

Reputation with Endogenous Information

Daniel García-González*

First Draft: March 2011 This Draft: November 2011

Job Market Paper

Abstract

In this paper I study a dynamic game between a principal who has to take a decision every period and an informed agent whose preferences over the decision are his private information. The decision maker has access to a costly monitoring technology that potentially allows her to learn whether the agent has revealed his information truthfully before taking a decision. I show that the principal's incentives to actively monitor are reduced by reputational concerns of the agent. In particular, I show that in any Markov Perfect Equilibrium the principal monitors less and the agent lies more often when the principal has long-run motivations if she believes that the agent is bad with sufficiently high probability. The reason is that when the type of the agent is uncertain, the principal can free-ride on his reputational concern, but once she discovers his type, this benefit is lost.

Keywords: Dynamic Games, Monitoring, Reputation

JEL Classification codes: D82, C73

*I gratefully acknowledge my advisor Antonio Cabrales for many comments and support. I would also like to thank Marco Celentani for his insightful comments and conversations throughout the development of the project. I also thank Juan Pablo Rincón-Zapatero, Angel Hernando-Veciana, Carlos Ponce, Jesper Rudiger, Luis C. Corchón and seminar participants at Universidad Carlos III, ECARES and the 2011 SAET Conference at Faro for their comments that greatly improved both the content and the exposition of the paper. Needless to say, all remaining errors are my sole responsibility. Please send me your comments to dggonzal@eco.uc3m.es

1 Introduction

Many economic organizations rely on the long-run concerns of their members to sustain cooperation and reduce the impact of opportunism. Firms may outperform anonymous markets whenever contract enforcement is difficult or costly, since they can identify those who take inefficient actions¹. For instance, in environments with adverse selection, patient players may be able to build a reputation by mimicking the behavior of cooperative types or by signaling away from defectors². They trade-off short-run gains against future losses that occur whenever their partners identify their behavior as "bad", and get to expect this bad behavior to be maintained in the future.

Therefore, for reputations to emerge, playing partners must be able to identify "bad" behavior, and equilibrium payoffs are effectively determined by the underlying distribution of signals conditional on action profiles. Accordingly, the literature has devoted most of its attention to the identification of conditions in the information structure that result in efficient outcomes. This information structure is, however, taken to be exogenous and independent of the behavior of the individuals. In many instances of interest output or other measures of performance are not readily observable and effort may be needed to identify potential misbehavior. Incentives must be in place for players to monitor intensively their partners, identify and deter defectors. In this paper, I am interested in such situations and analyze the impact of this assumption on the outcomes of the interaction³.

More specifically, I study a very simple, infinite-horizon game between two patient players. In each period, one of the players (the principal) has to choose between two alternatives. The principal is ex-ante indifferent between them. However, his payoff of choosing each alternative depends upon the realization of the state of nature, which is drawn from an iid distribution in every period. The principal is uninformed about this

¹An excellent review is presented in Mailath and Samuelson (2006)

²Seminal contributions are Kreps and Wilson (1982) and Milgrom and Roberts (1982), where a firm may be able to credibly threaten to predate entrants. Benabou and Laroque (1992) study a model in which a financial advisor is able to convey information credibly. Bar-Isaac (2003) studies a market equilibrium with firms providing high quality, even if this is unverifiable.

³To the best of my knowledge there are only two papers with endogenous monitoring in infinite-horizon games. Liu (2011) studies a dynamic game with short-run buyers who may acquire information about previous trades of a seller. Kandori and Obara (2005) who prove a sort of folk theorem in a repeated game where every player may devote effort to monitor others.

draw but she has access to an informed agent, who makes a report conditional on his information. The informativeness of this report depends on the type of the agent who sends it, and this is also his private information. A *good* agent is "committed" to telling the truth every period, while a *bad* agent is strategic and biased towards one alternative. The principal may verify the report of the agent, by engaging in costly monitoring activities. I exclude monetary transfers. Finally, I also assume that the principal cannot decide to fire the agent after any history.

I concentrate the analysis on (stationary) Markov Perfect Equilibria (MPE) where each player's strategy is a measurable function of payoff-relevant components of the history. I show that whenever the principal is sufficiently pessimistic about the type of the agent, in any MPE of the infinite-horizon game she monitors less often and the agent lies more often if the principal is a long-run player (compared with a situation where the principal is myopic). The intuition is simple. The strategic type of the agent has incentives to build a reputation for being truthful, in order to get his preferred decision more often in the future. This reputation is constructed by mimicking a truthful, non-strategic type. Reputation is valuable for the principal since it improves the informativeness of the report and saves on monitoring costs. However monitoring is also costly because it may bring bad news, i.e., the agent may be revealed to be strategic and continuation equilibrium yields lower value to the principal. This implies that for a patient principal monitoring intensity becomes lower and, consequently, the strategic type lies more often in equilibrium.

I argue that these effects may be present in a wide variety of economic interactions, where a combination of adverse selection and moral hazard is present. In those environments, monitoring is a fundamental economic activity and the incentives to undertake it crucially depend on the time-horizon of the relationship. According to the conventional view, in a long-term relationship, agents will acquire more information since they can use it in the future⁴. However, monitoring not only yields information but changes the nature of the relationship and constraints the behavior of the monitored agent. This, in turn, may affect the incentives to monitor and the value of the relationship. Hence, by reducing the effort in monitoring activities, a patient principal can reduce the risk of discovering a bad type and losing a valuable relationship⁵.

⁴See Fama (1985), Diamond (1991)

⁵See Subsection 2.4 for a discussion on the assumptions required for the result.

For instance, relationship financing is commonly assumed to reduce conflicts of interest and increase the effort that banks put in monitoring the actions of their borrowers. This is because banks appropriate the returns of their information in the future and therefore invest more today in acquiring it. This assumes, however, that those returns are positive, which may not be the case in the game I present. Therefore, whenever the mechanism presented in this paper is at work, banks entering in arm's length financing may provide more intensive monitoring than those engaging in relationship finance.

Similar examples abound. Parents may not have the appropriate incentives to monitor whether their kids do their homework, start smoking or get into troubles. Married individuals may face similar problems when monitoring their potentially unfaithful partner. Finally, supranational authorities may struggle to monitor intensively countries under their authority. Notice that in all these situations, the "principal" has a long-run concern and very limited ability to "fire" their partner.

More generally, I argue that this paper provides a rationale for why "fraudulent types" survive longer than external observers would expect in many organizations. As an example consider the case of double spies. Intelligence Agencies are unsure whether their spies are double agents and would like to crosscheck their reports. Nevertheless, if every report is required to be checked, the agent is of no value. Even more, if an agent is known to be a double agent, he is useless and the agency may not have access to similar agents or sources of information. Thus, double agents tend to survive undercover longer than would seem reasonable. This was the case of Juan Pujol, a Spanish spy working for both UK and Germany. Although, from the point of view of historical evidence, he had given enough evidence to the Germans that he was a double agent, they maintained their confidence in his reports until the very end⁶.

In all these situations, long-term incentives for the principal may fail to yield higher incentives to monitor. Therefore, one would expect that in organizations where this problem is severe, short-run incentives are used to encourage information acquisition. Indeed this is the case in the case analyzed in a recent contribution (Hertzberg, Liberti, and Paravisini (2010)). They document agency problems related with the time-horizon of relationships between the loan-officer of a bank and the borrowing firm. They show that, whenever the relationship of a given officer with a firm is about to end, the officer

⁶Indeed, Juan Pujol was key in the cover up of the D-Day. More information about him is to be found in Andrew (2009)

sends more accurate reports, and these reports are more likely to include bad news. In order to alleviate these agency problems, the bank has introduced policies of rapid turnover and task reallocation, and compensation packages depending on short-term objectives.

This model may also be interpreted as an expert-agent type of game. The decision-maker is matched with an informed expert and has to take a sequence of decisions of unknown length. The decision-maker may verify the report of the agent but obtains no more feedback until the sequence ends. Finally, she cannot decide to substitute the expert, or, at least, finds it very costly to do so. This feature is the main departure from previous literature (e.g. Ely and Välimäki (2003)) and can be interpreted in two different ways. First, it may be that the agent has been hired prior to this sequence of decisions and has secure tenure⁷. Alternatively the agent works on his own behalf and, therefore, he cannot be fired⁸. Finally, my model does also contribute to the (broad) literature on auditing games. The novelty from this perspective is the fact that the audited player can build a reputation of being "honest" and, therefore, affect subsequent auditing. We highlight the role of the time-horizon for the auditor under strong incentives.

This paper contributes to the literature on reputation in dynamic games of asymmetric information. In particular, it contributes to a small but important literature that shows the limits of reputation as a way to restrain the behavior of patient players. The first paper to identify a shortcoming of the reputation effect is Morris (2001) who presents a model where an advisor with long-run concerns and no intrinsic bias against any alternative has an incentive to misreport his information in order to look "good". This incentive results in an equilibrium without any useful information provision. In a similar fashion, Ely and Välimäki (2003) study the sequential equilibrium of a dynamic game where the interaction between reputation and the interest of the agent to maintain his employment status, leads to no trade in equilibrium. The no-trade result is originated, as opposed to my paper, by the fact that the principals are short-lived and there is an information externality between them. Each of the principals would like to know whether she is facing a good or a bad agent but no one has incentives to incur in short-run losses in order to learn it. This externality is internalized by a long-lived principal who is able

⁷For instance a newspaper hires a journalist to write a sequence of news pieces. The veracity of these is uncertain for the journal. Firing the worker is costly and may yield a reputational loss for the newspaper.

⁸Newspaper columnists, blogging economists or lobbyists offering advice on public policy are examples of this

to commit and delegates decision-rights to the agent.

Notice that in both Ely and Välimäki (2003) and my paper the time horizon and patience of the uninformed principal play a key role in determining the equilibrium. In Ely and Välimäki (2003), information comes only after trade, so that for a myopic principal such information is of no use. In my paper, however, monitoring yields information ex-ante and a myopic principal may therefore have incentives to experiment. What is more, I show that, under some conditions, she has more incentives to experiment when she is myopic than when she is patient.

