1-p_1) - g\Big(p_1 \cdot f(p_1) + (1-p_1) \cdot g\big(f(1-p_1)\big)\Big). \tag{28}$$

We will show that $Z < \frac{1}{\delta}$, which implies that $\widehat{g}_{av}(x) > \widehat{g}_{aa}(x)$. By Eq. (19), and since $p_1 \in (\frac{c}{1+c}, \frac{1}{1+c})$ by Step 1, we have

$$f(p_1) \geq p_1 \cdot f(p_1) + (1-p_1) \cdot g\big(f(1-p_1)\big).$$

By Proposition 7(i) the function g is non-increasing, which, together with the fact that $p_1 < 1$, implies that

$$p_1 \cdot g\big(f(p_1)\big) < g\big(f(p_1)\big) \leq g\Big(p_1 \cdot f(p_1) + (1-p_1) \cdot g\big(f(1-p_1)\big)\Big). \tag{29}$$

By Eqs. (28), (29), and (6),

$$\begin{aligned}
Z &= p_1 \cdot g\big(f(p_1)\big) + (1-p_1) \cdot f(1-p_1) - g\Big(p_1 \cdot f(p_1) + (1-p_1) \cdot g\big(f(1-p_1)\big)\Big) \\
&< (1-p_1) \cdot f(1-p_1) = \frac{1}{\delta} - \frac{(1+c)(1-p_1)}{\delta} < \frac{1}{\delta},
\end{aligned} \tag{30}$$

as claimed.

Step 4: $\widehat{g}_{vv}(x) \geq \widehat{g}_{av}(x)$.

Let $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ be a PPE in $\mathscr{E}_*^{nc}$ that implements the payoff $(v_1 = x, v_2 = \widehat{g}_{vv}(x))$, under which both agents violate in the first period. By the definition of $\mathscr{E}_*^{nc}$, both agents are inspected with probability 0 in the first period. Denote $y := \widehat{g}_{vv}(x)$. In the first period both agents gain 1, and therefore the agents' continuation payoffs are $\frac{x-1}{\delta}$ and $\frac{y-1}{\delta}$, respectively. When Agent 1's payoff is $\frac{x-1}{\delta}$, the lowest PPE payoff of Agent 2 is $g(\frac{x-1}{\delta})$. The definition of

$\widehat{g}_{vv}(x)$ implies that $\frac{y-1}{\delta} = g(\frac{x-1}{\delta})$. It follows that $\widehat{g}_{vv}(x) = y = 1+\delta\cdot g(\frac{x-1}{\delta})$. By Eq. (26) and since the function g is non-increasing, we have $\widehat{g}_{av}(x) = 1+\delta\cdot g\left(\frac{x}{\delta}\right) \leq 1+\delta\cdot g(\frac{x-1}{\delta}) = \widehat{g}_{vv}(x)$, as claimed.

Step 5: $\widehat{g}(x) = \widehat{g}_{aa}(x)$.

By Steps 1–4, we have $\widehat{g}(x) = \min\big(\widehat{g}_{aa}(x), \widehat{g}_{va}(x)\big)$. To prove that $\widehat{g}(x) = \widehat{g}_{aa}(x)$, we consider two cases, depending on whether $x = \widehat{g}\big(\widehat{g}(x)\big)$ or not.

Denote by $\mathcal{X}$ the set of $x \in (0, \frac{1-c^2}{1+c-\delta})$ that satisfies $x = \widehat{g}\big(\widehat{g}(x)\big)$. Let $x \in \mathcal{X}$ and let $y := \widehat{g}(x)$. Since $x = \widehat{g}(\widehat{g}(x))$, we have $x = \widehat{g}(y)$. Since $x \in (0, \frac{1-c^2}{1+c-\delta})$ and by Lemma 4, we have $y \in (0, \frac{1-c^2}{1+c-\delta})$. The payoff vector $\big(x, \widehat{g}(x)\big)$ can therefore be written as $\big(\widehat{g}(y), y\big)$, with $y \in (0, \frac{1-c^2}{1+c-\delta})$. By reversing the roles of the two agents and by the arguments given in Steps 1–3, we have $\widehat{g}_{va}(y) > \widehat{g}_{aa}(y)$. Let $\mathcal{Y}$ be the set of all $\widehat{g}(x)$ where $x \in \mathcal{X}$. The above argument implies that for every $x \in \mathcal{Y}$, we have $\widehat{g}_{va}(x) > \widehat{g}_{aa}(x)$, and hence $\widehat{g}(x) = \widehat{g}_{aa}(x)$.

We now argue that if $x \in \mathcal{X}$, then necessarily $x \in \mathcal{Y}$. This result, together with the previous argument, will imply that $\widehat{g}(x) = \widehat{g}_{aa}(x)$ for every $x \in \mathcal{X}$. Let $x \in \mathcal{X}$. By definition, we have $x = \widehat{g}\big(\widehat{g}(x)\big)$. Let $y := \widehat{g}(x)$. Then (i) $x = \widehat{g}(y)$, and (ii) $y = \widehat{g}(x) = \widehat{g}\big(\widehat{g}(y)\big)$ so that $y \in \mathcal{X}$. Therefore, $x \in \mathcal{Y}$, as claimed.

Consider next the case $x \notin \mathcal{X}$. By the definition of $\widehat{g}$ and since $(v_1 = x, v_2 = \widehat{g}(x))$ can be supported as a PPE outcome, we have $x \geq \widehat{g}\big(\widehat{g}(x)\big)$. Since $x \notin \mathcal{X}$, we have $x > \widehat{g}\big(\widehat{g}(x)\big)$. We will prove that $\big(v_1 = \widehat{g}\big(\widehat{g}(x)\big), v_2 = \widehat{g}(x)\big)$ can be supported as a PPE outcome in $\mathscr{E}_*^{nc}$ in which both agents adhere in the first period. Since $x \in \big(\widehat{g}\big(\widehat{g}(x)\big), \frac{1-c^2}{1+c-\delta}\big)$, by Lemma 3 and its proof, this will imply that the payoff vector $(v_1 = x, v_2 = \widehat{g}(x))$ can also be supported as a PPE outcome in $\mathscr{E}_*^{nc}$ in which both agents adhere in the first period, and therefore we will obtain that $\widehat{g}(x) = \widehat{g}_{aa}(x)$.

Denote $y := \widehat{g}(x)$. By the definition of $\widehat{g}$, we have $\widehat{g}(y) \leq x$. Since $\widehat{g}$ is non-increasing, we have $\widehat{g}\big(\widehat{g}(y)\big) \geq \widehat{g}(x) = y$. Since $(v_1 = y, v_2 = \widehat{g}(y))$ can be supported as a PPE in $\mathscr{E}_*^{nc}$, we have $\widehat{g}\big(\widehat{g}(y)\big) \leq y$, and therefore $\widehat{g}\big(\widehat{g}(y)\big) = y$. By the previous argument, the payoff vector $(v_1 = y, v_2 = \widehat{g}(y))$ can be supported as a PPE outcome in which both agents adhere in the first period. By reversing the roles of the two agents, we conclude that there exists a PPE that implements $\big(v_1 = \widehat{g}\big(\widehat{g}(x)\big), v_2 = \widehat{g}(x)\big)$ where both agents adhere in the first period, as desired. $\square$

Lemma 6. *Let $x \in (0, \frac{1-c^2}{1+c-\delta})$ and suppose that $\sigma \in \mathscr{E}_*^{nc}$ is a PPE that implements $(v_1 = x, v_2 = \widehat{g}(x))$. Then without loss of generality we can assume that the principal is not idle in the first period, that is, $\sigma_0(\emptyset) = 0$.*

Proof. Suppose that there exists a PPE $\sigma = (\sigma_0, \sigma_1, \sigma_2)$ under which $v_1(\sigma) = x$, $v_2(\sigma) = \widehat{g}(x)$, and $\sigma_0(I_1)+\sigma_0(I_2)<1$. Denote $p_1:=\sigma_0(I_1)$ and $p_2:=\sigma_0(I_2)$. Figure 8 describes the first period under σ.

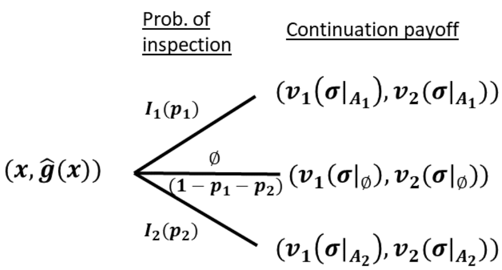

Figure 8: The first period under σ.

