Repeated Interaction and the Revelation of the Monitor's Type: A Principal-Monitor-Agent Problem.*

António M. Osório-Costa†

Universitat Rovira i Virgili and CREIP

(Dated: 12th December 2010)

Abstract

This paper studies a dynamic principal-monitor-agent relation where a strategic principal delegates the task of monitoring the effort of a strategic agent to a third party. The latter we call the monitor, whose type is initially unknown. Through repeated interaction the agent might learn his type. We show that this process damages the principal's payoffs. Compensation is assumed exogenous, limiting to a great extent the provision of incentives. We go around this difficulty by introducing costly replacement strategies, i.e. the principal replaces the monitor, thus disrupting the agent's learning. We found that even when replacement costs are null, if the revealed monitor is strictly preferred by both parties, there is a loss in efficiency due to the impossibility of benefitting from it. Nonetheless, these strategies can partially recover the principal's losses. Additionally, we establish upper and lower bounds on the payoffs that the principal and the agent can achieve. Finally we characterize the equilibrium strategies under public and private monitoring (with communication) for different cost and impatience levels.

JEL: C73, D82, D86, L23.

KEYWORDS: Relational Contracts, Replacement Strategies, Public and Private Monitoring, Bayesian Learning.

*Financial support from FCT grant SFRH/BD/17157/2004 is gratefully acknowledged. I would like to thank Venkataraman Bhaskar, Marco Celentani, Luis Corchón, Jacques Crémer, Martin Cripps, Mathias Dewatripont, Mikhail Drugov, Patrick Legros, Albert Ma, David Rahman, Juan Pablo Rincón-Zapatero, Tomas Sjöström and Galina Zudenkova, as well as the participants at XV SMYE, 2010 RES Conference and PhD Meeting, U. Carlos III, U. Rovira i Virgili, High School of Economics, U. Groningen, U. Bonn, 34th SEA Meeting, UECE Meeting, for helpful comments and discussions.

†Universitat Rovira i Virgili; Department of Economics; Av. de la Universitat, 1; 43204 Reus; Spain; Tel. +34 977 759 891; E-mail: superacosta@hotmail.com.

I. INTRODUCTION

Inside firms, delegation is a common and natural practice as the organizational structure grows large. Delegation is the assignment of authority from one organizational level to a lower one. In the manufacturing industry, it is common to employ a monitor to supervise the activity of a set of blue collar workers. Still, the high rank manager remains accountable for the outcome of the monitor's and the agent actions which affect the firm's performance.

A distinct structure of delegation occurs in the auditing industry and banking/finance supervision, typically an auditor/supervisor is sent to the client office to look at his accounts and other documents that might provide useful evidence about the client's behavior. The payoff of the auditing company or the supervising authority can be measured in monetary or in reputational terms. In either case, the auditor/supervisor actions affect the value of these variables.

A monitor is an individual with personal characteristics which interfere with his professional performance. Returning to the manufacturing industry example, when a worker effort is observed with noise, two different monitors may disagree on whether a given output realization is a signal of high or low effort.[1]

The table below shows the actual effort choices made by a worker and the possible interpretations that a particular monitor can make when these choices are not perfectly observable.

Monitor Perception

Worker Choice		High Effort	Low Effort
	High Effort	*Correct*	*Type I Error*
	Low Effort	*Type II Error*	*Correct*

For example, some monitors might be more permissive or tolerant with respect to their subordinates than others. Such is not necessarily a bad character trait. However, they are more likely to be the object of strategic behavior from the workers, because they tend to incur more often in type II errors. Strict or demanding individuals are also not necessarily better

[1] There are some connections to subjective performance evaluations. See for example Baker, Gibbons and Murphy (1994), Bull (1987), MacLeod (2003) or MacLeod and Malcomson (1989). The main issue here is how a low signal of the agent effort is interpreted in terms of actual effort choice.

or worse in professional terms. They tend to incur more often in wrong type I judgements, but at the same time, they bring more discipline to the relation.

Typically, after a sequence of low performance observations, the monitor might decide to punish the particular worker in question. These punishments can take several forms; one is the possibility of firing the individual. In spite of this, to satisfy their natural tendency for low effort, workers may explore the monitor's personal characteristics and limitations, patterns of behavior and working methods. This information might be of strategic relevance, allowing them to revise their initial prior beliefs and readjust their behavior accordingly. Repeated interaction facilitates this potential corrosive learning, with negative impact on the firm's performance.

In the auditing/supervision industry there are monitoring standards that have to be followed. However, different supervisors may differ with respect to specific aspects and working methods. Bernard Madoff, known to have run the largest Ponzi scheme in world history, describes in the following way his experience with two different supervisors from the SEC at different moments in time in the following way:[2]

Madoff stated that Mr. X was "doing things that make no sense to us at all."... Mr. X "talked tough, but didn't look at anything".

Madoff ... recalled Mr. Y was the supervisor.... Mr. Y, "knew what he was looking at and that was it."