Another paper that is close to this one is Sobel (1985). He analyzes a similar game but assumes that output is perfectly and immediately observed by the principal and that the agent is perfectly informed about the state of nature. He concentrates on the conditions for equilibria involving information revelation by the agent at the beginning of the game. At some point, however, the agent will lie and the principal will know it for sure. Once this happens the game ends. Therefore, the goals of his paper are very different from mine.

As mentioned above, I also contribute to a small literature on endogenous information acquisition in games with repeated interaction. Liu (2011) presents a model with a sequence of short-lived buyers who have access to a costly technology that allows them to observe a number of previous trades of the seller. The seller may be a commitment or a strategic type, and buyers acquire a finite number of observations in equilibrium. The model predicts a cyclical behavior in reputation, with strategic sellers creating reputation until no buyer can distinguish them from a commitment type and then "milking it down". In my model past trades are observable and information acquisition concerns current actions. Kandori and Obara (2005) develop a sort of Folk Theorem for games of costly, private monitoring. Players monitor (audit) other players' behavior randomly and during cooperative phases, players both cheat and audit at the same time. There is, however, no reputation effects since there is no asymmetric information about players' types. In any case, my paper is the first to show that dynamic incentives for information acquisition may fail to give lead to more effort, in the presence of reputation effects.

More broadly, my results are also related to a branch of the literature on mechanism design that highlights the benefits of ignorance and uncertainty in the provision of incentives. Holmstrom (1999) presents a signal-jamming model in which agents are

rewarded depending on the market belief about their ability. Ability and effort are substitutes, and so there is an incentive to provide more effort whenever the market belief is less precise. The main difference is that in this model ignorance fosters incentives *through the expected learning* and the dynamic value of information is always positive, and, therefore, the longer the horizon the higher the incentives to acquire information.

The remaining of the paper is organized as follows. In Section 2 I introduce the (infinite horizon) model and analyzes the main trade-offs in the simple static case. At the end of this Section I present a discussion of the main assumptions. In Section 3 I solve the game played by a long-run agent and a myopic principal. This is the benchmark for my main result. In Section 4 I analyze the game played by two patient players. In Section 5 I study the role of commitment for the principal. In Section 6, I apply the model to some of the environments already mentioned. Finally, in Section 7 I conclude the paper. All proofs are contained in the Appendix.

2 The Model

2.1 Environment

Consider a game played by two infinitely-lived players: a principal (she) and an agent (he). The principal has to make a decision every period $d_t \in D = \{a, b\}$. In order to fix ideas, you may think of this decision as whether to invest in a given project. This decision generates a flow payoff that depends on the realization of the state of the nature $\omega_t \in \Omega = \{A, B\}$. I assume, for simplicity, that the ex-ante probability of both events is the same, so that $\Pr(\omega_t = A) = \Pr(\omega_t = B) = \frac{1}{2}$. She does not observe the realization of ω_t but she has access to a costless report by an agent (he) who has received a costless signal s_t correlated with the realization of ω_t . I assume a simple, binary structure for the signal, so that $s_t \in \{S, F\}$, where $s = S$ represents the case where the agent has acquired evidence in favor of alternative A and $s = F$ the case in which he has not. There is no hard evidence in favor of alternative B ⁹. I denote by q be the posterior that the agent holds if he has obtained evidence and $1 - q$ the posterior in case he has not, so that q is the precision of the signal.

⁹Under the interpretation of B as no investment and A as investment. While it is possible to show that a project is profitable, it is very difficult to show that there does not exist a profitable project.

The principal can monitor the agent. By paying a fixed cost c , the principal verifies his report, requiring the agent to disclose the evidence he may have. Following with the banking example, the loan officer may decide to audit the accounts presented by the firm. Notice that given the asymmetric structure, and the fact that the agent may be biased in favor of A , the principal will never monitor after a report recommending B . This simplifies the analysis without affecting the main trade-offs.

The principal is a self-interested, forward-looking, rational player. She discounts the future by $\delta \in [0, 1)$ and maximize their expected discounted stream of utilities. The principal's per-period utility function is

$$v_p(d, \omega, s) = u(d, \omega) - cm$$

where $u_p(d, \omega) = 1$ if $d = a$ and $\omega = A$ or if $d = b$ and $\omega = B$, and $u_p(d, \omega) = 0$ otherwise. Finally cm is the cost of monitoring. If the principal verifies the signal of the agent $m = 1$, and if he does not $m = 0$.

There are two types of agents. The *bad* agent does not share the same preferences over actions than the principal. In particular, the agent may enjoy a private benefit $\lambda > 0$ if action A is implemented. I will assume throughout that $\lambda > 2q - 1$ so that he always prefers action A independently of the state. On the other hand, if the agent reports the truth in every period, we say that the agent is *good*¹⁰. The agent's type is also his private information and is set fixed through time. The principal holds an initial prior belief μ_0 that the agent is good. Finally, as in Sobel (1985) assume that the "importance of the decision" for the agent varies from period to period, so let $x_t \in X$ represent this weight. Assume that $X = [x_l, x_h]$ is distributed according to $F(x)$, iid through time, and is private information for the agent. As a normalization let $\mathbb{E}[x_t] = 1$. You may think x as a device to purify mixed strategies for the agent. Thus, if the agent is bad, his flow utility is

$$v_A(d, \omega) = [u(d, \omega) + \lambda \mathbf{1}_{d=A}] x$$

The following assumption simplify the problem.

¹⁰We assume that the *good* agent is a commitment type. It is straightforward to construct a payoff type choosing the same strategy. The complication is with respect to beliefs off-the-equilibrium path, since commitment types never deviate and payoff types may deviate. For a discussion see Mailath and Samuelson (2006), Chapter 15

Assumption 1 $x_h(1 - 2q + \lambda) < x_lq$

Assumption 1 guarantees that no matter how low is the realization of x_t the agent never prefers to recommend decision B when the signal he has received is S . It is sufficient to simplify both the problem of the principal and the agent without distorting the main intuitions.

The timing of the stage game is as follows. At the beginning of the period the principal holds a belief $\mu_t \in [0, 1]$ that the agent is good. The state of nature and the signals are drawn according to their distributions. The agent sends a report $r_t \in R$, the set of admissible messages¹¹, and the principal decides whether to verify the report or not ($m_t \in \{0, 1\}$). Finally, the principal chooses $d_t \in \Omega$ and the game moves to the following period.

2.2 Strategies and Equilibrium Concept

The main purpose of this paper is to analyze the (Stationary) Markov-Perfect Equilibrium of the infinite-horizon game. This equilibrium concept captures the idea that players have no commitment power and condition their strategies only on payoff-relevant information. Define a (private) history for the agent at the beginning of period t ,

$$h_t^a = \{(s_k, r_k), (m_k, d_k)\}_{k < t}.$$

Let \mathcal{H}_t be the space of such histories. The *bad* agent observes the realization of his private information (s_t, x_t) and decides which report to send. Therefore, a pure strategy for the agent is

$$r_t : h_t^a \times \{S, \emptyset\} \times X \rightarrow \Omega.$$

Recall that the good agent always tells the truth. A public history when the principal has to decide whether to monitor is $h_t^p = \{(r_k), (m_{k-1}, d_{k-1})\}_{k \leq t}$. Define \mathcal{H}_t^p to be the space of such histories. A pure monitoring strategy for the principal

$$m : h_t^p \rightarrow \Omega$$

¹¹Since the *good* type only receives information about the state, if the *bad* type pretends to be a *good*, she may send at most two messages in equilibrium. Thus, without loss of generality R contains two elements. Hence, I shall assume that $R = \Omega$

Stationary Markov strategies map public histories into actions with the constraint that for every two different histories h_t, h'_t generating the same posterior probability that the agent is good, the equilibrium play is the same. Thus, Markov strategies are measurable with respect to the posterior probability after every history. In what follows I describe these Markov strategies and present the concept of Markov Perfect Equilibrium.

The principal has to choose his actions using only payoff-relevant information. Given that both states are equally likely his choice of d_t is trivial. She will choose $d_t = a$ if and only if her posterior belief that the state $\omega = A$ exceeds $\frac{1}{2}$. For this reason, in the following, I will consider exclusively on the principal's choice of whether to monitor or not after a report A . Let $\sigma(\mu) \in [0, 1]$ be the monitoring mixed strategy for a principal who holds a belief μ that the agent is good and receives a report A .

The bad agent, on the other hand, will misreport his information whenever it is in his best interest to do so. Thus, let $r(x; \mu) \in \{0, 1\}$ be the probability that a bad agent, who has not received information and weights this period with x , and is believed to be good with probability μ by the principal sends report A .

Denote $V^k : [0, 1] \rightarrow \mathbb{R}_+$ be the value function of the principal who has received a report $k \in \{A, B\}$ as a function of the current belief. Similarly, denote $U : [0, 1] \times \{S, F\} \times X \rightarrow \mathbb{R}_+$ be the value function of the (bad) agent who has obtained a signal s_t and weights the current period with value x_t and is believed to be good with probability μ_t .

Definition 2 *A Markov-Perfect Equilibrium is a pair of Markov strategies (σ^*, r^*) and a belief μ satisfying;*

- i) for every m in the support of σ^* , $m \in \arg \max_{s' \in \{0, 1\}} \mathbb{E}_\mu [v_P(s'; r^*, \mu)] + \delta \mathbb{E}_{\mu, r^*} [V^k(\mu')]$
- ii) $r = 1$ if and only if $A \in \arg \max_{r' \in \{A, B\}} \mathbb{E}_{\sigma^*} [u(r'; x, B)] + \delta \mathbb{E}_{\sigma^*, x, s} [U(\mu'; x, s)]$
- iii) μ' is computed from μ using Bayes Rule¹²

Finally, some more notation would be useful. Let $\bar{F}(y) = 1 - F(y)$. Define $\phi_k(\mu; x)$ for $k \in \{A, B\}$ to be the posterior probability that the agent is good conditional on a report

¹²From the point of view of the principal, every report of the agent has positive probability as long as $\mu \in (0, 1)$. If μ is degenerate, it will remain constant in the future. Therefore, off-the-equilibrium path beliefs are irrelevant.

k if the prior was μ and the bad agent uses strategy x . With some abuse of notation I will write

$$\phi_k(\mu) = \int \phi_k(\mu; r(x; \mu)) dF(x)$$

to be the equilibrium posterior probability from the point of view of the principal. Let π_μ be the expected payoff for the principal after a recommendation to choose A of an agent believed to be good with probability μ , if the principal is not monitoring.