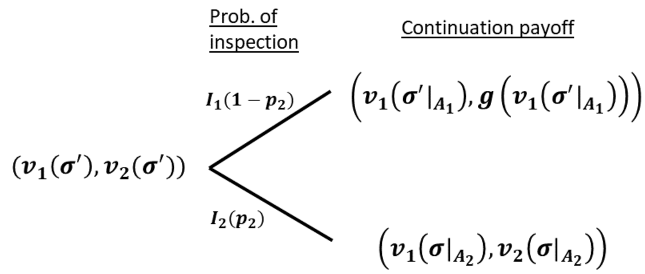

Figure 9: The first period under σ'

We will show that there is another PPE $\sigma'=(\sigma_0',\sigma_1',\sigma_2')$ that yields the principal the same payoff and satisfies $\sigma_0'(\emptyset)=0$. The structure of the proof is similar to Step 2 in the proof of Lemma 5, yet additional effort is needed to show that adhering is the best response of Agent 1 in the first period under σ'. Let $\sigma'=(\sigma_0',\sigma_1',\sigma_2')$ be the strategy profile that is similar to σ except for the following modifications (see Figure 9):
(i) In the first period the principal inspects Agent 1 with probability $1-p_2$ and Agent 2 with probability p_2. Both agents adhere.
(ii) If Agent 1 is inspected in the first period and found adhering, the players implement the PPE that yields $v_1(\sigma'|_{A_1})=\frac{p_1}{1-p_2}v_1(\sigma|_{A_1})+\frac{1-p_1-p_2}{1-p_2}v_1(\sigma|_{\emptyset})$ and $v_2(\sigma'|_{A_1})=g\big(v_1(\sigma'|_{A_1})\big)$.

We now verify that σ' is a PPE. For Agent 2, the probability of being inspected in period 1 and the continuation payoff if being found adhering are the same under σ' and σ. Therefore, by Proposition 4, Agent 2 chooses the same action (Adhere) in the first period under σ' and σ. It is left to verify that adhering is Agent 1's best response in the first period of σ'. By the definition of σ',

$$\begin{aligned}\sigma_0'(I_1)\cdot v_1(\sigma'|_{A_1})=\big(\sigma_0(I_1)+\sigma_0(\emptyset)\big)\cdot v_1(\sigma'|_{A_1})&=\sigma_0(I_1)\overbrace{v_1(\sigma|_{A_1})}^{\geq f(\sigma_0(I_1))}+\sigma_0(\emptyset)v_1(\sigma|_{\emptyset})\\&\geq\sigma_0(I_1)\cdot f\big(\sigma_0(I_1)\big)\\&>\sigma_0'(I_1)\cdot f\big(\sigma_0'(I_1)\big).\end{aligned}\tag{31}$$

Indeed, since σ is a PPE in $\mathscr{E}_*^{nc}$, by Lemma 5 Agent 1 adheres in the first period, and hence $v_1(\sigma|_{A_1})\geq f\big(\sigma_0(I_1)\big)$. The last inequality in Eq. (31) holds since $p\cdot f(p)$ is decreasing in p and since $\sigma_0(I_1)<\sigma_0'(I_1)$. Further, since $\sigma_0(I_1)$ is non-negative, $\sigma_0'(I_1)>0$ and hence by Eq. (31), we have $v_1(\sigma'|_{A_1})>f\big(\sigma_0'(I_1)\big)$. Proposition 4 now implies that Agent 1 is better off adhering in the first period of σ'.

Agent 1's payoff under σ' is $v_1(\sigma') = v_1(\sigma) = x$, while Agent 2's payoff under σ' is

$$v_2(\sigma') = 0 + \delta \cdot \Big((1 - p_2) \cdot \overbrace{v_2(\sigma'|_{A_1})}^{=g\left(v_1(\sigma'|_{A_1})\right)} + p_2 \cdot v_2(\sigma'|_{A_2})\Big). \tag{32}$$

By the definition of σ' and the convexity of g (Proposition 7(ii)), we have

$$v_2(\sigma'|_{A_1}) = g\big(v_1(\sigma'|_{A_1})\big) = g\left(\tfrac{p_1}{1-p_2} v_1(\sigma|_{A_1}) + \tfrac{1-p_1-p_2}{1-p_2} v_1(\sigma|_{\emptyset})\right) \leq \tfrac{p_1}{1-p_2} g\big(v_1(\sigma|_{A_1})\big) + \tfrac{1-p_1-p_2}{1-p_2} g\big(v_1(\sigma|_{\emptyset})\big). \tag{33}$$

Since σ' is a PPE in $\mathscr{E}_*^{nc}$, and by Eqs. (32) and (33),

$$\widehat{g}(x) \leq v_2(\sigma') \leq \delta \cdot \Big(p_1 \cdot g\big(v_1(\sigma|_{A_1})\big) + (1 - p_1 - p_2) \cdot g\big(v_1(\sigma|_{\emptyset})\big) + p_2 \cdot v_2(\sigma|_{A_2})\Big) = v_2(\sigma) = \widehat{g}(x). \tag{34}$$

Hence $v_2(\sigma') = v_2(\sigma)$, and therefore σ' satisfies the desired properties. $\square$

Let $x \in (0, \frac{1-c^2}{1+c-\delta})$ and let $\sigma \in \mathscr{E}_*^{nc}$ be a PPE that implements $(v_1 = x, v_2 = \widehat{g}(x))$ and where the principal is not idle in the first period. Denote $p_1 := \sigma_0(I_1)$, it follows that $\sigma_0(I_2) = 1 - p_1$. By Lemma 5, both agents adhere in the first period under σ, hence $p_1 > 0$ and $1-p_1 > 0$. Let $A := v_1(\sigma|_{A_1})$, $B := v_2(\sigma|_{A_1})$, $C := v_1(\sigma|_{A_2})$, and $D := v_2(\sigma|_{A_2})$. The first period under σ is summarized in Figure 10. By Proposition 4, $A \geq f(p_1)$ and $D \geq f(1-p_1)$.

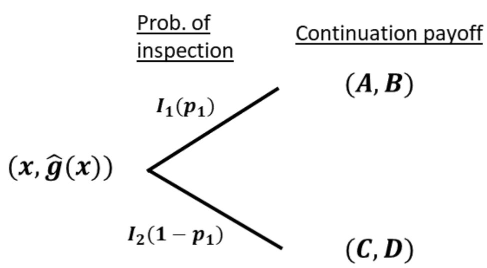

Figure 10: The first period under σ.

Proposition 7(v) states that $g\big(g(x)\big) = x$ for every $x \in (0, \frac{1-c^2}{1+c-\delta})$. We now prove that the analogous property holds for the function $\widehat{g}$.

Lemma 7. *For every $x \in (0, \frac{1-c^2}{1+c-\delta})$, we have $\widehat{g}\big(\widehat{g}(x)\big) = x$.*

Proof. Step 1: The lemma holds if $\widehat{g}$ is continuous on $(0, \frac{1-c^2}{1+c-\delta})$.

Suppose that $\widehat{g}$ is continuous on $(0, \frac{1-c^2}{1+c-\delta})$ and let $x \in (0, \frac{1-c^2}{1+c-\delta})$. Since $(v_1 = x, v_2 = \widehat{g}(x))$ can be implemented by a PPE in $\mathscr{E}_*^{nc}$, we have $\widehat{g}\big(\widehat{g}(x)\big) \le x$. Suppose by contradiction that $\widehat{g}\big(\widehat{g}(x)\big) < x$. If $\widehat{g}(x) = 0$, then by Lemma 4(ii), $\widehat{g}\left(\widehat{g}(x)\right) = \frac{1-c^2}{1+c-\delta} \ge x$, a contradiction. Hence $\widehat{g}(x) > 0$. Since $\widehat{g}$ is non-increasing, since $\widehat{g}(0) > x > \widehat{g}\big(\widehat{g}(x)\big)$, and since by assumption $\widehat{g}$ is continuous on $(0, \frac{1-c^2}{1+c-\delta})$, there exists $0 < y < \widehat{g}(x)$ such that $\widehat{g}(y) = x$. This implies that $(v_1 = x, v_2 = y)$ can be implemented by a PPE in $\mathscr{E}_*^{nc}$, contradicting the definition of $\widehat{g}$ at x. Therefore, it is sufficient to prove that $\widehat{g}$ is continuous on $(0, \frac{1-c^2}{1+c-\delta})$.

Step 2: For every $x \in (0, \frac{1-c^2}{1+c-\delta})$ there exists a PPE σ in $\mathscr{E}_*^{nc}$ that implements $(v_1 = x, v_2 = \widehat{g}(x))$ and satisfies in addition $\frac{c}{1+c} \le \sigma_0(I_1), \sigma_0(I_2) \le \frac{1}{1+c}$.