In Madoff's words it is clear that Mr. Y's working methods were different from the ones employed by Mr. X. More relevant for us is that Madoff, through observation of the both supervisors' actions, was able to rank them in terms of the likelihood of uncovering his Ponzi scheme. In fact, Madoff recognized that if the 2006 exam conducted by Mr. X had been conducted by Mr. Y, the Ponzi scheme would have been found. Translated to the setting of the present paper, Madoff would have more incentives to continue with his Ponzi scheme if he learns that the supervisor is of Mr. X type.

[2] The words in italics were taken from the description of an interview conducted by Inspector General H. David Rotz and Deputy Inspector General Noelle Frangipane with Bernard Madoff on June 17, 2009 about interaction between Madoff, his company and the U.S. Securities and Exchange Commission (SEC).

In general terms a fund manager that is rewarded with a performance fee would have more incentives to build a risky portfolio if he knew in advance what risk measures a particular supervisor would pay more or less attention to. Similarly, an unscrupulous manager would feel more tempted to enter into illicit activities of a given kind if he knew beforehand that the auditor would pay little or not attention to issues of that nature.

Blue collar workers and most managers are typically paid according to a flat compensation. Bonus or performance stimulus are possible in the case of good performance, but typically they do not share the losses. This fact limits, to a great extent, the provision of incentives. Monitoring is then the mechanism that disciplines these agents.

This paper studies a dynamic *principal-monitor-agent* relation where a strategic principal delegates the task of monitoring the effort choices of an agent to a third party. The latter we call the *monitor*. His actions are fully characterized by his type. Exogenous circumstances require the delegation of the monitoring activities, without which trade would not be possible. The agent is strategic and has a natural tendency to supply low effort.[3]

Through repeated interaction and the observation of the monitor's actions with respect to the effort signal, the agent may learns his type, which in our setting represents the flexibility of the monitor towards the observations of the realized output. The agent strategically lowers his effort if he finds that the monitor is *tolerant*. We show that this revelation process damages the principal's payoff. When the principal strategic influence is restricted to deciding exclusively on whether or not to trade with the agent, we are able to characterize the worst case scenario.

In order to solve the principal's problem, we formalize the idea of replacement strategies, i.e. the principal replaces the monitor when she finds it convenient, paying a cost but disrupting any learning that the agent might have acquired.[4] When the replacement costs are null, she obtains the largest possible payoffs that can be achieved with replacement strategies. We are thus able to establish upper and lower bounds on the payoffs that both parties can

[3] The model history is particularly tailored to capture the firm versus blue collar workers type of relations. With appropriate changes in the text, the generalization for the auditing and financial supervision problems is immediate.

[4] Holmström (1982) and Cripps, Mailath and Samuelson (2004) in different contexts show the existence of a similar revelation effect. They also mention the possibility of permanent replacement as a potential mechanism to solve the problem.

achieve independently of the information structure, for different degrees of impatience.

In any realistic setting, replacement strategies cannot fully solve the principal's problem when contrasted with a reference measure. Surprisingly, this is true even if replacement costs are zero and high effort is always played in equilibrium. The reason is that, for the replacement mechanism to work, the parties cannot benefit from the potential revelation of a tolerant monitor - which in our setting is preferred in payoff terms when incentives are met, since he incurs less often in mistaken punishments.

Nonetheless, replacement strategies turn out to be useful to solve the principal's problem, reducing the losses associated with the agent learning and enlarging the spectrum of discounting, where equilibrium effort can be sustained. The success of these strategies depends crucially on the replacement costs and on the agent impatience.

We also characterize the sequential equilibrium under public and private monitoring, for varying replacement costs and impatience levels. When the noise signal of the agent effort is publicly observed the principal is able to make more precise replacement choices, because she knows exactly the moment in time at which the agent learns that the monitor type is tolerant. However, when the realized output is the agent's private monitoring, this piece of information is not available anymore; she has to infer the informational state of the relation. For that reason her replacement choices are always limited either by being premature, in the sense that the agent is still uninformed, incurring in an unnecessary cost, or by being late, in the sense that the monitor type is already revealed and the agent is providing low effort. Consequently and not surprisingly, replacement strategies are payoff inferior under private observation than under public monitoring.

This paper contributes to the theory of incentives by presenting replacement strategies as a mechanism to solve problems where an agent might acquire information with strategic value that penalizes the principal's payoffs. This paper provides recommendations on how a principal should optimally rotate a monitor in situations of this type. Such a solution is particularly relevant in situations where compensation is exogenously determined, limiting the provision of incentives to a great extent. In multiagent situations, the power of replacement strategies is amplified.

Discussion on the Main Assumptions - There is a set of persistent facts which this paper attempts to capture. They justify important specificities of the present paper and, in some sense, novel departures from the existing literature.

We introduce a distribution of monitor types, differing with respect to their flexibility towards the observations of the noisy effort measure. These individuals are not necessarily the ideal choice. The scarcity of "perfect profiles", the subjectivity and not necessarily well defined characteristics of an "ideal type" are not easy to identify, together with restrictions and biases in the recruiting process might lead to a selected candidate that is simply perceived as the best of a limited pool of screened individuals.[5]

Firms are aware of these limitations, but they also know the necessity of hiring these individuals for the regular functioning and expansion of their businesses.[6] These arguments rationalize the existence of the monitor in our model.