3 Static Game

3.0.1 Known Type

It is useful to begin the analysis looking at the case in which the principal knows at the outset the type of the agent¹³. In this case, if the agent is good, there is no monitoring and the relationship yields an expected discounted payoff $\bar{V} = q$ for both players. On the other hand, if the agent is biased, there is a continuum of equilibria. Let p^* be the probability with which the agent reports A without supporting evidence in every period. This is the perceived probability of the principal, but the agent may use x as a randomization device. In every equilibrium, the principal will only monitor with probability 1 after an A -report whenever

$$c \leq p^* \left[q - \frac{1}{2} \right]$$

and never monitor otherwise. On the other hand, for the agent it is a (weakly) dominant strategy to report A after failing to obtain any signal since

$$1 - q + \lambda > q.$$

Let p_0 satisfy

$$c = \frac{1}{2} p_0 [2q - 1]$$

Assumption 3 $p_0 < 1$

The assumption guarantees that monitoring occurs in equilibrium.

¹³Formally, the type of the agent does also include the collection of realizations of the uncertainty (x, s) . However, for lack of a better word, I will devote the word type to refer to his preference profile.

A particular class of equilibria that will turn out to be useful are threshold equilibria. For every $x^* \in [x_l, \bar{F}^{-1}(p_0)]$ there is an equilibrium in which the agent reports A if $x \geq x^*$ and the principal monitors if A is reported. Each of these equilibria yield the same payoff for the agent, but not for the principal. To see this, define V_{x^*} as the value for the principal in an equilibrium where the agent uses x^* as his threshold strategy.

$$V_{x^*} = q - \frac{1}{2}c [\bar{F}(x^*) + 1].$$

For the remaining of the paper, I shall consider only undominated strategies (see discussion below). It is obvious that reporting B is (weakly) dominated for the agent, in the symmetric information game, so that I set $x^* = x_l$. Thus, the value for the principal is

$$\underline{V} = q - c$$

3.0.2 Unknown Type

We consider now the case in which the principal is uncertain about the type of the agent. Suppose that the probability she attaches to the event that the agent is good is μ . Then, she will monitor the agent if and only if $\mu \leq \bar{\mu}$, where $\bar{\mu}$ is defined in

$$c = \frac{1}{2}(1 - \bar{\mu})[2q - 1]$$

and so the principal gets a value

$$V(\mu) = \mu V_g(\mu) + (1 - \mu)V_b(\mu)$$

with $V_g(\mu) = \frac{1}{2}q + \frac{1}{2}(q - c) > V_b(\mu) = q - c$ if $\mu \leq \bar{\mu}$ and $V_g(\mu) = q > V_b(\mu) = \frac{1}{2}$ if $\mu > \bar{\mu}$

Finally, the value for the principal of learning the type of the agent, in this static framework is

$$\begin{aligned} \Pi &= \mu(\bar{V} - V_g(\mu)) + (1 - \mu)(\underline{V} - V_b(\mu)) \\ &= \begin{cases} \frac{1}{2}\mu c & \text{if } \mu \leq \bar{\mu} \\ (1 - \mu)(q - \frac{1}{2}) & \text{if } \mu > \bar{\mu} \end{cases} \geq 0 \end{aligned}$$

Obviously in this simple game, the principal is willing to pay to know the type of the

agent. In the rest of the paper we show that this incentives need not be present in the infinite-horizon version of the model, since information *does* change the continuation equilibrium.

3.1 Discussion

Although in a very stylized way, this model is able to capture the main features of the strategic interactions presented above. Think, for instance, in the game played by the bank and an entrepreneur. The entrepreneur may come, every period, with a potentially profitable project. He may also "save for better times" if there is no profitable project at hand. The bank would like to finance only profitable projects, but ascertain whether a project is profitable is costly. The bank would like to use the informational content of the decision of the entrepreneur to offer a project, in order to save resources from costly monitoring. This is possible, if the entrepreneur builds a reputation of being "trustworthy", by behaving as a "good" type.

As mentioned in the Introduction, the main feature of the environment is that the agent is paired with the principal and there is no exit option. This assumption may seem restrictive, but I feel that it captures quite well the environment described. First, in family relations exit is usually infeasible. Moreover, in those relations that allow for exit, it is usually very costly since the agent may be very hard to substitute. Our results would not change if verifiable evidence of misbehavior is required for dismissing the agent, and would be qualitatively similar if the principal dismisses the agent whenever his belief on his type falls below some threshold, since the good agent is committed to telling the truth and not strategic.

I also assume, that players do not learn their payoffs as they play. I make this assumption both for simplicity and realism. Again, think about the problem of a bank financing an entrepreneurial project. Projects require investments at the beginning and are unlikely to yield profits early on. For instance, advertisement, brand positioning and the like require big investments but their contribution to output is difficult to ascertain until they are already sunk. The bank finds it difficult to disentangle the individual contribution of each investment to the total revenue of the firm. Nonetheless, all results will go through as long as output is not perfectly observed (or $q < 1$) since information about payoffs is received after decisions are made and, therefore, does not affect the strategic

interaction in the interim stage¹⁴. The limiting case in which output is immediately and perfectly observed and $q = 1$ is analyzed in Sobel (1985).

Finally, I assume that stage-game payoffs are independent of the current belief that the market has. This is a simplification, and it is unlikely to hold in real-world examples. It requires, for instance, that, *a priori*, both alternatives are equally likely and that the importance of the period x_t is distributed independently of the reputation of the agent. Removing any of these restrictions will not affect the results, as long as the stage-game payoff of the agent is monotonically increasing in the belief that the principal holds about his type. This is satisfied in all the examples introduced. For instance, in the credit market example, interest rates will be decreasing in the belief that the lender has about the agent's type and, therefore, the payoff of the agent would be increasing in this belief.

Solution Concept: In the remaining of the paper I focus on undominated equilibria. In particular, I rule out weakly dominated strategies in subgames that involve perfect information and monitoring with probability one. Notice that these equilibria rely on dominated strategies for the agent but are weakly preferred for the principal. Alternatively, I could have assumed that the monitoring technology fails to give any information with some positive, albeit small, probability. Undominated equilibria will be the only equilibria surviving in such perturbed model. In this sense, the equilibria I look at is the limit with respect to a class of imperfect monitoring technologies, while the rest are not.

I also restrict attention to Markov strategies. I believe that Markov strategies are the natural way to introduce lack of commitment and impose sequential rationality in infinitely-lived relations. Notice that a strategy is Markov if it is only measurable with respect to payoff-relevant information in every period in time, so that for every two histories originating the same beliefs Markovian strategies specify the same distribution over actions. In particular, the principal is unable to commit to future punishments conditional on deviations that do not trigger a change in his beliefs about the type of the agent. This kind of punishments require players to coordinate either through correlating devices or their history, which may be difficult in many situations of interest. Notice that my aim in this paper is to compare the situation where both the agent and the principal are long-run players with one in which the principal is short-run. Equilib-

¹⁴Details of this extension are presented in the Appendix. Modified proofs are available from the author upon request.

ria involving future punishments and rewards renders the comparison between the two environments difficult.

Markov Perfect Equilibria are formed by simple strategies. This simple structure is well-suited for studying interactions within organizations, where decision-making follows standardized protocols and routines requiring the use of limited information. Organizations tradeoff the benefits of more accurate decision-making versus the costs of information transmission and acquisition (Arrow (1974)). Therefore, strategies requiring infinite recall and the use of unbounded information are not useful for understanding the equilibrium behavior of real-world organizations.

Monitoring Technology: I assume that monitoring is both public and perfect. This means that the agent knows for certain whether the principal has verified his report and the result of such verification. This assumption is justified in many of the environments described above, e.g. relationship finance, where the agent is asked to disclose his information.

The assumption is made mainly for the sake of simplicity. First, if monitoring were private but perfect, the state would not only include the current belief of the principal but also, the second-order belief of the agent. In general, this second-order belief is distribution over the set of possible beliefs that the principal may have (depending on her past actions). In my simple model, however, the principal is not willing to distort her decision in order to conceal her information, since the very reason to monitor is to make better decisions. Therefore, the second-order belief will be degenerate and no additional insights will be obtained.

Second, one could assume that the monitoring technology is imperfect, in the sense, that it yields no information with some positive probability. This is the technology used in Diamond (1991) among others. My results are robust to this extension and, moreover, our equilibrium selection criterion picks the limit equilibria of the class of games indexed by these technologies where the probability of failure goes to zero. But allowing monitoring to be both imperfect and private would affect the results non-trivially and is left for future research.

4 A myopic Principal

I will present now the results for the benchmark model where the principal is myopic while the agent is patient and discount the future with a factor δ . The strategic agent maximizes, then, the present expected discounted sum of his stream of payoffs. After the agent observes his signal s_t and the importance of the period x_t , his value function is

$$U(\mu; x_t, s_t) = \max_r x \mathbb{E}_\sigma [v_A(d, s) | r] + \delta \mathbb{E}_{x, \sigma} (U(\mu) | r)$$

with

$$U(\mu) = \frac{1}{2} \int [U(\mu; x, a) + U(\mu; x, b)] dF(x)$$

where the agent takes expectations over the (mixed) strategy of the principal and the future realizations of the uncertainty. The principal will never monitor with probability 1 since in that case the agent will not be lying with positive probability. On the other hand, if the agent plays a threshold strategy x^{15} , the principal will randomize, and monitor with positive probability only if

$$c = (1 - \mu) \bar{F}(x^*) (2q - 1)$$

This determines x as part of the equilibrium for those μ such that $\sigma(\mu) > 0$. Call $x_s(\mu)$ the solution to the equation (in case it exists). We need to find $\sigma \in (0, 1)$ to make $x_s(\mu)$ optimal

$$[x_s(\mu)(1 - q + \lambda) + U(\phi_A(\mu))] (1 - \sigma) + \sigma [qx_s(\mu) + U(0)] = qx_s(\mu) + U(\phi_B(\mu))$$

The following Proposition summarizes this equilibrium.