Let σ be a PPE in $\mathscr{E}_*^{nc}$ that satisfies Lemma 6 and implements the payoffs $(v_1 = x, v_2 = \widehat{g}(x))$. Denote by $p_1 := \sigma_0(I_1)$ the inspection probability for Agent 1 in the first period under σ (see Figure 10). We will show that we can assume without loss of generality that $p_1 \in [\frac{c}{1+c}, \frac{1}{1+c}]$. Since under σ the principal is not idle in the first period, this implies that the inspection probability for Agent 2, $\sigma_0(I_2)$, is in $[\frac{c}{1+c}, \frac{1}{1+c}]$ as well.

Suppose that one of the agents is inspected with probability lower than $\frac{c}{1+c}$. Without loss of generality, suppose $p_1 < \frac{c}{1+c}$, and hence $1 - p_1 > \frac{1}{1+c}$. Define an alternative PPE, $\widehat{\sigma}$, where we set the inspection probability for Agent 2 to $\frac{1}{1+c}$, and assign the remaining probability $(1 - p_1 - \frac{1}{1+c})$ to no inspection. Under $\widehat{\sigma}$, the continuation payoff if no one is inspected is the same as the continuation payoff if Agent 2 is inspected and found adhering (see Figure 11). Since Agent 2 is inspected with probability $\frac{1}{1+c}$ in the first period under $\widehat{\sigma}$, by Proposition 4 and Eq. (6), Agent 2 adheres in the first period. The reader can verify that $\widehat{\sigma}$ is a PPE that implements the payoffs $\big(v_1(\widehat{\sigma}) = x, v_2(\widehat{\sigma}) = \widehat{g}(x)\big)$. Note that $\widehat{\sigma} \notin \mathscr{E}_*^{nc}$.

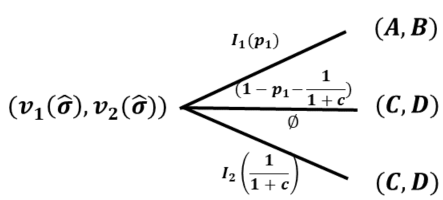

Figure 11: The first period under $\widehat{\sigma}$.

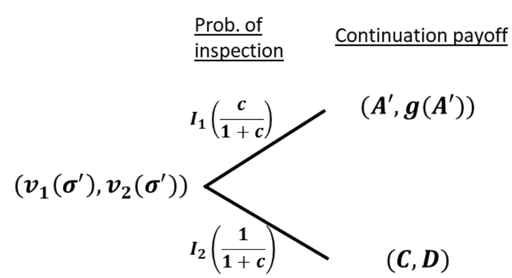

Figure 12: The first period under σ'.

Now we construct an alternative PPE in $\mathscr{E}_*^{nc}$, denoted σ', by adding the probability for no inspection under $\widehat{\sigma}$ to the inspection of Agent 1 (see Figure 12). Under σ', if Agent 1 is inspected, the players implement the PPE that yields Agent 1 the payoff $A' := \frac{p_1 A + \left(\frac{c}{1+c} - p_1\right) C}{\frac{c}{1+c}}$

and Agent 2 the payoff $g(A')$. Using the same argument as in the proof of Lemma 6, σ' is a PPE in $\mathscr{E}_*^{nc}$ where both agents adhere in the first period and implements the payoffs $(v_1(\sigma') = x, v_2(\sigma') = \widehat{g}(x))$.

Step 3: The function $\widehat{g}$ is continuous on $[0, \frac{1-c^2}{1+c-\delta}]$.

Let $X := [0, \frac{1-c^2}{1+c-\delta}]$ and $W \subset [0,1] \times [0, \frac{1}{1-\delta}]^4$. Let $G : X \twoheadrightarrow W$ be the compact-valued correspondence defined by

$$G(x) := \Big\{(p_1, A, B, C, D) : \quad \begin{aligned} & \delta \cdot \big(p_1 A + (1-p_1)C\big) = x, \\ & f(p_1) \le A \le \frac{1}{1-\delta}, \\ & g(A) \le B \le \frac{1}{1-\delta}, \\ & 0 \le C \le \frac{1}{1-\delta} \\ & \max\big(f(1-p_1), g(C)\big) \le D \le \frac{1}{1-\delta}, \\ & \frac{c}{1+c} \le p_1 \le \frac{1}{1+c} \Big\}. \end{aligned} \tag{35}$$

A tuple (p_1, A, B, C, D) that is in $G(x)$ can define a strategy profile in which both agents adhere in the first period and where the payoff to Agent 1 is x. Indeed, p_1 (resp. $(1-p_1)$) represents the inspection probability for Agent 1 (resp. Agent 2) and A, B, C, D are the continuation payoffs as described in Figure 10. The first condition in the definition of $G(x)$ implies that the expected payoff of Agent 1 is x; the second and the third conditions imply that it is optimal for Agent 1 to adhere in the first period and the payoffs $(v_1 = A, v_2 = B)$ can be supported as a PPE outcome; the forth and the fifth conditions imply that it is optimal for Agent 2 to adhere in the first period and the payoffs $(v_1 = C, v_2 = D)$ can be supported as a PPE outcome; the sixth condition requires that the inspection probability for Agent 1 is in $[\frac{c}{1+c}, \frac{1}{1+c}]$, which is made without loss of generality as shown in Step 2.

For every $x \in X$, the set $G(x)$ is not empty. Indeed, it can be verified that the tuple $(\widehat{p}_1, \widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ defined by $\widehat{p}_1 := w^{-1}(\frac{x}{\delta})$, $\widehat{A} := f(\widehat{p}_1)$, $\widehat{B} := g(f(\widehat{p}_1))$, $\widehat{C} := g(f(1-\widehat{p}_1))$, and $\widehat{D} := f(1-\widehat{p}_1)$, where the function w is defined in Eq. (8), lies in $G(x)$.

Denote by $h : W \to \mathbb{R}$ the continuous function

$$h(\overbrace{p_1, A, B, C, D}^{v \in W}) := \delta \cdot \big(p_1 B + (1-p_1)D\big). \tag{36}$$

This is Agent 2's expected payoff under the strategy profile that is defined by the tuple (p_1, A, B, C, D). Then, for every $x \in X$, we have

$$\widehat{g}(x) = \min\Big(h(v)|v \in G(x)\Big). \tag{37}$$

By Berge's Maximum Theorem, to prove that the function $\widehat{g}$ is continuous on $[0, \frac{1-c^2}{1+c-\delta}]$, it is sufficient to prove that the correspondence G is continuous, i.e., both upper and lower hemicontinuous, on $[0, \frac{1-c^2}{1+c-\delta}]$.

The upper hemicontinuity of G follows from the continuity of f and g. We now turn to proving that G is lower hemicontinuous: for every $x \in X$, every $v \in G(x)$, and every sequence $x_n \to x$, there exist $N \geq 1$ and a sequence $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n \to v$ and $v_n \in G(x_n)$ for every $n \geq N$. Fix $v = (p_1, A, B, C, D) \in G(x)$, $x \in X$, and a sequence $(x_n)_{n\in\mathbb{N}}$ in X that converges to x.

Case 1: $0 < C < \frac{1}{1-\delta}$.

For every $n \geq 1$, define $v_n := (p_1^n, A_n, B_n, C_n, D_n)$ by $p_1^n := p_1$, $A_n := A$, $B_n := B$, $C_n := C - \frac{x-x_n}{\delta(1-p_1)}$, and $D_n := \max\big(g(C_n), D\big)$. Since the function g is continuous, we have $v_n \to v$. When x_n is sufficiently close to x, the quantity C_n is sufficiently close to C, and therefore $C_n \in (0, \frac{1}{1-\delta})$. If follows that $g(C_n) < g(0) = \frac{1-c^2}{1+c-\delta} < \frac{1}{1-\delta}$, and $v_n \in G(x_n)$.

Case 2: $C = 0$ and $A = f(p_1)$.

For every $n \geq 1$, define $v_n := (p_1^n, A_n, B_n, C_n, D_n)$ as follows. (i) If $x_n > x$, then $p_1^n := p_1$, $A_n := A$, $B_n = B$, $C_n = \frac{x_n-x}{\delta(1-p_1)}$, and $D_n := \max\big(g(C_n), D\big) = D$, while (ii) if $x_n < x$, then $p_1^n := p_1 + \frac{x-x_n}{1+c}$, $A_n := f(p_1^n)$, $B_n := \max\big(g(A_n), B\big)$, $C_n := C = 0$, and $D_n := \max\big(f(1-p_1^n), D\big)$. Because of the continuity of the functions f and g, we have $v_n \to v$. It can be verified that $\delta \cdot \big(p_1^n A_n + (1-p_1^n) \cdot C_n\big) = x_n$, and hence the first condition in Eq. (35) holds for x_n and v_n. We now verify that $p_1^n \in [\frac{c}{1+c}, \frac{1}{1+c}]$. By construction, $p_1^n \geq p_1 \geq \frac{c}{1+c}$. Since $x > 0$ and $C = 0$, we have $A = f(p_1) > 0$ and hence $p_1 < \frac{1}{1+c}$. Therefore, when x_n is sufficiently close to x, we have $p_1^n < \frac{1}{1+c}$, as desired. It follows that for such n, A_n, B_n, C_n and D_n are in the proper range and we have $v_n \in G(x_n)$.