We restrict the role of compensation as an incentive mechanism. The assumption parallels the vision of Alchian and Demsetz (1972) of monitoring as a way of providing the agent with incentives.[7] The monitoring role is motivated by the independence between compensation and performance. Job insecurity is the mechanism that disciplines the relation.

This environment comes as natural in many economies and industrial sectors where base salaries are determined by social norms and political or legal aspects.[8] It also captures rigidities observed in the labor markets. When unsatisfied with an employee's performance, the employer is more prone to fire him than to decrease his compensation.

We assume that independently of the performance, once employed, the worker receives the promised end of period compensation. This is a "no-slavery" condition that the principal must respect ex-post. Punishments are executed after the compensation is paid.

The strategic aspect of firms' behavior tends to be mainly directed towards their costumers and competitors. In relation with its employees, firms usually pay the agreed compen-

---

[5] The classical "secretary problem" and its multiple extensions provides sufficient intuition about the recruiter screening difficulties. See Ferguson (1989) and Freeman (1983).

[6] Leibenstein (1987) discusses a great number of inefficiencies that exist inside organizations.

[7] See also Shapiro and Stiglitz (1984) where they link wages, unemployment, monitoring and efficiency. The worker incentives for high effort are also induced by the fear of being fired.

[8] Clearly, this is not an universal fact, it is easy to find exceptions, in particular in white-collar jobs where the employee has attributes/skills, recognized by both parties, that endow him with bargaining power or in tasks where performance and compensation are linked in a very sensitive way. Even in these cases, we should also expect a market reference measure to base the negotiations. See Kennan and Wilson (1989) for a survey on strategic bargaining models between unions and firms and for empirical studies.

sation and only fail to do so in special circumstances (e.g. financial difficulties, bankruptcy, etc.) that have little to do with incentives. A failure in commitment with an employee is not only seen by all the other employees but also spreads outside the firm's halls. Reputational considerations of this kind provide theoretical foundations for this assumption.

We abstain from discussing potential renegotiation from the initial agreement. This allows us to focus on the main issue of this paper without extra complexities.

Related Literature - The revelation of a player's type is not a new issue. However, it has never been applied to a principal-monitor-agent relation with the kind of structure as in this paper.

In a setting where incentives are driven by career concerns, Holmström (1982a) shows that an individual's ability is revealed over time through observation of his performance. On the limit, this result holds no matter how noisy the performance measure is and no matter how much the agent tries to bias the principal beliefs.[9]

Cripps, Mailath and Samuelson (2004) recently showed that if monitoring is imperfect, the type of player that is building a reputation is revealed in the long-run. Either the opponent players will become convinced that the reputation builder strategy will not change or they will come to learn his type. The former case cannot be an equilibrium; otherwise the reputation builder would take advantage of that fact to deviate from the potentially costly reputational action. In equilibrium, the revelation of the true reputation builder type occurs almost surely.[10]

One common characteristic that is reported in these papers is that the revelation of the player's type has negative payoff consequences for the revealed player. In order to deal with the problem, a solution that has been proposed by these authors is the permanent replacement of the player whose type is initially unknown, disrupting any potential for learning.[11] Such a costless solution does not fit and is hard to motivate in many economic

[9] Fudenberg and Tirole (1986) present an explicit theory of predation where an incumbent firm attempts to bias the learning process of an entrant firm about the market profitability. See also Mirman, Samuelson and Urbano (1993).

[10] See Cripps, Mailath and Samuelson (2007) where they establish analogous results for the case where the uninformed player is long-lived. See also Cripps, Ely, Mailath and Samuelson (2008). Revelation issues due to repeated interaction are also common in dynamic games with incomplete information; Renault (2009) provides a survey on the topic.

[11] Cole, Dow and English (1995) and Phelan (2006) model situations where the government type in power

problems. The present paper extends these ideas for different informational structures, in a simple setting where the replacement possibility is costly.

The literature in three party relations, principal-supervisor-agent, is in particular concerned with the resolution of the potential breakout of collusive arrangements between the agent and the supervisor. Tirole (1986) points out that problems related with the monitor's conflicting interests might arise. See also Laffont and Tirole (1991) and Kofman and Lawarree (1993) for further developments and extensions on this literature. In this paper, we are not concerned with delegation effects of this kind. Instead, we focus our attention on the negative effects associated with the revelation of strategic relevant information.

This paper is also related to the theory of self-enforced contracts where the provision of incentives is guaranteed by the sensitivity of the continuation value of the infinitely repeated game to changes in the player's actions. Future rewards and punishments provide the incentives for present behaviour.[12] Some important contributions to the large and growing literature in relational contracts are Klein and Leffler (1981), Bull (1987), and MacLeod and Malcomson (1989), all these papers assume perfect monitoring; while Shapiro and Stiglitz (1984), Baker, Gibbons and Murphy (2002), Levin (2003) and Fuchs (2007), assume imperfect monitoring. The present paper is in the spirit of the former set of contributions since we assume a different spectrum of monitoring imperfections.