Proposition 4 *In any equilibrium of the game with one long-run player, the principal monitors with positive probability if and only if μ is small enough, the agent lies with positive probability for every μ . The value for the agent is increasing in μ and discontinuous at $\mu = 0$*

The typical play is as follows. If the principal is confident that the agent is good, she does not monitor and the agent fully chooses the future path of posterior beliefs. In this case, he restrains his behavior because if the belief goes down, the principal will

¹⁵In the long-run case, every equilibrium is outcome equivalent to a threshold equilibrium.

monitor and thus his influence in the decision will be lower. If the agent is bad, beliefs will follow a stochastic path with negative drift and, at some point, the principal starts to monitor with positive probability. In this case, the agent's payoff from suggesting option A is decreased but he still lies with positive probability, until he is discovered. Importantly, the agent does always restraint his behavior as compared with the one-shot interaction case if $\mu \in (0, 1)$, in the sense that he cooperates and tells the truth with positive probability. This increases the equilibrium payoff of the principal. Finally, notice that this allows the principal to reduce his monitoring (if the agent chooses B the principal does not bear the cost of monitoring). This is the key observation for what follows.

5 Patient Principal

The problem for the agent is similar than the one presented above, given that he takes σ as given. The problem of the principal changes in two dimensions. First, the principal now has an incentive to learn the information of the agent to make better decisions and thus increases his sustained payoff. But second, he faces a dynamic inconsistency problem that will lead him to reduce the speed of learning and thus, the reputational concern of the agent. Characterizing this second force is the main aim of this paper.

Whenever the agent chooses action A , the principal can choose between verifying his signal or not. If he does, his future discounted expected payoff is

$$V_{A,1}(\mu) = q - c + \delta [(1 - \mu) \bar{F}(x(\mu))V(0) + (\mu + F(x(\mu))(1 - \mu))V(\mu)]$$

while if he does not verify

$$V_{A,0}(\mu) = \pi_\mu + \delta V(\phi_A(\mu))$$

Notice that the principal randomizes if and only if

$$V_{A,1}(\mu) = V_{A,0}(\mu)$$

In the Appendix we show that both V and U as well as the policy functions are monotone and well-behaved except around $\mu = 0$. Now we are ready to state the main result of the paper. I show that the principal would like to avoid learning the type of the agent, if the

agent is bad, and, hence, will optimally reduce his monitoring intensity at some beliefs.

Proposition 5 *In any equilibrium of the game in which both players have reputational concerns, there exists some $\mu_1 > 0$, such that if $\mu_t < \tilde{\mu}$, $x_s(\mu_t) > x^l(\mu_t)$ and $\sigma_s(\mu_t) > \sigma^l(\mu_t)$*

Another way to see this result is considering the value function as a weighted average of the two conditional value functions $V(\mu | t)$ for $t \in \{g, b\}$. It is clear that $V(\mu) = \mu V(\mu | g) + (1 - \mu)V(\mu | b)$ where μ defines the equilibrium strategies and the type determines the payoffs. Notice that the martingale property of the beliefs is such that the dynamic value of information is pinned down by a weighted average of the change in those two value functions as information arrives. In particular $V(\mu | g)$ is decreasing in μ since monitoring is redundant in this case, and takes the value $\frac{q}{1-\delta}$ for all μ such that $\sigma(\mu) = 0$. However, $V(\mu | b)$ is not maximized at $\mu = 0$ since the reputation effect is absent. In particular, $V(\mu | b)$ is bounded away from its supremum at $\mu = 0$. Intuitively, if the principal is indifferent of whether to monitor or not, conditional on receiving a report recommending action A , the informativeness of such report must be bounded away from zero. Therefore, as the belief goes to zero, the probability of a lie must remain bounded away from 1. This generates a positive value for the principal and leads to the result.

The following proposition shows that this may increase the value for the agent for every $\mu \in (0, 1)$

Proposition 6 *There exists some $c_1 > 0$ such that if $c \geq c_1$ then the agent lies more often when matched with a long-run principal ($x^s(\mu) \geq x^l(\mu)$) and gets a higher payoff ($U^l(\mu) \geq U^s(\mu)$) for every $\mu \in (0, 1)$*

As mentioned above, these are the main results of the paper and show how a principal facing an agent who is likely to have conflicting interests with himself, faces a trade-off when deciding whether to investigate him. Absent any dynamic concerns, she may find it optimal to monitor him closely, in order to avoid being cheated, but whenever she has a long-term concern, she will reduce her monitoring intensity and try to free-ride on the dynamic concern of the agent. This will in its turn lead to a lower incentive for the agent to restrain his behavior and thus lower payoffs for the principal.

6 The Role of Commitment

In this Section I explore commitment of the principal. I will first allow the principal to commit at the beginning of the period to a monitoring intensity. This implies that the principal need not be indifferent whether to monitor or not conditional on a report recommending A and that she internalizes the effect of a higher monitoring intensity in the behavior of the principal. In short, the principal becomes the stackleberg leader. Her problem becomes

$$V^C(\mu) = \max_{\sigma \in [0,1]} p_\mu(\sigma) q + (1 - p_\mu(\sigma)) [\sigma(q - c) + (1 - \sigma)\pi_\mu(\sigma)] + \delta \mathbb{E}_{\sigma,p} [V^C(\phi(\mu))] \quad (1)$$

where $p_\mu(\sigma)$ is the expected probability of a report B when the belief is μ and the principal committed to a strategy σ . Notice that $V^C(\phi(\mu))$ may be not differentiable. However, it is easy to see that the maximization problem is globally concave¹⁶. The following proposition shows that the solution to this problem $\sigma^C(\mu) \geq \sigma^l(\mu)$ for every μ .

Proposition 7 *If the principal has long-run concerns and is able to commit to a monitoring strategy before the agent reports, she will monitor more intensively.*

The proof uses the fact that at $\sigma^l(\mu)$ the principal is ex-post indifferent between monitoring or not, but ex-ante has an incentive to increase monitor to lead to a higher accuracy of the report and save on monitoring costs. Therefore, if $\sigma^l(\mu) > 0$, $\sigma^C(\mu) > \sigma^l(\mu)$. Finally notice that this implies that whenever $\sigma^S(\mu) \geq \sigma^l(\mu)$ the principal may find it useful to delegate monitoring to a third party and give him short-term incentives.

The preceding discussion suggests that there is a role to commitment. But, can a principal who fully commits to a monitoring intensity get his first-best payoff? Ely and Välimäki (2003) show that in an environment similar to ours, if the principal is a long-run player and she can commit to a (stochastic) participation-rule as a function of the history, social first-best is attainable (under the average discounted expected payoff criterion). The idea of their mechanism is to "promote" the agent after every period with positive probability, so that monitoring stops altogether. The probability is chosen to separate types, so that a *bad* type is not willing to incur in the cost of mimicking the

¹⁶Informally, increasing σ reduces both the probability of a report A but also the difference in continuation values (less informative signals.)

good type and prefers to deviate in the first period. Thus, two things are required in this mechanism. First, commitment is needed since after the first period, the agent knows whether the agent is good or bad. Second, and more importantly for my analysis, it is required that deviating today yields high payoffs for the *bad* type. In particular, the decision-rule today must follow his advice -although this advice is also known to be wrong.

In my model, unlike in Ely and Välimäki (2003), however, the decision is made by the principal. This implies that if the agent expects the principal to verify his report with high enough probability, the agent will pretend he is a good type and tell the truth. Hence, the principal will be unable to tell them apart fast enough and fail to get her first-best payoff. To understand the result, let $V^i : \mathcal{H} \rightarrow \mathbb{R}$ be the value that the principal can get after a given history $h \in \mathcal{H}$ where she to play against a type $i \in \{g, b\}$, and let $\mu(h)$ be the implied belief of the principal.

Proposition 8 *For every $\mu(h) \in (0, 1)$, if $V^g(h) = \frac{q}{1-\delta}$ then $V^b(h) < \frac{q-c}{1-\delta}$.*

The argument is simple. Since the bad agent may pretend to be a good type, in order to give incentives for revelation the principal must give him higher continuation payoff. But absent any commitment to follow the advice, the principal can only give higher payoffs by reducing her monitoring intensity. Lower monitoring implies delay in information and more cheating, so that her payoff is bounded away from first best. On the other hand, if she were to monitor more, her payoff on the good agent will be lower since monitoring is inefficient in such a case.

As it is clear from the proof the IC constraint of the agent is binding and this implies that revelation will be slow enough for the principal to obtain his first best value. If the principal could commit also to delegate the decision to the agent, this would soften this constraint and increase the value of the principal. Thus, an important insight of this paper is that delegating monitoring and decision-making may be superior to delegating monitoring and relying on communication.

7 Applications

Organizational Design: As mentioned in the Introduction, this paper offers novel insights about the use of different institutions in many organizations. First, we provide a new rationale for delegated monitoring. In the framework presented here, monitoring is not only required to provide incentives but also to acquire information about the type of the agent. The value of such information is not always positive, since it reduces future incentives. Therefore, delegating monitoring to a third party who does not internalize this negative value of information may increase monitoring and provide better incentives now.

Second, we provide an intuitive explanation for the pervasive use of short-term incentives in many organizations where information is revealed over time and cooperation is required for efficiency (Hertzberg, Liberti, and Paravisini (2010)). Agents with long-run concerns may have low incentives to monitor agents with whom they interact, since such monitoring may bring bad news and decreases the continuation payoff.

Third, in my model, even if monitoring is delegated to a third party, the principal may want to delegate the decision to the agent. Under delegation, the *bad* agent is willing to reveal his type more quickly and first best may be approximated. If the agent is to report to the decision maker, however, the agent will conceal his information and preclude first best. Hence, I give an additional rationale for the widespread use of delegation, even when information is potentially verifiable.

Relationship Finance: As mentioned in the Introduction, one of the main applications of the model presented here is the banking industry. In particular, my model is very similar to that in Diamond (1991), who is the first to study the interplay between monitoring and reputation but under the assumption that the monitor is short-lived. Among other things, he discovers a "paradox of monitoring" whereby cheap and effective monitoring fails to create incentives since lenders face a lack of commitment to cut defaulters off the credit market. I extend the analysis to incorporate long-run motives for the monitor. I show that, if the credit score of the firm is sufficiently low and the bank has a lack of commitment, dynamic interactions may not be able to solve such problems¹⁷.