Case 3: $C = 0$ and $A > f(p_1)$.

For every $n \geq 1$, define $v_n := (p_1^n, A_n, B_n, C_n, D_n)$ as follows. (i) If $x_n > x$, then $p_1^n := p_1$, $A_n := A$, $B_n = B$, $C_n = \frac{x_n-x}{\delta(1-p_1)}$, and $D_n := D$, while (ii) if $x_n < x$, then $p_1^n := p_1$, $A_n := A - \frac{x-x_n}{\delta p_1}$, $B_n := g(A_n)$, $C_n := C = 0$, and $D_n := D$. Since the function g is continuous, we have $v_n \to v$. As in Case 2, it can be verified that $\delta \cdot \big(p_1^n A_n + (1-p_1^n) \cdot 0\big) = x_n$. For x_n sufficiently close to x, we have that A_n is sufficiently close to A, and hence $A_n > f(p_1^n) = f(p_1)$. It can then be verified that for such n, we have $v_n \in G(x_n)$.

Case 4: $C = \frac{1}{1-\delta}$.

We first argue that $A < \frac{1}{1-\delta}$. Indeed, since $x = p_1\delta A + (1-p_1)\delta C \leq \frac{1-c^2}{1+c-\delta}$ and since $\delta C = \frac{\delta}{1-\delta} > \frac{1-c^2}{1+c-\delta}$ (see Assumption 2), we have $\delta A < \frac{1-c^2}{1+c-\delta} < \frac{\delta}{1-\delta}$, as claimed. For every $n \geq 1$, define $v_n := (p_1^n, A_n, B_n, C_n, D_n)$ as follows. (i) If $x_n > x$, then $p_1^n := p_1$, $A_n := A + \frac{x_n - x}{\delta p_1}$, $B_n = \max\big(g(A_n), B\big) = B$, $C_n = C$, and $D_n := D$, while (ii) if $x_n < x$, then $p_1^n := p_1$, $A_n := A$, $B_n := B$, $C_n := C - \frac{x - x_n}{\delta(1-p_1)}$, and $D_n := \max\big(g(C_n), D\big)$. It can be verified that $x_n = \delta\big(p_1 A_n + (1-p_1)C_n\big)$. When x_n is sufficiently close to x, we have $A_n < \frac{1}{1-\delta}$ and $g(C_n) = g(\frac{1-c^2}{1+c-\delta}) = 0$. It follows that for such n, we have $v_n \in G(x_n)$. This completes the proof that the correspondence G is lower hemicontinuous. $\square$

For $x \in (0, \frac{1-c^2}{1+c-\delta})$, let $\overline{\sigma}^* \in \mathscr{E}_*^{nc}$ be the PPE that implements the payoffs $(v_1 = x, v_2 = \widehat{g}(x))$ and satisfies the properties stated in Lemmas 5–7 (see Figure 13). One can expect that $B = g(A)$ and $C = g(D)$. This is the content of the next lemma.

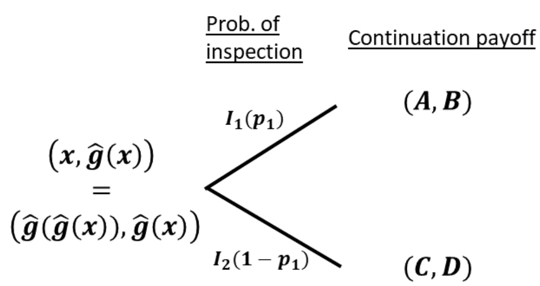

Figure 13: The first period under $\overline{\sigma}^*$.

Lemma 8. $B = g(A)$ *and* $C = g(D)$.

Proof. By the definition of g, we have $B \geq g(A)$ and $C \geq g(D)$. Assume first that $B > g(A)$. Consider the strategy profile σ' under which $\sigma'_0(I_1) = p_1$, $\sigma'_0(I_2) = 1 - p_1$, $v_1(\sigma'|_{A_1}) = A$, $v_2(\sigma'|_{A_1}) = g(A)$, $v_1(\sigma'|_{A_2}) = C$, and $v_2(\sigma'|_{A_2}) = D$, and both agents adhere in the first period. That is, if Agent 1 is inspected in the first period the players implement the PPE that supports the payoffs $(A, g(A))$, and if Agent 2 is inspected in the first period the players implement the PPE that supports the payoffs (C, D).

Since $A \geq f(p_1)$ and $D \geq f(1-p_1)$, the strategy profile σ' is a PPE in $\mathscr{E}_*^{nc}$. Agent 1's payoff under σ' is

$$v_1(\sigma') = \delta \cdot \big(p_1 \cdot A + (1-p_1) \cdot C\big) = x,$$

and Agent 2's payoff is

$$v_2(\sigma') = \delta \cdot \big(p_1 \cdot g(A) + (1-p_1) \cdot D\big) < \widehat{g}(x),$$

contradicting the definition of $\widehat{g}(x)$. This implies that $B = g(A)$. By Lemma 7, we have $\widehat{g}\big(\widehat{g}(x)\big) = x$, hence an analogous argument shows that $C = g(D)$. □

Lemma 9. $A = f(p_1)$ *and* $D = f(1-p_1)$.

Proof. We only prove that $A = f(p_1)$. The proof that $D = f(1-p_1)$ is analogous.

Step 1: If $\frac{1-c^2}{1+c-\delta} < A \leq \frac{1}{1-\delta}$ then $A = f(p_1)$.

By Proposition 7(i,iii), in this case $g(A) = 0$. The structure of $\overline{\sigma}^*$ in the first period is summarized in Figure 14.

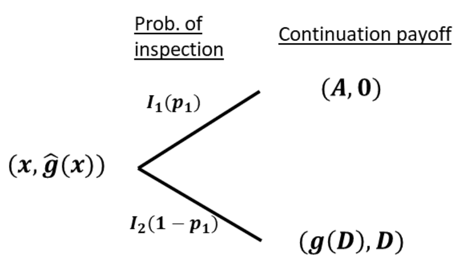

Figure 14: The first period under $\overline{\sigma}^*$.

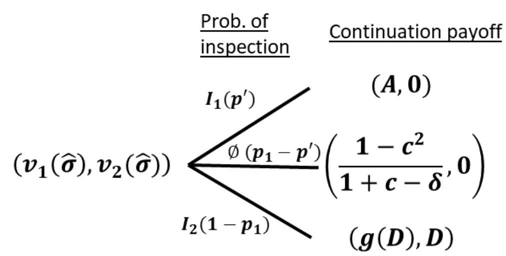

Figure 15: The first period under $\widehat{\sigma}$.

Suppose to the contrary that $A > f(p_1)$. Since f is decreasing, there exists $p' < p_1$ such that $f(p') = A$. Consider the following strategy profile $\widehat{\sigma} = (\widehat{\sigma}_0, \widehat{\sigma}_1, \widehat{\sigma}_2)$ (see Figure 15):

(i) $\widehat{\sigma}_0(I_1) = p'$, $\widehat{\sigma}_0(\emptyset) = p_1 - p'$, and $\widehat{\sigma}_0(I_2) = 1 - p_1$: in the first period the principal reduces the inspection probability for Agent 1 from p_1 to p', and increases the probability of "no inspection" from 0 to $p_1 - p'$. Both agents adhere.

(ii) $\widehat{\sigma}|_{A_1} = \overline{\sigma}^*|_{A_1}$ and $\widehat{\sigma}|_{A_2} = \overline{\sigma}^*|_{A_2}$. The continuation play in the eventuality that one of the agents is inspected in the first period is the same under $\widehat{\sigma}$ and under $\overline{\sigma}^*$.

(iii) $v_1(\widehat{\sigma}|_\emptyset) = \frac{1-c^2}{1+c-\delta}$ and $v_0(\widehat{\sigma}|_\emptyset) = 0$: if no agent is inspected in the first period, the players implement the PPE that yields Agent 1 the payoff $\frac{1-c^2}{1+c-\delta}$ and Agent 2 the payoff 0. Section 3.4 guarantees the existence of such an inspection strategy. Note that by assumption $\frac{1-c^2}{1+c-\delta} < A$.