Levin (2003) shows, that an optimal relational contract has an equivalent stationary representation that achieves the same payoffs. Because of the uncertainty about the moment in time when the monitor type is revealed (if it is revealed) and due to the replacement possibility, multiple alternations between informational sates are possible. For that reason, Levin's result does not generalize to our model; we have to solve the dynamic problem with dynamic constraints.

The paper is organized as follows. Section II presents the general model structure and

is not permanent. These situations are easy to motivate. See also Mailath and Samuelson (2001), where they study a problem where a competent firm might become inept, this mechanism keeps the competent firm with incentives and is an equilibrium.

[12] The present paper mixes concepts both from the theory of repeated games and the incentives theory. See Fudenberg and Tirole (1991) and Mailath and Samuelson (2006) for complete surveys of the former theory, and see for example Bolton and Dewatripont (2005) and Salanié (2005) for surveys of the latter theory.

discusses the learning process. Section III discusses the principal worst case scenario in detail and defines the reference measure. Section IV introduces the possibility of the monitor being replaced and considers the costless replacement solution. Section V characterizes the equilibrium replacement strategies under public monitoring. Similarly, Section VI characterizes the case where the agent privately observes the noisy effort signals. Section VII concludes. Proofs for the results in this paper can be found in Appendix A.

II. THE MODEL

We model a dynamic *principal-monitor-agent* relation that incorporates both elements of adverse selection and moral hazard. The former is motivated by the existence of a monitor, whose type is initially unknown to the agent and to the principal. We assume that the monitor has no incentives to choose an action that is against his type. Consequently, the monitor's actions are revealing. The problem can then be treated as one of dynamic moral hazard where nature first selects the monitor type.

The Principal and the Agent - Both parties are risk neutral. The timing of the relation is the following: at the beginning of each period $t = 0, 1, 2...$, the principal offers the fixed market compensation $w_t \in \mathbb{R}_+$ in exchange for the agent effort. The latter has the option to accept or refuse it. Once accepted, the agent privately chooses an effort level $e_t \in \left\{e^L, e^H\right\}$, where e^H denotes high effort and e^L denotes low effort, and $e^H > e^L \geq 0$. The effort level e^H is exogenous and it is the highest effort that the agent can physically supply per period. The low effort level e^L will be endogenously determined below.

At the end of each period, before players choose their actions, the realized output $y_t \in \mathbb{R}$ is observed by the monitor and the agent.[13] The principal might observe it or not, depending on the information structure considered. The realized output is a noisy measure of the agent's effort, i.e. $y_t = be_t + \varepsilon_t$ where ε_t is stochastic with $E(\varepsilon_t) = 0$ and $b \in \mathbb{R}_+$ is a productivity parameter. High (low) effort choices make higher (lower) outputs more likely.

The principal's ex-post payoff is the realized output subtracted from the wage paid to the agent. The principal's ex-ante expected payoff is then

$$\pi_{pt}^{j} = be_t - w_t, \tag{2.1}$$

[13] There is no loss in generality when allowing for potentially negative output values.

with $j = L$ when $e_t = e^L$ and $j = H$ when $e_t = e^H$.

Independently of the observed output, the principal has necessarily to pay the wage w_t corresponding to that period. We look at the principal as having one period commitment. She has, however, the option of not starting the relation with the agent or terminate it, if she finds convenient.

In exchange for the compensation paid at the end of the period, the agent suffers a disutility from effort equal to ce_t, where $c \in \mathbb{R}_+$ is the marginal cost of effort. The agent payoff is given by

$$\pi_{at}^j = w_t - ce_t, \tag{2.2}$$

with $j = L$ when $e_t = e^L$ and $j = H$ otherwise. Assume $b > c$, to assure an expected positive surplus.

Notice that there is no uncertainty about the agent's per period payoff. The principal holds all the idiosyncratic production risk.

If there is no trade both parties obtain their respective outside options denoted as $\underline{v}_p$ and $\underline{v}_a$. There is a wide range of interpretations that can be given to the values $\underline{v}_a$ and $\underline{v}_p$, we let these values to be exogenous.[14]

In the stage game, low effort is preferred by the agent but not by the principal. In order for the agent to always have an incentive to participate, we assume $\pi_a^L > \pi_a^H > \underline{v}_a$. The principal stage game payoff follows the relation $\pi_p^H > \underline{v}_p \geq \pi_p^L$. We assume that the last inequality binds, i.e. $\underline{v}_p = \pi_p^L$. This assumption simplifies the problem later. However, it has implications on the stage game.

	e^H	e^L	$\varnothing$
E	π_p^H, π_a^H	$\underline{v}_p, \pi_a^L$	$\underline{v}_p, \underline{v}_a$
NE	$\underline{v}_p, -ce^H$	$\underline{v}_p, -ce^L$	$\underline{v}_p, \underline{v}_a$

The principal is the row player, who has the option of employing E or not employing the agent NE. The agent is the column player, together with the effort choices he has the option of not trading with the principal, denoted as $\varnothing$, guaranteeing the outside option $\underline{v}_a$.