¹⁷There is one additional difference between both frameworks. He carefully models the credit market and, therefore, allows interest rates to be endogenous. Interest rates will be monotonic in the belief that

More generally, my model sheds light on the use of relationship finance versus arm's-length or directly placed financing. Relationship finance is widely understood as a way to overcome informational asymmetries between borrowers and lenders. Banks accumulate information about their borrowers over time, both through communication and monitoring. The value of communication depends on the incentives of the incentives of the borrower to maintain a reputation and the threat of monitoring. Communication saves on information acquisition costs and, therefore, is valuable for the bank. Our model predicts that banks starting long-term relationships will use less external or formal monitoring and rely more on communication and soft information than borrowers engaging in arm's-length financing. This is consistent with empirical evidence presented in Kano, Uchida, Udell, and Watanabe (2011) who show that firms benefit most from bank-borrower relationships when they do not have audited financial statements. Similarly, Blackwell and Winters (1997) finds that banks less frequently monitor firms with whom they have closer relationships. Even more, many studies have found that banks lending to insolvent firms are more likely to execute their guarantees or liquidate assets if they do not have a long-run relationship with the borrower. The previous literature has relied on two different theoretical frameworks to understand this evidence.

First, banks may suffer a *soft-budget constraint problem* so that they are unable to commit ex-ante not to refinance inefficient projects since the initial investment is sunk Dewatripont and Maskin (1995). This lack of commitment, together with an agency problem, is most likely to be present in bank-oriented economies where banks perform relationship financing and have more "at stake". Even so, these refinancing decisions are sequentially efficient, in the sense that the expected value of those projects at the time of the refinancing decision exceed their liquidation value. Nevertheless, this argument is insufficient to explain "zombie-lending", i.e. the concession of debts to inefficient firms at a subsidized rate, e.g. Caballero, Hoshi, and Kashyap (2008), since these rates are computed using all the available information at the time of the decision.

The second common explanation relies on an *efficiency-wage* type of argument. Banks may lower their monitoring effort because they provide rents to entrepreneurs (lower interest rates) and, therefore, alleviate the moral hazard problem. This hypothesis requires that the entrepreneur is easily substitutable for the bank, and that the bank has

the market has about the agent, so that the per-period payoff of the agent will be endogenously increasing in this belief. Adding this monotonicity would not change the results of my model, but in the interest of simplicity I ignored it.

commitment to do so.

According to the hypothesis developed in my paper, however, the terms of refinancing offered by banks to long-term borrowers will be the less sensitive to the financial situation behavior than that of firms with whom they have short-run relationships, and more so, whenever these firms are in financial distress. Banks face the risk of losing the relationship they created with a firm if it goes under and this reduces the incentives they have to monitor their activities. This, in turn, decreases the incentives for firms to devote resources to efficient activities and, thus destroys value for the bank.

Nonetheless, it would be a mistake to conclude for this argument that relationship financing is not profitable for banks. As in Diamond (1991), monitoring generates valuable information about the type of the agent, which has a future value only in the case of relationship financing. But in our model, the behavior of the agent is also affected by the current belief of the bank about his type. This change in behavior decreases the incentives to monitor for the principal and it eventually overcomes the positive effect of information generation. This is the trade-off analyzed in the current paper.

8 Conclusions

In this paper I have presented a simple dynamic framework to understand the interaction between an uninformed decision-maker and an informed agent who may be bad in favor of one of the alternatives. The decision-maker may actively monitor the agent but lacks commitment and is bound to suffer losses in case that she discovers that the agent is bad. In this environment, we show that dynamic concerns for the agent increase the payoff of the principal, but dynamic concerns for the principal may decrease her payoff, if her belief about the type of the agent is sufficiently pessimistic.

The key observation of the paper is as follows. In any equilibrium with positive value of reputation, the principal must be (at most) indifferent between auditing or not. Therefore, the agent will not lie with probability 1 in any equilibrium. This increases the average value for the principal as long as the agent has not been found to be bad with certainty. Monitoring more intensively increases the probability that this happens, and, therefore, reduces the future value of the principal. This implies that the dynamic value of monitoring becomes negative when the probability that the agent is bad is sufficiently

high, and, therefore, the long-run principal monitors less intensively. This increases the value of the agent and gives him more incentives to cheat.

We have also shown that even if the principal is able to commit to a fully contingent plan of future monitoring intensity, he can not approximate her first-best payoff. This is because the principal retains the decision rights even if monitoring was delegated, and once she discovers the type of the agent, has no incentive to bias his decision. Therefore, the incentive for the agent to reveal his type is diminished and the agent has to monitor for a sufficiently long period of time that his payoff is bounded away from first-best.

Finally, we have presented a number of different environments where this insight seems to be present, and discussed the implications of our findings for organizational design, highlighting the benefits of delegation and short-term incentives and the shortcomings of communication. As for future research, it seems interesting to exploit other organizational arrangements that may allow the principal to commit and limit wrongdoing by the agent.

References

- ANDREW, C. (2009): *The Defense of the Realm: The Authorized History of MI5*. Allen Lane, London.
- ARROW, K. (1974): *The Limits of Organization*, The Fels Lectures on Public Policy Analysis. W.W.Norton and Company, New York.
- BAR-ISAAC, H. (2003): "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies*, 70(2), 231–251.
- BENABOU, R., AND G. LAROQUE (1992): "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *The Quarterly Journal of Economics*, 107(3), 921–58.
- BLACKWELL, D. W., AND D. B. WINTERS (1997): "Banking Relationships and the Effect of Monitoring on Loan Pricing," *Journal of Financial Research*, 20(2), 275–89.
- CABALLERO, R. J., T. HOSHI, AND A. K. KASHYAP (2008): "Zombie Lending and Depressed Restructuring in Japan," *American Economic Review*, 98(5), 1943–77.

- DEWATRIPONT, M., AND E. MASKIN (1995): "Credit and Efficiency in Centralized and Decentralized Economies," *Review of Economic Studies*, 62(4), 541–55.
- DIAMOND, D. W. (1991): "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," *Journal of Political Economy*, 99(4), 689–721.
- ELY, J. C., AND J. VÄLIMÄKI (2003): "Bad Reputation," *The Quarterly Journal of Economics*, 118(3), 785–814.
- FAMA, E. F. (1985): "What's different about banks?," *Journal of Monetary Economics*, 15(1), 29–39.
- HERTZBERG, A., J. M. LIBERTI, AND D. PARAVISINI (2010): "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation," *Journal of Finance*, 65(3), 795–828.
- HOLMSTROM, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *The Review of Economic Studies*, 66(1), 169–182.
- KANDORI, M., AND I. OBARA (2005): "Endogenous Monitoring," *Unpublished Manuscript*.
- KANO, M., H. UCHIDA, G. F. UDELL, AND W. WATANABE (2011): "Information verifiability, bank organization, bank competition and bank-borrower relationships," *Journal of Banking & Finance*, 35(4), 935–954.
- KREPS, D. M., AND R. WILSON (1982): "Reputation and imperfect information," *Journal of Economic Theory*, 27(2), 253–279.
- LIU, Q. (2011): "Information Acquisition and Reputation Dynamics," *Review of Economic Studies*, forthcoming.
- MAILATH, G., AND L. SAMUELSON (2006): *Repeated games and reputations: long-run relationships*. Oxford University Press.
- MILGROM, P., AND J. ROBERTS (1982): "Predation, reputation, and entry deterrence," *Journal of Economic Theory*, 27(2), 280–312.
- MORRIS, S. (2001): "Political Correctness," *Journal of Political Economy*, 109(2), 231–265.
- SOBEL, J. (1985): "A Theory of Credibility," *Review of Economic Studies*, 52(4), 557–73.

A Appendix A: Main Results

Lemma 9 *A Markov Perfect Equilibrium in undominated strategies exists*

Proof. An MPE exists by standard arguments. Below, I show that equilibrium outcomes and strategies must be monotonic in the state, and then construct the equilibrium by finding the fixed point of the best response mapping with augmented payoffs. To construct the equilibrium fixed point, notice that, for $\mu \in (0, 1)$ best replies are continuous and monotone. It suffices then to see that the value functions are also well-defined and monotone. Assume first that σ is decreasing in μ ¹⁸. Let $T : \mathcal{F}([0, 1]) \rightarrow \mathcal{F}([0, 1])$, where $\mathcal{F}([0, 1])$ is the set of bounded functions from $[0, 1]$ into the reals. Let u be the expected value for the agent Let $\Delta = (1 - \sigma)(1 - 2q + \lambda)$. The value function for the agent is defined as the fixed point of T

$$2T(f(\mu)) = u \int x dF(x) + \delta [(1 - \sigma)f(\phi_A(\mu)) + \sigma(f(\mu))] + \\ + \max_{x^*} \left[\Delta \int_{x^*}^{x_h} x dF(x) + \delta \bar{F}(x^*) ((1 - \sigma)f(\phi_A(\mu)) + \sigma(f(\mu))) + \delta F(x^*) f(\phi_B(\mu)) \right]$$

which clearly satisfies both monotonicity and discounting, so that the value function is well-defined by Blackwell's Theorem and is decreasing. Now, let p_μ be decreasing, the value function for the agent is the fixed point of $S : \mathcal{F}([0, 1]) \rightarrow \mathcal{F}([0, 1])$.

$$S(f(\mu)) = \frac{1}{2} p_\mu [q + \delta f(\phi_B(\mu))] + \\ \frac{1}{2} (2 - p_\mu) \max_{\sigma} \left[\sigma \left(\pi - c + \delta \left(\frac{p_\mu}{1 + p_\mu} f(0) + \frac{1}{1 + p_\mu} f(\mu) \right) + (1 - \sigma) (\pi_\mu + f(\phi_A(\mu))) \right) \right]$$

which does again satisfy Blackwell's Sufficiency Condition. In what follows we show that indeed the only equilibria satisfy these properties. ■

Lemma 10 (Monotonicity) *(U, V) are increasing, σ is decreasing and x is increasing*

This proof proceeds in various steps.

Claim 11 *Let $\mu > 0$. $V(\mu) > V(\phi_A(\mu))$*

¹⁸If the principal is infinitely impatient, this is guaranteed by construction. If he is patient, we give sufficient conditions below, for this to be true.