Since $A = f(p')$ and $D \geq f(1-p_1)$, by Proposition 4 it is optimal for both agents to adhere in the first period, and hence $\widehat{\sigma}$ is a PPE in $\mathscr{E}^{nc}_*$. Agent 2's payoff under $\widehat{\sigma}$ is

$$v_2(\widehat{\sigma}) = 0 + \delta\big(p_1 \cdot 0 + (1-p_1) \cdot D\big) = \widehat{g}(x), \tag{38}$$

and Agent 1's payoff under $\widehat{\sigma}$ is

$$\begin{aligned} v_1(\widehat{\sigma}) &= 0 + \delta\big(p' \cdot A + (p_1 - p') \cdot \frac{1-c^2}{1+c-\delta} + (1-p_1) \cdot g(D)\big) \\ &< 0 + \delta\big(p_1 \cdot A + (1-p_1) \cdot g(D)\big) = x, \end{aligned} \tag{39}$$

where the inequality holds because $A > \frac{1-c^2}{1+c-\delta}$. Thus, the payoffs $(v_1 = v_1(\widehat{\sigma}), v_2 = \widehat{g}(x))$ is implemented by a PPE in $\mathscr{E}_*^{nc}$, and therefore $x = \widehat{g}\big(\widehat{g}(x)\big) \le v_1(\widehat{\sigma}) < x$, a contradiction. This contradiction implies that if $\frac{1-c^2}{1+c-\delta} < A \le \frac{1}{1-\delta}$, then $A = f(p_1)$.

We now study the case $A \le \frac{1-c^2}{1+c-\delta}$. We first prove a technical result.

<u>Step 2</u>: Suppose that $g(B) \in \left(0, \frac{1-c^2}{1+c-\delta}\right]$. Then

$$p \cdot g(B) + (1-p) \cdot g(D) > g\big(p \cdot B + (1-p) \cdot D\big), \quad \forall p \in (0,1). \tag{40}$$

Since g is convex, Eq. (40) holds with weak inequality. In this step we show that the inequality is in fact strict. By construction (as shown in Figure 17),

$$x = \delta \cdot \big(p_1 \cdot g(B) + (1-p_1) \cdot g(D)\big), \tag{41}$$

$$\widehat{g}(x) = \delta \cdot \big(p_1 \cdot B + (1-p_1) \cdot D\big). \tag{42}$$

Since $g\big(\widehat{g}(x)\big) \le \widehat{g}\big(\widehat{g}(x)\big) = x$, by Eqs. (41)–(42), we have

$$g\Big(\delta \cdot \big(p_1 \cdot B + (1-p_1) \cdot D\big)\Big) \le \delta \cdot \big(p_1 \cdot g(B) + (1-p_1) \cdot g(D)\big). \tag{43}$$

Since $x > 0$, we have $\delta x < x$, that is,

$$\delta \cdot \big(p_1 \cdot g(B) + (1-p_1) \cdot g(D)\big) < p_1 \cdot g(B) + (1-p_1) \cdot g(D). \tag{44}$$

Since g is non-increasing, we deduce that

$$g\big(p_1 \cdot B + (1-p_1) \cdot D\big) \le g\Big(\delta \cdot \big(p_1 \cdot B + (1-p_1) \cdot D\big)\Big). \tag{45}$$

By Eqs. (43)–(45),

$$g\big(p_1 \cdot B + (1-p_1) \cdot D\big) < p_1 \cdot g(B) + (1-p_1) \cdot g(D). \tag{46}$$

Since g is (weakly) convex and since there exists $p_1 \in (0,1)$ such that Eq. (46) holds, we have $p \cdot g(B) + (1-p) \cdot g(D) > g\big(p \cdot B + (1-p) \cdot D\big)$ for every $p \in (0,1)$. Indeed, if there were $p_2 \in (0,1)$ such that $g\big(p_2 \cdot B + (1-p_2) \cdot D\big) = p_2 \cdot g(B) + (1-p_2) \cdot g(D)$, then g would not be convex. Denote $Z_1 := p_1 \cdot B + (1-p_1) \cdot D$. If, for example, $p_1 < p_2 < 1$, then the line between $(Z_1, g(Z_1))$ and $(D, g(D))$ does not pass above the graph of g (see Figure 16 for an illustration).

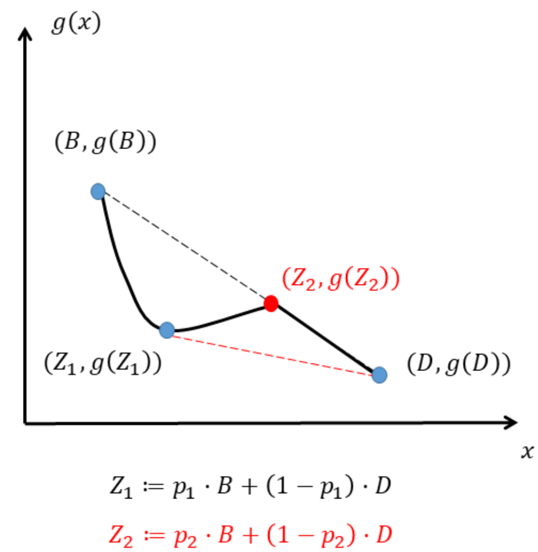

Figure 16: The inequality (40) holds for every $p \in (0,1)$.

Step 3: If $A \leq \frac{1-c^2}{1+c-\delta}$, then $A = f(p_1)$.

Since $A \leq \frac{1-c^2}{1+c-\delta}$ and by Lemma 8, we have $B = g(A) \leq \frac{1-c^2}{1+c-\delta}$. By Proposition 7(v), we have $A = g\big(g(A)\big) = g(B)$. The play in the first period under $\overline{\sigma}^*$ is depicted in Figure 17.

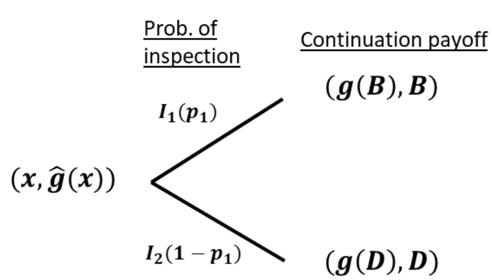

Figure 17: The first period under $\overline{\sigma}^*$.

We are now ready to show that $A = f(p_1)$. Suppose to the contrary that $g(B) = A >$

$f(p_1)$. Since f is decreasing, continuous, and satisfies $\lim_{p \searrow 0} f(p) = \infty$, there exists $p' < p_1$ such that $f(p') = g(B)$.

Consider an alternative PPE, $\widetilde{\sigma}$, where we set the inspection probability for Agent 1 to p', and the remaining probability we assign to no inspection (see Figure 18). Since $g(B) = f(p')$ and $D \geq f(1 - p_1)$, by Proposition 4 both agents adhere in the first period under $\widetilde{\sigma}$. The reader can verify that $\widetilde{\sigma}$ is a PPE in $\mathscr{E}_*^{nc}$ that yields the payoffs $v_1(\widetilde{\sigma}) = x$ and $v_2(\widetilde{\sigma}) = \widehat{g}(x)$.

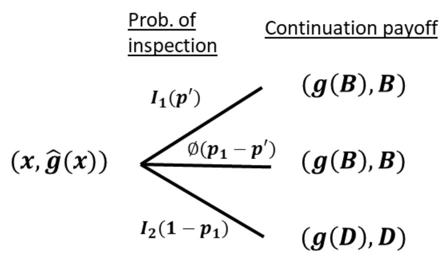

Figure 18: The first period under $\widetilde{\sigma}$.

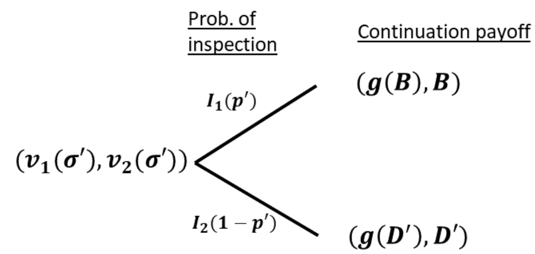

Figure 19: The first period under σ'.

We next construct a new PPE in $\mathscr{E}_*^{nc}$, denoted σ', by adding the probability for no inspection under $\widetilde{\sigma}$ to the inspection of Agent 2 (see Figure 19). Under σ', if Agent 2 is inspected, the players implement the PPE that yields Agent 2 the payoff $D' := \frac{p_1 - p'}{1 - p'} \cdot B + \frac{1 - p_1}{1 - p'} \cdot D$ and Agent 1 the payoff $g(D')$.