[14] These values can be easily endogenized at the cost of a more complex problem. For example; $\underline{v}_p$ could represent the value of a new relation with another agent, subtracted from the associated searching costs. A similar interpretation can be done for $\underline{v}_a$.

With the assumption $\underline{v}_p = \pi_p^L$, instead of a single pure strategies Nash equilibrium $(NE, \varnothing)$ we have two, i.e. $\left(E, e^L\right)$ and $(NE, \varnothing)$. None of these equilibria guarantee the principal more than the outside option. The equilibrium $(NE, \varnothing)$ is not interesting for either party but the equilibrium $\left(E, e^L\right)$ is preferred by the agent. The assumption is without loss of generality because, in the repeated relation, we want to sustain the repetition of the outcome associated with the equilibrium $\left(E, e^H\right)$.[15]

The Monitor - The monitor's task is to supervise the agent effort. He has autonomy to punish the agent when he observes a low output realization and subsequently reports this event to the principal.

The employed monitor can be of two types $\theta \in \left\{\theta^T, \theta^S\right\}$, with $\theta \in \mathbb{R}$ and $be_t^H > be_t^L \geq \theta^S > \theta^T$.[16] The type θ^S occurs with probability $\beta \in (0,1)$ and the type θ^T with the remaining probability. Here θ^S denotes a "strict" monitor, i.e. an individual that is more likely to perceive a low output realization as a signal of low effort. The strict type considers a low output realization every signal $y_t \leq \theta^S$, i.e. $Y^S \equiv \left\{y_t : y_t \leq \theta^S\right\}$, and a high output realization otherwise. A less strict monitor, call it "tolerant", is denoted by θ^T. This type attempts to capture a more flexible individual towards the output observations, so low output realizations are less likely to be interpreted as signalling low effort. For a tolerant monitor any output in the set $Y^T \equiv \left\{y_t : y_t \leq \theta^T\right\}$ is a signal of low effort.

Consequently, when the agent provides high effort, low output is observed with different probabilities depending on whether the monitor is "tolerant" or "strict",[17] i.e.

$$\Pr\left(y_t \in Y^T | e_t = e^H\right) \equiv p^T \textit{ and } \Pr\left(y_t \in Y^S | e_t = e^H\right) \equiv p^S,$$

respectively. High output is interpreted with the remaining probabilities.

Similarly, when the agent chooses low effort, the output is low with probabilities

$$\Pr\left(y_t \in Y^T | e_t = e^L\right) \equiv q^T \textit{ and } \Pr\left(y_t \in Y^S | e_t = e^L\right) \equiv q^S,$$

[15] Notice that the lower bound on effort supplied by the agent becomes endogenously determined, i.e. $e^L = \left(w + \underline{v}_p\right)/b$.

[16] The choice of two monitor types is made for simplicity. The model is robust to the introduction of more types and a continuum of efforts choices. The addition of more types simply increases the complexity of the problem.

[17] Without loss of generality we can consider another interpretation for θ, such as, for example, the state of nature. Clearly the story of the problem would have to be adjusted accordingly. Another interpretation is to look at θ as a monitoring technology, rather than an individual. The latter interpretation was suggested by Jacques Crémer. Nonetheless, we prefer to look at the monitor as a human.

depending on whether the monitor is "tolerant" or "strict", respectively.

To shorten notation, given the prior beliefs about the persistence of each type, we define the "expected" type $\theta^E \equiv \beta\theta^S + (1-\beta)\,\theta^T$.[18] In this case, the "expected" probability of observing low output when the agent is providing high and low effort are respectively

$$\beta p^S + (1-\beta)\,p^T \equiv p^E \text{ and } \beta q^S + (1-\beta)\,q^T \equiv q^E. \tag{2.3}$$

Since $\theta^S > \theta^T$ we have $Y^T \subset Y^S$, then for a same effort choice, a low output interpretation is more likely when the monitor type is "strict",

$$p^S > p^E > p^T > 0 \;\text{ and } q^S > q^E > q^T > 0. \tag{2.4}$$

Within the same type, high effort has associated a lower punishment probability

$$q^T \geq p^T, q^S \geq p^S \text{ and } q^E \geq p^E. \tag{2.5}$$

Putting together the inequalities (2.4) and (2.5) we obtain, without specifying an order, that q^T and p^S must lie in the interval $\left(p^T, q^S\right)$.

We ignore any payment made to the monitor by the principal and we do not specify the monitor's payoff functions. We assume that a "tolerant" monitor has no incentives to misbehave pretending to be of a "strict" type and vice versa. The value of θ not only denotes the type of monitor but also determines his behavior. The presence of the monitor is crucial, otherwise no trade would be possible and both parties would get their outside options. We see the monitoring task as more complex, specialized and with more responsibilities than simply observing output realizations. This assumption justifies the presence of the monitor even when the principal observes the realized output.

The Revelation Probabilities - The output observation carries a signal concerning the effort supplied by the agent. The monitor's reaction to the signal conveys information about his type to the agent. It is then natural to expect that a strategic agent would take advantage of this aggregated information. This is the intuition behind the revelation process.