Proof. First, assume that $\sigma > 0$. Then, for a contradiction assume there exists some μ for which $V(\mu) \leq V(\phi_A(\mu))$. Notice that

$$\pi - c + \delta \left[\frac{p_\mu}{1 + p_\mu} V(0) + \frac{1}{1 + p_\mu} V(\mu) \right] = \pi_\mu + \delta V(\phi_A(\mu))$$

so that

$$\begin{aligned} \pi - c + \delta \left[\frac{p_\mu}{1 + p_\mu} V(0) + \frac{1}{1 + p_\mu} V(\mu) \right] &\geq \pi_\mu + \delta V(\mu) \\ \pi - c + \delta \frac{p_\mu}{1 + p_\mu} V(0) &\geq \pi_\mu + \delta \frac{p_\mu}{1 + p_\mu} V(\mu) \end{aligned}$$

which is true only for $\mu = 0$.

Assume now that $\sigma = 0$. Since

$$\begin{aligned} V(\mu) &= \pi_\mu + \delta \left[\frac{1}{2}(1 + p_\mu)V(\phi_A(\mu)) + \frac{1}{2}(1 - p_\mu)V(\phi_B(\mu)) \right] \\ V(\mu) &\geq \frac{2\pi_\mu + \delta(1 - p_\mu)V(\phi_B(\mu))}{2 - \delta(1 + p_\mu)} \end{aligned}$$

Notice that if $V(\phi_B(\mu)) < V(\mu)$, we have that $V(\phi_B(\mu)) \leq V(\phi_A(\mu))$ and we know that $V(\phi_B(\mu)) \leq V(\phi_A(\phi_B(\mu)))$ so substituting μ for $V(\phi_B(\mu))$ for a number of times, we will get as close as we want to $\mu = 1$ and we have a contradiction since $\lim_{\mu \rightarrow 1} V(\mu) = V(1) = \frac{\pi}{1 - \delta} \geq V(\mu)$. So then assume that $V(\phi_B(\mu)) \geq V(\mu)$, then

$$V(\mu) \geq \frac{\pi_\mu}{1 - \delta} = \frac{\pi}{1 - \delta} - \frac{2q - 1}{2(1 - \delta)}(1 - \mu)\bar{F}(x(\mu))$$

To see that this cannot be true if $\sigma = 0$, notice that $\bar{F}(x(\mu)) > \frac{1}{2}$ (to see why, just consider the simplified problem of the agent who is not allowed to go below the cutoff of the principal; maybe add a restriction of F so that this holds). It is clear then that

$$\begin{aligned} V(\mu) &\leq \mu \frac{\pi}{1 - \delta} + (1 - \mu) \left[\frac{\pi}{(1 - \delta)} + \frac{\pi - c}{2(1 - \delta)} \right] \\ &= \frac{\pi}{1 - \delta} - (1 - \mu) \frac{c}{2(1 - \delta)} \end{aligned}$$

Hence a necessary condition is that

$$\begin{aligned}\frac{2q-1}{2(1-\delta)}(1-\mu)\bar{F}(x(\mu)) &\geq (1-\mu)\frac{c}{2(1-\delta)} \\ (2q-1)\bar{F}(x(\mu)) &\geq c\end{aligned}$$

but Assumption 1 guarantees that there exists $x_1 > 0$ for which $c = \bar{F}(x_1)(2q-1)$ and we know that this is the static equilibrium, where the value for the principal is much lower. Since for x higher than this, $c > (2q-1)\bar{F}(x(\mu))$ we reach a contradiction and $V(\mu) > V(\phi_A(\mu))$ ■

Claim 12 V is increasing

Proof. For all μ such that that $\sigma = 0$ we have that

$$V(\mu) = q - (1 - \frac{p_\mu}{2})(2q - 1) + \delta \left[\frac{1 + p_\mu}{2}V(\phi_A(\mu)) + \frac{1 - p_\mu}{2}V(\phi_B(\mu)) \right]$$

where $p_\mu = (1 - \mu)\bar{F}(x(\mu))$.

Since $\sigma > 0$, V increases if p_μ decreases. We can write $V(\mu)$ as

$$\begin{aligned}V(\mu) &= (1 - \frac{1}{2}p_\mu)q + \frac{1}{2}p_\mu(1 - q) + \delta \left[\frac{1}{2}(1 + p_\mu)V(\phi_A(\mu)) + \frac{1}{2}(1 - p_\mu)V(\phi_B(\mu)) \right] \\ &= q - \frac{1}{2}(1 + p_\mu)c + \delta \left[\frac{1}{2}p_\mu V(0) + \frac{1}{2}V(\mu) + \frac{1}{2}(1 - p_\mu)V(\phi_B(\mu)) \right]\end{aligned}$$

Now, suppose that there exists two different beliefs such that $\mu_H > \mu_L$ and $V(\mu_H) = V(\mu_L) = V$

$$\begin{aligned}V &= q - \frac{1}{2}(1 + p_H)c + \delta \left[\frac{1}{2}p_H V(0) + \frac{1}{2}V + \frac{1}{2}(1 - p_H)V(\phi_B(\mu_H)) \right] \\ &= q - \frac{1}{2}(1 + p_L)c + \delta \left[\frac{1}{2}p_L V(0) + \frac{1}{2}V + \frac{1}{2}(1 - p_L)V(\phi_B(\mu_L)) \right]\end{aligned}$$

Hence, we need that

$$\frac{1}{\delta}(p_H - p_L)c = (p_L - p_H)V(0) + (1 - p_L)V(\phi_B(\mu_L)) - (1 - p_H)V(\phi_B(\mu_H))$$

Suppose first that $p_H \geq p_L$. This requires that $V(\phi_B(\mu_L)) \geq V(\phi_B(\mu_H))$, where $\phi_B(\mu_H) >$

$\phi_B(\mu_L)$. Now, notice that this would lead to a decreasing sequence of $\{\mu_{ki}; \mu_{ki} > \mu_{k-1i}\}_{k=1,2}$ but we know that $V(\mu)$ is strictly increasing when $\mu \rightarrow 1$. Hence, a contradiction. Suppose then that $p_H < p_L$. Notice that

$$q - c + \delta \left[\frac{p_H}{1 + p_H} V(0) + \frac{1}{1 + p_H} V \right] = \frac{q(1 - p_H) + 1}{1 + p_H} + \delta V(\phi_A(\mu_H))$$

and

$$q - c + \delta \left[\frac{p_L}{1 + p_L} V(0) + \frac{1}{1 + p_L} V \right] = \frac{q(1 - p_L) + 1}{1 + p_L} + \delta V(\phi_A(\mu_L))$$

so that

$$\delta [V - V(0)] \frac{p_L - p_H}{(1 + p_H)(1 + p_L)} = \frac{2q(p_L - p_H)}{(1 + p_H)(1 + p_L)} + [V(\phi_A(\mu_H)) - V(\phi_A(\mu_L))]$$

Since $V - V(0) \geq \lim_{\mu \rightarrow 0} V(\mu) = \frac{\epsilon}{1 - \delta}$, for δ large enough, this requires that

$$V(\phi_A(\mu_H)) - V(\phi_A(\mu_L)) < 0$$

which again will not be true since $V(\mu)$ is strictly increasing for μ small enough. ■

Claim 13 *If $\sigma(\mu) > 0$ then U is non-decreasing and $(1 - \mu)\bar{F}(x(\mu))$ is non-increasing.*

Proof. Notice that if σ is weakly decreasing, the value of the agent is monotone. This follows from a standard choice argument, since the agent is free to choose his report and facing a lower monitoring intensity yields higher payoffs.

Now, for every μ such that $\sigma(\mu) > 0$ assume that V is non decreasing. There exists an open set of such μ by Assumption 2. From this it is easy to see that $(1 - \mu)\bar{F}(x(\mu))$ must be non-increasing if σ is non-increasing. To see why, assume for a contradiction that $(1 - \mu)\bar{F}(x(\mu))$ is increasing in some open interval (μ_0, μ_1) . The principal must be indifferent between monitoring and not monitoring at every μ such that $\sigma(\mu) > 0$. Therefore

$$V_{A,1}(\mu) = V_{A,0}(\mu) \forall \mu \in (\mu_0, \mu_1)$$

but then the principal's value at the beginning of the period could be written as

$$V(\mu) = p_\mu q + (1 - p_\mu)\pi_\mu + \delta [p_\mu V(\phi_B(\mu)) + (1 - p_\mu)V_{A,0}(\mu)]$$

But $(1 - p_\mu)$ and π_μ are increasing in μ so that either V is decreasing at μ or $V(\phi_B(\mu))$ is decreasing at μ . Finally notice that since at higher μ the agent lies more often, a report of B is more informative, thereby contradicting the monotonicity of V . ■

Claim 14 *Every equilibrium is monotone.*

Proof. To see this notice that in non-monotone equilibria p_μ is increasing and σ is increasing (at least in some open interval). But then in that open interval U must be decreasing. But in such a case (write the expression down)

$$U(\phi_B(\mu)) < U(\phi_A(\mu))$$

But then there exists a sequence of values for agents whose belief converges to zero that is increasing. But the limit of all these sequences is the infimum of all the values that are consistent with the principal randomizing. Therefore we have a subsequence converging from below to the infimum of all the equilibrium values. This is a contradiction. ■

Proof of Proposition 4. Let this equilibrium pair (σ_s, x_s) . Notice that in equilibrium the current period's profit of monitoring is equated to that of not monitoring if $\sigma_s > 0$. In particular, since the principal is short-lived, information has no future value for the principal. We shall see that there is a region for which this value would be negative, thus triggering a lower required threshold x_s and hence, a lower level of monitoring.