Using the same argument as in the proof of Lemma 6, σ' is a PPE in $\mathscr{E}_*^{nc}$ where both agents adhere in the first period, and $v_2(\sigma') = v_2(\widetilde{\sigma}) = \widehat{g}(x)$. Because of the strict convexity of g shown in Step 2, we have

$$g(D') < \frac{p_1 - p'}{1 - p'} \cdot g(B) + \frac{1 - p_1}{1 - p'} \cdot g(D), \tag{47}$$

and hence,

$$v_1(\sigma') < \delta \cdot \big(p' \cdot g(B) + (p_1 - p') \cdot g(B) + (1 - p_1) \cdot g(D)\big) = x. \tag{48}$$

Since $(v_1 = v_1(\sigma'), v_2 = \widehat{g}(x))$ can be implemented by a PPE in $\mathscr{E}_*^{nc}$, the inequality (48) contradicts $x = \widehat{g}\big(\widehat{g}(x)\big)$. This completes the proof that if $A \leq \frac{1 - c^2}{1 + c - \delta}$, then $A = f(p_1)$. $\square$

A.7.2 Proof of Proposition 9

In this section we prove that the one-stage optimal operator θ is convex. We will first define a family $\mathscr{G}$ of functions that contains the function g (Definition 2). We will then study for every $h \in \mathscr{G}$ the one-stage operator θ_h that is defined by the right-hand side of Eq. (24), where w is defined by h instead of g. We will next prove that if the function $h \in \mathscr{G}$ is twice differentiable, then θ_h is convex (Lemma 10). Finally, we will drop the differentiability constraint on h and prove that if $h \in \mathscr{G}$, then θ_h is convex (Lemma 11).

Definition 2. *Let $\mathscr{G}$ be the set of functions $h : [0,\infty) \to [0,\infty)$ that satisfy the following conditions: (i) h is continuous and non-increasing, (ii) h is convex, (iii) $h(x) = 0$ for every $x \geq \frac{1-c^2}{1+c-\delta}$, and (iv) $h(0) \leq \frac{1-c^2}{1+c-\delta}$.*

By Proposition 7, the function g is in $\mathscr{G}$. For every $h \in \mathscr{G}$, let $w_h(p) : [0,1] \to [0,\infty)$ be the function defined as follows:

$$w_h(p) := \begin{cases} p \cdot f(p) + (1-p) \cdot h\big(f(1-p)\big) & \text{if } 0 < p \leq 1, \\ \frac{1}{\delta} + h(0) & \text{if } p = 0. \end{cases} \tag{49}$$

As argued before, the function w_h is continuous at $p = 0$, and strictly decreasing on $[0, \frac{1}{1+c}]$. Therefore, the minimal value of $w_h(p)$ on $[0, \frac{1}{1+c}]$ is achieved at $p = \frac{1}{1+c}$. By Assumption 2, $f(\frac{c}{1+c}) > \frac{1-c^2}{1+c-\delta}$, hence, by the definition of $\mathscr{G}$, we have $h(f(\frac{c}{1+c})) = 0$. By Eq. (6), $f(\frac{1}{1+c}) = 0$, hence the minimal value of $w_h(p)$ on $[0, \frac{1}{1+c}]$ is $w_h(\frac{1}{1+c}) = 0$. Therefore, the inverse function of $w_h(p)$ for $p \in [0, \frac{1}{1+c}]$, denoted $w_h^{-1}(x) : [0, w_h(0)] \to [0, \frac{1}{1+c}]$, is well defined and strictly decreasing.

Denote by $\theta_h(x) : [0,\infty] \to [0,\infty)$ the function

$$\theta_h(x) := \begin{cases} \varphi_h(x) & \text{if } 0 \leq x \leq \delta w_h(0), \\ 0 & \text{if } x > \delta w_h(0), \end{cases} \tag{50}$$

where $\varphi_h(x) : [0, \delta w_h(0)] \to [0,\infty)$ is defined by

$$\varphi_h(x) := \delta \cdot w_h\left(1 - w_h^{-1}\left(\frac{x}{\delta}\right)\right). \tag{51}$$

One can view θ_h as the one-stage optimal operator; that is, the lowest payoff to Agent 2 in PPEs that yield Agent 1 the payoff x, if the continuation payoffs of the agents from the second period and on is given by h. Since w_h is continuous at $p = 0$, and since both f and h are continuous, the function w_h is continuous on $[0, \frac{1}{1+c}]$. This implies that the inverse function w_h^{-1} is continuous on $[0, w_h(0)]$, and therefore the function φ_h is continuous on $[0, \delta w_h(0)]$. Since w_h^{-1} is strictly decreasing on $[0, w_h(0)]$, the function $1 - w_h^{-1}$ is strictly increasing on $[1 - w_h(0), 1]$, hence the composition $w_h\left(1 - w_h^{-1}\left(\frac{x}{\delta}\right)\right)$ is strictly decreasing on $[0, \delta w_h(0)]$. Since $\varphi_h\big(\delta w_h(0)\big) = 0$, the function $\theta_h(x)$ is continuous and non-increasing on

$[0, \infty)$.

Lemma 10. *If the function $h \in \mathscr{G}$ is twice differentiable, then $\theta_h \in \mathscr{G}$.*

Proof. Step 1: The function w_h^{-1} is convex.

Since h is twice differentiable, and since $f(p)$ is twice differentiable on $[0, \frac{1}{1+c}]$, by Eq. (49), the function w_h is twice differentiable on $[0, \frac{1}{1+c}]$. This implies that the function w_h^{-1} is twice differentiable on $[0, w_h(0)]$. For every $x \in \big(0, w_h(0)\big)$, we have

$$(w_h^{-1})''(x) = \frac{-w_h''\big(w_h^{-1}(x)\big) \cdot (w_h^{-1})'(x)}{\Big(w_h'\big(w_h^{-1}(x)\big)\Big)^2} \geq 0. \tag{52}$$

Consequently, the function w_h^{-1} is convex.

Step 2: The function θ_h is convex.

Define the function $p_h : [0, \delta w_h(0)] \to [0, \frac{1}{1+c}]$ by $p_h(x) := w_h^{-1}(\frac{x}{\delta})$ for every $x \in [0, \delta w_h(0)]$. Since w_h^{-1} is strictly decreasing and convex on $[0, w_h(0)]$, it follows that p_h is strictly decreasing and convex on $[0, \delta w_h(0)]$. By Eqs. (49) and (51), and by the definition of p_h and f, for every $x \in [0, \delta w_h(0)]$, we have

$$\begin{aligned}\varphi_h(x) &= \delta \cdot (1 - p_h(x)) \cdot f\big(1 - p_h(x)\big) + \delta \cdot p_h(x) \cdot h\big(f(p_h(x))\big) \\ &= \underbrace{1 - (1+c) \cdot (1 - p_h(x))}_{\text{part E}} + \delta \cdot \underbrace{p_h(x) \cdot h\big(f(p_h(x))\big)}_{\text{part F}}.\end{aligned} \tag{53}$$

Since $p_h(x)$ is convex, Part E is convex. We argue that Part F is convex. For every $x \in \big(0, \delta w_h(0)\big)$,

$$\begin{aligned}\frac{\partial^2\Big(p_h(x) \cdot h\big(f(p_h(x))\big)\Big)}{\partial^2 x} =& p_h''(x) \cdot h\big(f(p_h(x))\big) \\ &+ \underbrace{2 \cdot h'\big(f(p_h(x))\big) \cdot f'(p_h(x)) \cdot (p_h'(x))^2}_{\text{part G}} \\ &+ p_h(x) \cdot h''\big(f(p_h(x))\big) \cdot \big(f'(p_h(x)) \cdot p_h'(x)\big)^2 \\ &+ \underbrace{p_h(x) \cdot h'\big(f(p_h(x))\big) \cdot f''(p_h(x)) \cdot (p_h'(x))^2}_{\text{part H} \leq 0} \\ &+ p_h(x) \cdot h'\big(f(p_h(x))\big) \cdot f'(p_h(x)) \cdot p_h''(x).\end{aligned} \tag{54}$$

Since the functions h and p_h are non-negative, the functions h and f are non-increasing, and the functions p_h and h are convex, except for Part H, all terms in Eq. (54) are non-negative. By Eq. (6), $2 \cdot f'(p_h(x)) + p_h(x) \cdot f''(p_h(x)) = 0$. Hence the sum of Part G and Part H is 0. Therefore, the left-hand-side of Eq. (54) is non-negative, which implies that Part F in Eq. (53) is convex. By Eq. (50), the function θ_h is convex on $[0, \infty)$.

Step 3: $\theta_h(\frac{1-c^2}{1+c-\delta}) = 0$ and $\theta_h(0) \leq \frac{1-c^2}{1+c-\delta}$.