Definition 1 *Conditional on no punishment, in the event $\left\{\theta^T < y_t \leq \theta^S\right\}$ we say that a revealing signal has occurred.*

[18] Outside the expected utility hypothesis, a more conservative approach would require the initial effort choice to be based on θ^S rather than on θ^E. There is no ambiguity about the likelihood of both types.

In the two types setting that we build in this paper, revealing information occurs only when the true type is θ^T. In this case the agent can update his believes and consequently revises his effort.

For simplicity, the event that the agent learns that the monitor is θ^S coincides with a punishment decision. Consequently, learning that the monitor is strict is irrelevant. Such is a consequence of punishment exclusively based on firing the agent.[19]

With two types, after observing a revealing signal and updating his beliefs there is nothing else for the agent to learn, so the type of monitor is fully revealed.

Remark. *If we added more types, conditional on no punishment, the agent would be learning more information about the monitor type with positive probability in every period. With punishment schemes based on the termination of the relation, the true value of θ could never be perfectly learned; the punishment disrupts this process.*[20] *However, the agent need not know exactly the type of monitor in order to discover profitable deviations.* ■

In every period of the relation with probability

$$\Pr\left(\theta = \theta^T\right) \Pr\left(y_t \in \underline{Y}^T \cap \underline{Y}^S | e_t = e^H\right) \equiv (1-\beta)\, r$$

the agent observes a revealing signal that excludes the type θ^S with probability one (in the spirit of Bayesian updating).[21] The larger the value of $(1-\beta)\, r$, the smaller the expected number of periods needed for a revealing signal to occur. In case of low effort e^L, the agent might observe a revealing signal with probability $(1-\beta)\, s$. We assume that $s \geq r$, this is the case for most distributions of interest. Consequently, when the agent provides low effort

[19] For such a case to be interesting, we would have to define an extra effort level, i.e. after the agent learns that the monitor is strict, he would adjust to a higher effort level. That would require the definition of more probabilities. Additionally we would have to consider other punishment schemes, e.g. review or forgiving strategies. See Footnote 20.

[20] If the punishment allows for forgiveness and there is unbound recall, the learning process will not be disrupted. It will continue after the punishment has been completed. In this case, θ might become common knowledge after a sufficiently large number of periods. In this case, the agent can also learn the value of θ from below, by updating the lower bound on the distribution of monitor types. (A possibility that we are not considering here) This would accelerate the convergence to common knowledge. See Cripps, Ely, Mailath and Samuelson (2008). Such a setting requires different strategic considerations.

[21] In general, this is different from saying that there is a signal clearly revealing the type of monitor. The former case does not necessarily lead to common knowledge, while the latter does. Here, with only two types, by exclusion both situations are equivalent. See the remark after Definition 1.

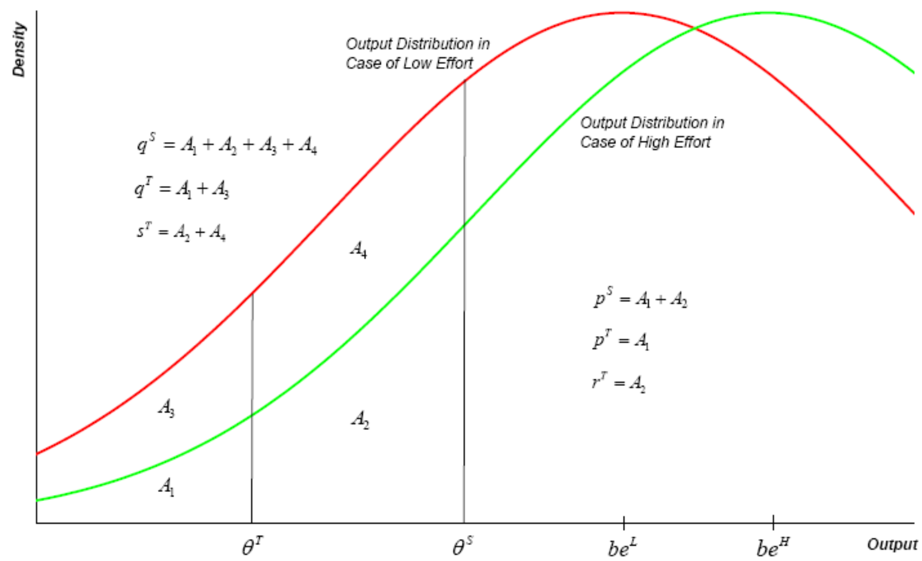

FIG. 1: The model - Relation between revelation and punishment probabilities.

he accelerates the potential revelation process, but at the same time he is also more likely to be punished.

Since $be^L \geq \theta^S$, we can relate the punishment and the unconditional learning probabilities, i.e. $p^S = p^T + r$ and $q^S = q^T + s$. We can also write $p^E = p^T + \beta r$ and $q^E = q^T + \beta s$. Without knowing the relation between $q^S - q^T$ and r, we cannot rank q^T and p^S. Figure 1 provides an illustration.

Actions, Histories and Strategies - The Nash equilibria of the stage game gives low payoffs for the principal. Through the provision of intertemporal incentives, we can achieve payoffs above the set of stage game Nash equilibria.