To show that U is well-defined, we use most of the analysis in Bar-Isaac except that we do not have continuity. In particular, given that the static equilibrium leads to monitoring for sure, the payoff of the agent is discontinuous at $\mu \rightarrow 0$. To see why, notice that if it were continuous at $\mu = 0$

$$\begin{aligned} [x_s(\mu) + \lambda + U(\phi_A(\mu))] (1 - \sigma) + \sigma [1 - x_s(\mu) + U(0)] &= 1 - x_s(\mu) + U(\phi_B(\mu)) \\ (1 - \sigma) [x_s(\mu) + \lambda] &= (1 - \sigma) [1 - x_s(\mu)] + \epsilon(\mu) \end{aligned}$$

where

$$\epsilon = U(\phi_B(\mu)) - (1 - \sigma)U(\phi_A(\mu)) + \sigma U(0)$$

Hence in any equilibrium with $\sigma < 1$, $p^M(\mu) \rightarrow x_1^*$ and so $c > (1 - \mu) \left[\pi_{\frac{1}{2}} - \pi_p \right]$ by Assumption 2, a contradiction. Hence, for every $\mu < \mu_l$, for some $\mu_l > 0$, $\sigma = 1$ if U is

continuous. However, in that case, there is another type μ' reachable from any $\mu > 0$ such that $\sigma(\mu') < 1$ and so $U(\mu') > U(0) + (1 - \sigma) \left[\lambda \int_{\frac{2q-1}{\lambda}}^{\infty} x dF(x) + 1 - 2q \right]$. But if $\sigma = 1$, there are no current gains from deviating and there are future gains, so that not deviating becomes optimal and the condition is not satisfied. Hence $\epsilon(\mu) \geq \tilde{\epsilon} > 0$ for every $\mu > 0$, so that $x_s(\mu) \leq x_0$ and $\sigma < 1$.

To conclude we show that there exists $\tilde{\mu}$ satisfying $\sigma_s(\mu) > 0 \Leftrightarrow \mu < \tilde{\mu}$. We know that there is no monitoring when $\mu = 1$ and full monitoring with $\mu = 0$. Now assume that $\mu \in (0, 1)$. First notice that if $\mu_1 < \mu_2$ and $\sigma_2 > 0$

$$q - c = q + (1 - \mu_2)\bar{F}(x(\mu_2))(1 - 2q)$$

Now assume that $\sigma_1 < \sigma_2$. Then, $\bar{F}(x(\mu_1)) > \bar{F}(x(\mu_2))$, so that

$$q + (1 - \mu_1)\bar{F}(x(\mu_1))(1 - 2q) > q - c \Rightarrow \sigma = 1$$

a contradiction. Similarly, if $\sigma_1 = 0$, we know that

$$\begin{aligned} q + (1 - \mu_2)F(x(\mu_2)) &< q + (1 - \mu_2)\bar{F}(x(\mu_1))(1 - 2q) \\ &< q + (1 - \mu_2)\bar{F}(x(\mu_1))(1 - 2q) \\ &< q - c \end{aligned}$$

so that $\sigma_2 = 0$ ■

Proof of Proposition 5. The argument is very simple. As $\mu \rightarrow 0$, $V(\phi_A(\mu))$ becomes arbitrarily close to $V(\mu)$ so that the indifference condition becomes

$$q - c + \delta \left[(1 - \mu)\bar{F}(x(\mu))V(0) + (\mu + F(x(\mu))(1 - \mu))V(\mu) \right] = \pi_x + \delta V(\phi_A(\mu))$$

so that

$$[q - \pi_x] - c = \delta \left[(1 - \mu)\bar{F}(x(\mu)) (V(0) - V(\phi_A(\mu))) + (\mu + F(x(\mu))(1 - \mu))\epsilon \right] < 0 \quad (2)$$

Notice that if $x(\mu) = x_s(\mu)$, LHS must be equal to zero, but it is clear that it is negative as long as $V(\phi_A(\mu)) > V(0)$. To see that this holds, notice that it will not hold if for $\mu \rightarrow 0$ either $\sigma = 1$ or $\sigma = 0$, since in the former case the agent is always indifferent about his own reputation and in the later he is never discovered to be bad. Assumption 1 rules out

the second one, while it is easy to see that if $\sigma = 1$, the agent can increase his reputation at no cost and get to μ close enough to 1 so that $\sigma = 0$ ¹⁹, and getting some benefit from it. Thus, $\sigma(\mu) = 1$ cannot be part of the equilibrium for any $\mu > 0$.

Hence (2) implies that $\pi_x > \pi_{x_s}$. This implies that $x(\mu) < x_s(\mu)$. Intuitively, the dynamic value of the information becomes negative and so the equilibrium level of cheating must increase compared with the case in which the dynamic value is zero.

Now we show that there must exist some open set of beliefs for which the monitoring is lower in the case where the principal has dynamic concerns. Suppose not, then $U(\mu) \leq U(\mu_s)$, but clearly in such a case, $x(\mu) < x_s(\mu)$ only if $\sigma_s > \sigma$, thus contradicting the initial assumption. Hence the result obtains. ■

Proof of Proposition 6. Let $r = \frac{(1-\bar{\mu})F(x_s(\bar{\mu}))}{1+(1-\bar{\mu})F(x_s(\bar{\mu}))}$ where $\bar{\mu}$ is such that

$$V(\bar{\mu}) = rV(0) + (1-r)V(\phi_A(\bar{\mu}))$$

Finally, let $c_1 = r(1-2q)$, which by Assumption 1 is guaranteed to exist.

For any two equilibrium strategies for the principal σ^1, σ^2 such that $\sigma^1 \geq \sigma^2$ for every μ . Clearly $U^2(\mu) \geq U^1(\mu)$ since any strategy for the agent leads to at least higher payoff in the second equilibrium. In Lemma 10 we show that in any equilibrium σ is non-increasing and continuous. Notice that the way we constructed c_1 , at $\bar{\mu}$, $\sigma_s = 0$, and $x(\mu) \leq x_s(\mu)$ since in the putative equilibrium $\sigma_s \geq \sigma$, so that $U^s(\mu) \geq U(\mu)$ and the indifference condition is now tighter for the agent. Now, clearly we have that

$$(1-\bar{\mu})\bar{F}(x(\bar{\mu})) \leq r$$

so that $\sigma > 0$ if and only if $\sigma_s > 0$. To conclude the proof notice that for $\mu > \bar{\mu}$, $V(\mu) < rV(0) + (1-r)V(\phi_A(\mu)) \leq p_\mu V(0) + (1-p_\mu)V(\phi_A(\mu))$ and hence, $\sigma < \sigma_s$.

If $U(\mu) \geq U^s(\mu)$ it also holds that $V(\mu | b) \leq V^s(\mu | b)$ where $V^s(\mu | b)$ is the average discounted payoff of a given principal with short-term concerns playing against a bad type. ■

¹⁹Note that if $\frac{1-\mu}{1-\delta} [2q-1] < c$, monitoring is strictly dominated.

B Appendix B: Commitment

Proof of Proposition 7. The proof is very simple. Given concavity, all we have to show is that at $\sigma^l(\mu)$ and if $\sigma^l(\mu) > 0$, the principal has a strict incentive to increase her monitoring intensity. This argument is valid since an increase in the monitoring intensity at some histories will lead to a lower value of reputation, and therefore, the agent will tell the truth with lower probability for the same monitoring intensity, further affecting the result. Now, the objective function in (1) can be written as

$$\max_{\sigma} \quad p(\mu)q + (1-p(\mu)) [\mathbb{E}_{\sigma,p}[u(\mu)] + \sigma(q-c - \mathbb{E}_{\sigma,p}[u(\mu)])] + \\ \delta \left[p(\mu)V^C(\phi_B(\mu)) + (1-p(\mu)) \left[\sigma \left[\frac{r}{1-p}V^C(0) + \frac{1-p-r}{1-p}V^C(\mu) \right] + (1-\sigma)V^C(\phi_A(\mu)) \right] \right]$$

Increasing σ from $\sigma^l(\mu)$ yields a benefit

$$B = \Delta p \left[q - u(1-\sigma) - \sigma(q-c) + \delta \left[V^C(\phi_B) - \left[\sigma \left[\frac{r}{1-p}V^C(0) + \frac{1-p-r}{1-p}V^C(\mu) \right] + (1-\sigma)V^C(\phi_A(\mu)) \right] \right] \right] \\ + \Psi(V^C, \mu) + (1-p) \left[q - c - u - \delta \left[\sigma \left[\frac{r}{1-p}V^C(0) + \frac{1-p-r}{1-p}V^C(\mu) \right] + (1-\sigma)V^C(\phi_A(\mu)) \right] \right]$$

where $\Psi(V^C, \mu)$ measures the change in the continuation values that is driven by a change in the informativeness of reports through an adjusted μ . The principal must be indifferent between monitoring or not at $\sigma^l(\mu)$ and therefore

$$q - c - \mathbb{E}_{\sigma,p}[u(\mu)] = \delta \left[\sigma \left[\frac{r}{1-p}V^C(0) + \frac{1-p-r}{1-p}V^C(\mu) \right] + (1-\sigma)V^C(\phi_A(\mu)) \right]$$

but then the condition becomes

$$B = \Delta p \left[\delta \left[V^C(\phi_B) - \left[\sigma \left[\frac{r}{1-p}V^C(0) + \frac{1-p-r}{1-p}V^C(\mu) \right] + (1-\sigma)V^C(\phi_A(\mu)) \right] \right] + \Psi(V^C, \mu) \right]$$

To see that this is always positive at σ^l notice that $\Delta p > 0$ and the term in brackets is also positive (since it is the difference in value between a report B and a report A). The last term is negative but can be shown to be second order. Therefore $\sigma^C(\mu) > \sigma^l(\mu)$ whenever $\sigma^l(\mu) > 0$. ■

Proof of Proposition 8. The argument proceeds in two steps. I first show that there cannot be a fixed time $t^* > 1$ such that, the type of the agent has been revealed for sure, and the principal attains V^g if the agent is good. In the second step I extend the result to stochastic time limits.

To see the first claim, notice that at time $t^* - 1$, and independently of the way in which the type is revealed, a strategic agent who has not been revealed must choose action A independently of his current information. In particular, it must be the case that

$$\begin{aligned} x_l [(1 - q + \lambda)(1 - \bar{\sigma}) + \bar{\sigma}q] + \delta U^b &\geq x_l q + \delta U^g \\ &= x_l q + \frac{\delta}{1 - \delta} \bar{x} \left(\frac{1}{2} + \lambda \right) \end{aligned}$$

while $U^b = \frac{q}{1 - \delta}$. Hence

$$x_l (1 - 2q + \lambda)(1 - \bar{\sigma}) + \delta \frac{q}{1 - \delta} \geq \frac{\delta}{1 - \delta} \bar{x} \left(\frac{1}{2} + \lambda \right) \quad (3)$$

since $\frac{1}{2} + \lambda > q$ by assumption, for every $\bar{\sigma}$, there is a δ^* , such that if $\sigma > \delta^*$, the condition is violated.