By Eq. (49) and the definition of $\mathscr{G}$, $w_h(0) = h(0) \leq \frac{1-c^2}{1+c-\delta}$. Hence, by Eq. (24), $\theta_h(\frac{1-c^2}{1+c-\delta}) = 0$. Finally, since $1 - w_h^{-1}(0) = \frac{c}{1+c} > 0$ and since w_h is strictly decreasing on the interval $[0, \frac{1}{1+c}]$, we have $\delta \cdot w_h\big(1 - w_h^{-1}(0)\big) \leq w_h\big(1 - w_h^{-1}(0)\big) < w_h(0)$. Therefore, $\theta_h(0) = \delta \cdot w_h\big(1 - w_h^{-1}(0)\big) < w_h(0) \leq \frac{1-c^2}{1+c-\delta}$. This completes the proof of Lemma 10. □

The next lemma drops the differentiability constraint on h in Lemma 10.

Lemma 11. *If the function $h \in \mathscr{G}$, then $\theta_h \in \mathscr{G}$.*

Proof. We start by showing that θ_h is convex. To this end, it is sufficient to prove that φ_h is convex. The idea is to approximate the convex non-differentiable function h by a sequence of convex C^∞-functions $(h^k)_{k\in\mathbb{N}}$. By Lemma 10, we will deduce that for every $k \in \mathbb{N}$, the function φ_{h^k} is convex. We will then prove that the sequence $(\varphi_{h^k})_{k\in\mathbb{N}}$ converges to φ_h, and therefore φ_h is convex.

Step 1: Approximating h by twice-differentiable functions.

Since the function $h \in \mathscr{G}$ is convex, for every $k \in \mathbb{N}$ there exists a convex C^∞-function $h^k : [0, \infty] \to [0, \infty)$ such that $|h^k(x) - h(x)| \leq \frac{1}{k}$ for every $x \in [0, \infty)$. This can be done, for instance, by taking a convolution of h with a smooth non-negative function ξ that is concentrated around 0 (see, e.g., Rockafellar and Wets, 2009, Theorem 2.26).

Step 2: The sequence of functions $(w_{h^k}^{-1})_{k\in\mathbb{N}}$ converges uniformly to w_h^{-1}.

Observe that

$$\begin{aligned} d\Big(w_h\big(w_{h^k}^{-1}(y)\big), w_h\big(w_h^{-1}(y)\big)\Big) &= d\Big(w_h\big(w_{h^k}^{-1}(y)\big), y\Big) \\ &= d\Big(w_h\big(w_{h^k}^{-1}(y)\big), w_{h^k}\big(w_{h^k}^{-1}(y)\big)\Big). \end{aligned} \tag{55}$$

Since the sequence $(h^k)_{k\in\mathbb{N}}$ converges uniformly to h, the sequence $(w_{h^k})_{k\in\mathbb{N}}$ converges uniformly to w_h. Therefore by Eq. (55), the sequence $\big(w_h(w_{h^k}^{-1})\big)_{k\in\mathbb{N}}$ converges uniformly to $w_h(w_h^{-1})$.

Since the function h is continuous on $[0, \frac{1-c^2}{1+c-\delta}]$ and since $h(x) = 0$ for every $x \geq \frac{1-c^2}{1+c-\delta}$, the function h is uniformly continuous on $[0, \infty)$. This implies that the function w_h is uniformly continuous. Since the function w_h^{-1} is continuous on $[0, w_h(0)]$, it is uniformly continuous on $[0, w_h(0)]$. Therefore, as $\big(w_h(w_{h^k}^{-1})\big)_{k\in\mathbb{N}}$ converges uniformly to $w_h(w_h^{-1})$, the sequence $\big(w_{h^k}^{-1}\big)_{k\in\mathbb{N}}$ converges uniformly to w_h^{-1}, as desired.

Step 3: The sequence $(\varphi_{h^k})_{k\in\mathbb{N}}$ converges uniformly to φ_h.

By the triangle inequality and by Eq. (51),

$$
\begin{aligned}
d\Big(\varphi_{h^k}(x), \varphi_h(x)\Big) \leq & \delta \cdot \overbrace{d\Big(w_{h^k}\big(1 - w_{h^k}^{-1}(\tfrac{x}{\delta})\big), w_{h^k}\big(1 - w_h^{-1}(\tfrac{x}{\delta})\big)\Big)}^{\text{part I}} \\
& + \delta \cdot \underbrace{d\Big(w_{h^k}\big(1 - w_h^{-1}(\tfrac{x}{\delta})\big), w_h\big(1 - w_h^{-1}(\tfrac{x}{\delta})\big)\Big)}_{\text{part J}}.
\end{aligned}
\tag{56}
$$

By Eq. (49) and since the function h^k is uniformly continuous, the function w_{h^k} is uniformly continuous. Since $\lim_{k\to\infty} w_{h^k}^{-1} = w_h^{-1}$ uniformly, Part I goes to 0 uniformly. Since $\lim_{k\to\infty} w_{h^k} = w_h$, Part J goes to 0 uniformly. Consequently, $\lim_{k\to\infty} \varphi_{h^k}(x) = \varphi_h(x)$, as desired.

Finally, since the sequence $(\varphi_{h^k})_{k\in\mathbb{N}}$ converges uniformly to $\varphi_h(x)$, it follows that $(\theta_{h^k})_{k\in\mathbb{N}}$ converges uniformly to θ_h. By Lemma 10, $\theta_{h^k}(\frac{1-c^2}{1+c-\delta}) = 0$ and $\theta_{h^k}(0) \leq \frac{1-c^2}{1+c-\delta}$ for every $k \in \mathbb{N}$. Therefore, $\theta_h(\frac{1-c^2}{1+c-\delta}) = \lim_{k\to\infty} \theta_{h^k}(\frac{1-c^2}{1+c-\delta}) = 0$ and $\theta_h(0) = \lim_{k\to\infty} \theta_{h^k}(0) \leq \frac{1-c^2}{1+c-\delta}$. This completes the proof of Lemma 11, and with it the proof of Proposition 9. $\square$

A.7.3 The existence of a minimal solution to Eq. (7), and an algorithm to approximate it

The notations in this section follow Section A.7.2. We first show that the operator $h \mapsto \theta_h$ is monotone.

Lemma 12. *Let h_1 and h_2 be two functions in $\mathscr{G}$ such that $h_1(x) \geq h_2(x)$ for every $x \in [0, \frac{1-c^2}{1+c-\delta})$. Then $\theta_{h_1}(x) \geq \theta_{h_2}(x)$ for every $x \in [0, \infty)$.*

Proof. Since $h_1 \geq h_2$, by Eq. (49), we have $w_{h_1}(p) \geq w_{h_2}(p)$ for every $p \in [0, 1]$. Hence, for every $x \in [0, w_{h_2}(0)]$ we have $w_{h_1}^{-1}(x) \geq w_{h_2}^{-1}(x)$. By Eq. (51) and since w_{h_1} is non-increasing,

$$\begin{aligned}\theta_{h_1}(x) = \delta \cdot w_{h_1}\left(1 - w_{h_1}^{-1}(x)\right) &\geq \delta \cdot w_{h_1}\left(1 - w_{h_2}^{-1}(x)\right) \\ &\geq \delta \cdot w_{h_2}\left(1 - w_{h_2}^{-1}(x)\right) = \theta_{h_2}(x).\end{aligned} \tag{57}$$

The result follows by Eq. (57) and by observing that for $x > w_{h_2}(0)$, we have $\theta_{h_2}(x) = 0$ and $\theta_{h_1}(x) \geq 0$. $\square$

We now provide an iterative procedure that allows us to approximate the minimal solution to Eq. (7).

Lemma 13. *The minimal solution to Eq. (7) exists and is given by* $\lim_{k\to\infty} g_k$*, where the sequence* $(g_k)_{k\in\mathbb{N}}$ *is defined by* $g_0 := 0$ *and* $g_k := \theta_{g_{k-1}}$ *for every* $k \geq 1$.

Proof. The function g satisfies Eqs. (7) and (8), hence these equations have at least one solution. Since $g_0 \in \mathscr{G}$, by Lemma 11 the function g_k is in $\mathscr{G}$ for every $k \in \mathbb{N}$. Lemma 12 implies that the sequence $(g_k)_{k\in\mathbb{N}}$ is increasing and bounded above by the minimal solution to Eq. (7). Therefore, the sequence $(g_k)_{k\in\mathbb{N}}$ converges pointwise to some function, denoted g^*. Since the limit of a non-increasing sequence of convex functions is also non-increasing and convex, and since $g^*(\frac{1-c^2}{1+c-\delta}) = 0$ and $g^*(0)$ is bounded by $\frac{1-c^2}{1+c-\delta}$, the function g^* is in $\mathscr{G}$. Since g^* is continuous and since $(g_k)_{k\in\mathbb{N}}$ is a sequence of monotone functions, Dini's Theorem ensures that the sequence $(g_k)_{k\in\mathbb{N}}$ converges uniformly to g^*.