For convenience, we consider strategies where a low output observation triggers the immediate termination of the relation. Both parts are then left with their outside options.[22] The punishment scheme interferes with the learning process, see Footnote 20. The important aspect is that these punishments are mutual and occur with positive probability along the equilibrium path, causing destruction of value, and consequently bounding the equilibrium payoffs away from full efficiency. Even the reference value is not fully efficient. Instead of

[22] In the real world principal-agent relations punishments are usually not so severe. Typically, a firing decision is only made after a sequence or a certain number of low output interpretations. Radner (1985), Rubinstein (1979) and Rubinstein and Yaari (1983) study strategies with a similar structure.

searching for a mechanism that implements an optimal effort level without surplus losses, we acknowledge the existence of inefficiencies. This assumption is without loss in generality and is motivated by tractability issues. Moreover, we assume that punishments based on a "money-back guarantee" cannot be contracted.[23]

The potential revelation of the monitor type creates an extra layer of inefficiency and damages the principal's interests. Our goal is to search for solutions that attempt to eliminate or minimize this latter effect on the principal's payoffs.[24]

Both the principal and the agent discount the future according to some $\delta \in (0,1)$.

Once the principal and the agent decide to participate, the monitor is hired and the parties take their actions. At any given moment in time t; an agent's action is an effort choice $e_t \in \left\{e^L, e^H\right\}$. A principal's action is a replacement choice $r_t \in \{0,1\}$, where $r_t = 0$ means not replacing the monitor and $r_t = 1$ otherwise. A monitor's action $m_t \in \{0,1\}$ is a choice between punishing or not, respectively $m_t = 1$ and $m_t = 0$.

Given the actions and the observed output up to time t, a history of play is built. Depending on the information structure considered, different private histories players will accumulate. The history of output realization at a given time t is $h_y^t \equiv \{y_0, y_1, ..., y_{t-1}\}$. The monitor history of actions is $h_m^t \equiv \{m_0, m_1, ..., m_{t-1}\}$ and is public observed in any setting, for that reason we are able to keep the recursive structure. The principal and the agent condition their actions on this public observed history, even holding private and different pieces of information.[25] This way we can apply the dynamic programming methods developed in

[23] We assume that the only way to provide incentives is through mutual punishment. Alternatively, given the one-sided moral hazard structure of our problem, in theoretical terms we could consider other ways of providing incentives that are less costly for the parts, i.e. by transfers of value between the parties involved in the relation. Even with noisy signals about players' actions Fudenberg, Levine and Maskin (1994) (see also Fudenberg and Levine (1994) and Sannikov (2007)) have shown that, in problems of this type and with arbitrarily patient players, we can obtain any feasible and individual rational payoffs that are fully efficient. However, for our particular problem, transfers of value from employees to employers as a punishment scheme are not usually observed in reality.

[24] In a bilateral risk neutral setting, when the agent owns or receives all the surplus from his work, it is possible to obtain full efficient payoffs. The agent has no incentive to deviate from high effort. See for example Stiglitz (1974).

[25] Kandori (2002) points out the difficulties that arise when dealing with private monitoring when a recursive structure is absent. Compte (1998) and Kandori and Matsushima (1998) present the first folk theorem for private monitoring with communication. See also Gossner and Tomala (2009) and Mailath and Samuelson (2006) for surveys on the subject.

Abreu, Pearce and Stacchetti (1986, 1990). The principal and the agent private histories are $h_p^t \equiv \{r_0, r_1, ..., r_{t-1}\}$ and $h_a^t \equiv \{e_0, e_1, ..., e_{t-1}\}$ respectively.

Throughout the paper we focus on two information structures. In either case both the agent and the monitor observe the realized output. When the principal also observes the realized output, we say that *monitoring is public*; then the public history is $h_y^t \cup h_m^t$, the principal's private history is $h_y^t \cup h_m^t \cup h_p^t$ and the agent's private history is $h_y^t \cup h_m^t \cup h_a^t$. When the principal does not observe the realized output we say that the signals are the *agent's private monitoring*: then the public history is h_m^t, the principal's and the agent's private histories are respectively $h_m^t \cup h_p^t$ and $h_y^t \cup h_m^t \cup h_a^t$.

A pure strategy for player $i \in \{a, p\}$ is a mapping from the set of $i's$ private histories into the set of $i's$ pure actions. When the distribution of signals generated by the effort choices has full support, a *perfect public equilibrium* is a *sequential equilibrium.* In order to be general enough to deal with private monitoring structures, we work with the latter concept.

Notation - Players select the same action for every period of the repeated game until the monitor type is revealed or the monitor is replaced, in which case a new action is selected; for that reason we remove the time index t. Instead, we distinguish between the uninformed state, denoted with the superscript 0, and the informed state, denoted with the superscript 1. The punishment event is an absorbing state.

III. THE REFERENCE VALUE AND THE REVELATION OF THE MONITOR TYPE

In this Section we restrict the principal's strategic role other than deciding on whether or not to hire the agent and stoping the relation if she finds it convenient. This strategic structure is particularly interesting since it highlights the destructive effect that the revelation of the monitor type has on the principal's payoff. This is the principal's worst case scenario.