Now, for the second claim, notice that, if the IC constraint is to be satisfied, the maximum intensity after every history h is

$$x(h)(1 - 2q + \lambda)(1 - \sigma(h)) + \delta \left[(1 - \sigma(h))U(h \sqcup a) + \sigma(h)U^b \right] \geq \delta \bar{x} U(h \sqcup b)$$

Since $U^b < U(h \sqcup a) \leq U(h \sqcup b)$, this gives an upper bound on $\sigma(h)$ for every $x(h)$. Hence, let $\bar{\sigma}(h)$ be such an upper bound. Since $\bar{\sigma}(h) \in (0, 1)$ and $x \in [x_l, x_h]$, both variables define a probability measure over histories. Let

$$\Omega^t = \{h_t \in \mathcal{H}^t \text{ s.t. } \mu \in \{0, 1\}\}$$

Let $p(\Omega^t)$ be the implied probability of such event. All we have to show is that

$$C = (1 - \delta) \sum_{h^t \in \Omega^t} p(h^t) \sum_{t=0} \delta^t \sigma(h^t) c > 0 \quad (4)$$

since in that event $V^g < \frac{q}{1-\delta} - \epsilon$ for some $\epsilon > 0$, or $V^b < \frac{q-c}{1-\delta}$ since

$$(1-\delta) \sum_{h^t \in \Omega^t} p(h^t) \sum_{t=0} \delta^t \bar{F}(x(h^t))(2q-1) > 0$$

. We have

$$\begin{aligned} p(\Omega^t) - p(\Omega^{t-1}) &= \left[\frac{1}{2} \int \sigma(h^t) F(x(h^t)) dp(h^t \mid h^{t-1} \notin \Omega^{t-1}) \right] [1 - p(\Omega^t)] \\ &\leq \frac{1}{2} [1 - p(\Omega^t)] \sup_{h^t \in \Psi(\mathcal{H}^{t-1} \setminus \Omega^{t-1})} \bar{\sigma}(h^t) F(x(h^t)) \end{aligned}$$

But because of the monotonicity property in the value function²⁰.

$$\sup_{h^t \in \Psi(\mathcal{H}^{t-1} \setminus \Omega^{t-1})} \bar{\sigma}(h^t) F(x(h^t)) \leq \bar{\sigma}(h^*) F(x(h^*))$$

where at h^* , $U(h^* \sqcup a) = U(h^* \sqcup b)$, and hence using (3)

$$x(h^*) = \delta \frac{\sigma(h^*)(U(h \sqcup r) - U^b)}{(1-2q+\lambda)(1-\sigma(h^*))}$$

Clearly this condition imposes a bound on $F(x(h))$ for every $\sigma(h)$. Let that bound be $\bar{F}(\sigma)$. Notice that the bound depends both explicitly and implicitly on δ through the continuation values of the agent. To further simplify the problem, let us fix the average expected continuation monitoring intensity from the point of view of the agent to be some $\sigma' < 1$ so that

$$U(h \sqcup r) - U^b \geq \frac{1}{1-\delta\sigma'} [1-2q+\lambda] \bar{x} \frac{1}{2}$$

And thus

$$x(h^*) \geq \delta \frac{\sigma}{1-\sigma} \frac{\bar{x}}{1-\delta\sigma'} \frac{1}{2}$$

As $\delta \rightarrow 1$

$$x(h^*) \geq \frac{\sigma \bar{x}}{2(1-\sigma)(1-\sigma')}$$

First assume that $\sigma \geq \frac{2(1-\sigma')}{1+2(1-\sigma')}$. In such a case, $F(x(h)) \leq \frac{1}{2}$. Thus $\bar{\sigma}(h^t) F(x(h^t)) < \frac{1}{2}$ for

²⁰In the optimal solution to the Full Commitment problem the value function is monotone as long as the value for the agent is monotone. This value, in turn, is monotone as long as the monitoring intensity of the agent is non-increasing. A simple local indifference argument shows that this must be true in the FC case.

every h^t and

$$p(\Omega^t) - p(\Omega^{t-1}) < \frac{1}{4} [1 - p(\Omega^t)]$$

while $\sigma \geq \frac{2(1-\sigma')}{1+2(1-\sigma')} > 0$. Thus (4) becomes

$$\begin{aligned} C &\geq (1-\delta) \sum_t \left(\frac{3\delta}{4}\right)^t c \\ &= (1-\delta) \frac{4}{4-3\delta} c \end{aligned}$$

Now if $\sigma < \frac{2(1-\sigma')}{1+2(1-\sigma')}$, since σ was the upper-bound of all possible σ we have that $\sigma \geq \sigma'$, and so $\sigma \leq 2(1-\sigma)^2$, so that $\sigma \leq \frac{1}{2}$

$$p(\Omega^t) - p(\Omega^{t-1}) \leq \frac{1}{4} [1 - p(\Omega^t)]$$

and hence (4) does also hold. ■

C Supplemental Appendix: Information about Output

In this Appendix I extend the main results in the paper to the possibility that the principal observes its own payoff at the end of the period. I shall assume throughout that $q < 1$, but all the results will again hold unchanged if $q = 1$ but the principal only observes a noisy signal about output. Notice that the strategic interaction takes place in the interim stage, so that information about payoffs only affects the continuation values in a Markov Perfect Equilibrium. Therefore, all we have to show is that the arguments derived in Appendix A guaranteeing that an MPE exists and value functions are monotone still hold under this modification.

Let $\tilde{\phi}(\mu; 1), \tilde{\phi}(\mu; 0)$ the updated equilibrium beliefs after observing a payoff $h_t = 1$ and a payoff $h_t = 0$ respectively if the decision was $d_t = A$ and μ is the conditional belief of the principal after observing a report A and not monitoring. Notice that after a decision $d_t = B$ there is no further update because the expected conditional payoff of

that decision is independent of the type of the agent. It is straightforward that

$$\tilde{\phi}(\mu; 1) = \mu \frac{q [1 + (1 - \mu) \bar{F}(x(\mu))]}{q + (1 - q)(1 - \mu) \bar{F}(x(\mu))} \geq \mu$$

and that

$$\tilde{\phi}(\mu; 0) = \mu \frac{(1 - q) [1 + (1 - \mu) \bar{F}(x(\mu))]}{(1 - q) + q(1 - \mu) \bar{F}(x(\mu))} \leq \mu$$

Therefore, the continuation value for the agent after lying with a signal of B if there is no monitoring becomes

$$\mathbb{E}[U(A; B, \mu)] = (1 - q)U(\tilde{\phi}(\mu; 1)) + qU(\tilde{\phi}(\mu; 0))$$

and the value of telling the truth after that signal does not change. For the principal, however, the change comes after she receives a report recommending A . If she decides to monitor, her value function is defined as before so that

$$V_{A,S}(\mu) = q - c + \delta [(1 - \mu) \bar{F}(x(\mu))V(0) + (\mu + F(x(\mu))(1 - \mu))V(\mu)]$$

However, if she is not to monitor, she faces a lottery now that implies

$$\begin{aligned} V_{A,\emptyset}(\mu) &= \pi_\mu + \delta [(1 - \mu) \bar{F}(x(\mu))q + (\mu + F(x(\mu))(1 - \mu))(1 - q)] V(\tilde{\phi}(\phi_A(\mu); 0)) \\ &\quad + \delta [(1 - \mu) \bar{F}(x(\mu))(1 - q) + (\mu + F(x(\mu))(1 - \mu))q] V(\tilde{\phi}(\phi_A(\mu); 1)) \end{aligned}$$

Notice that the main argument behind Propositions 5 and 6 remains unchanged. Further it is easy to see that these value functions are themselves monotonic from standard arguments. For an example, start assuming that the principal's strategy is monotone, then

Proposition 15 *If $\sigma(\mu)$ is non-decreasing, U is non-decreasing. p_μ is non-increasing if V is increasing.*

Proof. The monotonicity of the value function is guaranteed by the fact that since the monitoring strategy is non-decreasing, a higher μ agent can secure himself a (weakly) higher payoff by following the equilibrium strategy of any μ' below μ . It suffices to show then that $x^*(\mu)$ is non-increasing. For a contradiction, assume that it is increasing in

some open interval (μ_0, μ_1) . First, assume that $\sigma > 0$, and notice then that

$$V_{A,\emptyset}(\mu) = V_{A,S}(\mu)$$

at every $\mu \in (\mu_0, \mu_1)$ but $\gamma = (1 - \mu) \bar{F}(x(\mu))$ is clearly decreasing in the range. Therefore π_μ is decreasing and the future value of not monitoring

$$\begin{aligned} \Delta = & [\gamma q + (1 - \gamma)(1 - q)] V(\tilde{\phi}(\phi_A(\mu); 0)) + [\gamma(1 - q) + (1 - \gamma)q] V(\tilde{\phi}(\phi_A(\mu); 1)) - \\ & - [\gamma V(0) + (1 - \gamma)V(\mu)] \end{aligned}$$

must be increasing. Rewriting

$$\begin{aligned} \Delta = & \gamma \left[qV(\tilde{\phi}(\phi_A(\mu); 0)) + (1 - q)V(\tilde{\phi}(\phi_A(\mu); 1)) - V(0) \right] + \\ & + (1 - \gamma) \left[(1 - q)V(\tilde{\phi}(\phi_A(\mu); 0)) + qV(\tilde{\phi}(\phi_A(\mu); 1)) - V(\mu) \right] \end{aligned}$$

or

$$\begin{aligned} \Delta = & \gamma \left[(2q - 1) \left[V(\tilde{\phi}(\phi_A(\mu); 1)) - V(\tilde{\phi}(\phi_A(\mu); 0)) \right] + V(\mu) - V(0) \right] \\ & + \left[(1 - q)V(\tilde{\phi}(\phi_A(\mu); 0)) + qV(\tilde{\phi}(\phi_A(\mu); 1)) - V(\mu) \right] \end{aligned}$$

but γ is decreasing and it is weighted by a positive term. It must that the effect through the change in the corresponding value after the outcomes occurs more than compensates. However, under the hypothesis, a ■

Finally, notice that since the information about output is not conclusive, i.e. the posteriors generated by such information are not degenerate, the value of this signal is positive for the principal and negative for the (strategic) agent.