We now prove that the function θ_{g^k} converges uniformly to θ_{g^*}. Observe that

$$\begin{aligned}&d\Big(w_{g^k}\big(1 - w_{g^k}^{-1}(y)\big), w_{g^*}\big(1 - w_{g^*}^{-1}(y)\big)\Big) \\ \leq &d\Big(w_{g^k}\big(1 - w_{g^k}^{-1}(y)\big), w_{g^*}\big(1 - w_{g^k}^{-1}(y)\big)\Big) + d\Big(w_{g^*}\big(1 - w_{g^k}^{-1}(y)\big), w_{g^*}\big(1 - w_{g^*}^{-1}(y)\big)\Big).\end{aligned} \tag{58}$$

By a similar argument as in the proof of Lemma 11, the sequence $(w_{g^k})_{k\in\mathbb{N}}$ converges uniformly to w_{g^*} and the sequence $(w_{g^k}^{-1})_{k\in\mathbb{N}}$ converges uniformly to $w_{g^*}^{-1}$. Since, in addition, the function w_{g^*} is uniformly continuous, Eq. (58) implies that the sequence of functions $\Big(w_{g^k}\big(1 - w_{g^k}^{-1}(y)\big)\Big)_{k\in\mathbb{N}}$ converges uniformly to $w_{g^*}\big(1 - w_{g^*}^{-1}(y)\big)$. Therefore, the function θ_{g^k} converges uniformly to θ_{g^*}. Consequently, we have $\theta_{g^*} = g^*$, and by Eqs. (50) and (51), the function g^* satisfies Eq. (7). $\square$

Since any solution to Eq. (7) can be implemented by the PPE described in Proposition 8, Theorem 1 follows from Lemma 13.

A.8 Proof of Theorem 2

Let σ^* be an optimal PPE for the principal. We first show that in the first period of σ^* the principal inspects each agent with probability $\frac{1}{2}$. Denote by p_1 the inspection probability for Agent 1 in the first period under σ^*. By Eq. (10) and(11), $v_1(\sigma^*) = \delta \cdot w(p_1)$ and $v_2(\sigma^*) = \delta \cdot w(1-p_1)$. As argued before, under the optimal PPE, $v_1(\sigma^*) = v_2(\sigma^*)$. Since $w(p)$ is decreasing in p and since $\delta \cdot w(p_1) = \delta \cdot w(1-p_1)$ we have $p_1 = \frac{1}{2}$, as desired.

We now study some properties of the optimal PPE σ^*. Since Phase 2 of σ^* has been thoroughly analyzed in Section 3.4, in this section we focus on Phase 1 of σ^*. Recall that both agents adhere in Phase 1.

Proposition 10. *In Phase 1 of σ^* (i) the agent who is inspected (and found adhering) faces lower inspection probability in the next period, (ii) the continuation payoff of the inspected agent increases, and the continuation payoff of the uninspected agent decreases, and (iii) the inspection probability for each agent is in $(\frac{c}{1+c}, \frac{1}{1+c})$.*

Proof. Fix a history in Phase 1 of σ^* with length $t-2$ where no violation has been detected. Denote by p_i^{t-1} the inspection probability agent i faces in period $t-1$ under this history. If agent i is inspected in period $t-1$ and is found adhering, his continuation payoff is $f(p_i^{t-1})$. If $f(p_i^{t-1}) \geq \frac{1-c^2}{1+c-\delta}$ (or equivalently, $p_i^{t-1} \leq f^{-1}(\frac{1-c^2}{1+c-\delta})$), then agent i in period t is inspected with probability $\frac{c}{1+c}$. As argued before, $\frac{c}{1+c} \leq w^{-1}(\frac{x}{\delta})$ for every $x \in (0, \frac{1-c^2}{1+c-\delta})$, hence by Proposition 8(a), $\frac{c}{1+c} \leq p_i^{t-1}$, as desired. If $f(p_i^{t-1}) < \frac{1-c^2}{1+c-\delta}$ (or equivalently, $p_i^{t-1} > f^{-1}(\frac{1-c^2}{1+c-\delta})$), then by Proposition 8(a), agent i in period t is inspected with probability $p_i^t = w^{-1}\left(\frac{f(p_i^{t-1})}{\delta}\right)$. It is left to verify that $p_i^t < p_i^{t-1}$.

Denote by z agent i's payoff if he adheres in period $t-1$. By Eq. (21),

$$z = \delta \cdot \Big(p_i^{t-1} \cdot f(p_i^{t-1}) + (1-p_i^{t-1}) \cdot g\big(f(1-p_i^{t-1})\big)\Big). \tag{59}$$

By Eq. (19), we have $f(p_i^{t-1}) \geq g\big(f(1-p_i^{t-1})\big)$. By Eq. (59) and since $\delta < 1$ we have $f(p_i^{t-1}) > z$. By Proposition 8(a), in period $t-1$ agent i is inspected with probability $p_i^{t-1} = w^{-1}\left(\frac{z}{\delta}\right)$. Since w^{-1} is a decreasing function, we have $w^{-1}\left(\frac{f(p_i^{t-1})}{\delta}\right) < w^{-1}\left(\frac{z}{\delta}\right)$. That is, $p_i^t < p_i^{t-1}$, and Part (i) follows.

We now turn to prove Part (ii). Following the notation of Figure 7, it is sufficient to prove that for every $x \in (0, \frac{1-c^2}{1+c-\delta})$: (a) $f(p_1) > x$, (b) $g\big(f(1-p_1)\big) < x$, (c) $g\left(f(p_1)\right) < g(x)$, and (d) $f(1-p_1) > g(x)$.

As mentioned above, since $f(p_1) \geq g\big(f(1-p_1)\big)$ and since $\delta < 1$, by Eq. (20) we have $f(p_1) > x$. A similar argument shows that $f(1-p_1) > g(x)$ and this proves Parts (a) and (d) of the claim. Now suppose that $g\big(f(1-p_1)\big) \geq x$. Then by Proposition 7(i,v) $f(1-p_1) = g\big(g\big(f(1-p_1)\big)\big) \leq g(x)$, contradicting Part (d). Hence $g\big(f(1-p_1)\big) < x$ and Part (b) follows.[19] Part (c) follows from a similar argument.

Finally, we prove Part (iii). By construction, in every period in Phase 1, the expected payoffs of the two agents are both positive. Denote by v_1^t and v_2^t the payoffs of Agent 1 and Agent 2, respectively, at a certain period t in Phase 1. Since $v_1^t = \delta w(p_1^t)$ and $v_2^t = \delta w(p_2^t)$ (see Eqs. (20)–(21)), we have $w(p_1^t) > 0$ and $w(p_2^t) > 0$. Since $w(p)$ is decreasing in p and since $w(\frac{1}{1+c}) = 0$, we have $p_1^t < \frac{1}{1+c}$ and $p_2^t < \frac{1}{1+c}$. Since $p_1^t = 1-p_2^t$, we have $p_1^t, p_2^t \in (\frac{c}{1+c}, \frac{1}{1+c})$. $\square$

Finally, we take care of the histories that connect the two Phases. When we reach a finite public history where the payoff of one of the agents, say Agent 1, is leaving the range $(0, \frac{1-c^2}{1+c-\delta})$, that is, $\widehat{v}_1 \geq \frac{1-c^2}{1+c-\delta}$, then by the construction of the continuation payoffs $\widehat{v}_2 = g(\widehat{v}_1) = 0$. If $\widehat{v}_1 \leq \frac{1}{1-\delta}$, then the payoffs $(v_1 = \widehat{v}_1, v_2 = 0)$ can be implemented by a PPE following the inspection strategy described in Section 3.4. We next prove that $\widehat{v}_1 \leq \frac{1}{1-\delta}$ must hold. That is, $f\big(w^{-1}(\frac{x}{\delta})\big) \leq \frac{1}{1-\delta}$ for every $x \in (0, \frac{1-c^2}{1+c-\delta})$. Indeed, since $\frac{1-c^2}{1+c-\delta} = \delta \cdot w(\frac{c}{1+c})$ and since w is decreasing in p for $p \in [0, \frac{1}{1+c}]$, we have $w^{-1}(\frac{x}{\delta}) \geq w^{-1}(\frac{1-c^2}{1+c-\delta} \cdot \frac{1}{\delta}) = \frac{c}{1+c}$ for every $x \leq \frac{1-c^2}{1+c-\delta}$. Hence $f\big(w^{-1}(\frac{x}{\delta})\big) \leq f(\frac{c}{1+c})$. Since $f(\frac{c}{1+c}) \leq \frac{1}{1-\delta}$ by Assumption 2, we have $f\big(w^{-1}(\frac{x}{\delta})\big) \leq \frac{1}{1-\delta}$, as desired.

[19] Note that the inequality in Part (b) is strict, hence it does not follow directly from Part (d) and the fact that g is non-increasing.