The agent incentives are provided by the fixed market compensation and the uncertainty about the monitor type. The latter incentives might disappear as the game unfolds. With positive probability the agent might learn that the monitor type is tolerant and adjust his effort accordingly.

Given the strategic restrictions on the principal's behavior, whether she observes the realized output or not becomes irrelevant.

The Repeated Relation Payoffs - For $i \in \{a, p\}$, denote π_i^0 as the stage game payoff in the uninformed state 0, i.e. before the monitor type has been revealed. This value depends on the effort choices made by the agent, i.e. $\pi_i^0 \in \left\{\pi_i^L, \pi_i^H\right\}$. Similarly, in the informed state 1, i.e. when the monitor type has been found to be tolerant, the stage game payoffs are $\pi_i^1 \in \left\{\pi_i^L, \pi_i^H\right\}$. Let $v_i^{0,1}$ denote the normalized expected value, for $i \in \{a, p\}$, where the first superscript refers to the effort choice made in state 0 and the second superscript refers to the effort choice made in state 1.

Lemma 2 *Suppose the agent and the monitor observe the realized output. The infinitely repeated normalized expected payoff when the agent chooses $\{e^0, e^1\} = \left\{e^H, e^L\right\}$ is*

$$v_i^{H,L} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i}{D} + \delta\,(1-\beta)\,\frac{r}{D}\,\frac{(1-\delta)\,\pi_i^L + \delta q^T \underline{v}_i}{1-\delta\,(1-q^T)}. \tag{3.1}$$

for $i \in \{a, p\}$, where $D \equiv 1 - \delta\left(1 - p^E - (1-\beta)\,r\right)$.

Expression (3.1) has two components, the first ratio on the right-hand side is $i's$ ex-ante expected value of the uninformed state, when the agent provides high effort. The second part on the right-hand side is $i's$ ex-ante unconditional expected value of the informed state, when the agent provides low effort.

From expression (3.1) we can obtain the payoff associated with low effort and high effort in any informational state, $v_i^{L,L}$ and $v_i^{H,H}$ respectively. In the former case, replace respectively p^E, r and π_i^H for q^E, s and π_i^L. (Such an expression, for the case where $i = a$, can be found in the Proof of Proposition 4, expression (A3).) The latter case is obtained by replacing respectively q^T and π_i^L for p^T and π_i^H in (3.1). (This is expression (3.7) of Definition 5, below.)

The effort choices $\{e^0, e^1\} = \left\{e^L, e^H\right\}$ are never optimal. In other words, low effort and an associated high probability of punishment while uninformed, is not compatible with a later high effort when the agent learns that the monitor is tolerant. In our setting, it can be shown that $v_a^{L,H}$ is always dominated by the payoffs of some other strategy for all $\delta \in (0, 1)$.

The Agent Effort Incentives - As discussed in Section II, the principal pays the promised end of period compensation independently of the observed performance. The principal intertemporal incentives are then satisfied by assumption. The same does not happen with the agent who has a natural tendency to supply low effort. When the agent is sufficiently impatient the market compensation may not be enough to sustain high effort in both informational states. He provides high effort while uninformed and low effort once informed. It might even be the case, for larger impatience levels, that the agent finds it is in his best interest to provide low effort in any state.

We say that the agent effort choice $\{e^0, e^1\}$ is self-enforceable[26] when

$$\{e^0, e^1\} = \underset{e^0 \in \{e^L, e^H\},\ e^1 \in \{e^L, e^H\}}{\arg\sup} v_a^{0,1}. \tag{3.2}$$

The effort choices $\{e^0, e^1\}$ can be part of a non trivial equilibrium if, in addition, they guarantee that the principal has incentives to hire the agent. Before that, the following remark is in order, respecting to the methods employed to solve the dynamic problem that we are modelling.

Remark 3 *With two potential informational states, we can decompose the repeated relation into two relevant subgames. The whole game that starts in state* 0 *and extends for an unknown number of periods until the monitor interprets a low output realization, suggesting low effort. In this case the relation terminates. The state* 1 *subgame is initiated if the monitor type is revealed to be* θ^T*.*

When the informational state moves to state 1*, the agent might change his behavior. To solve the game, we first find the best strategy for the agent in any of the potentially infinitely repeated state* 1 *subgames and then search for the best strategy for the game that starts in the uninformed state* 0*. The approach is similar to backward induction; the difference is the timing uncertainty associated with the beginning of the informed state subgame.*

The Principal Participation - We require the ex-ante condition $v_p^{0,1} > \underline{v}_p$ as necessary for the principal to have interest in trading with the agent, i.e. the principal will only

[26] Throughout the paper we say that the agent has incentives to choose $\{e^0, e^1\}$ (or the agent incentives are satisfied). In the theory of incentives, expression 3.2 is an incentive compatible constraint.

participate if she can do strictly better than in the stage game. However, this does not guarantee the principal's participation in the informed state 1 subgame. We denote the value of this subgame as $v_p^{\cdot,1}$. In the worst case scenario, i.e. the agent provides low effort in the informed state, we want

$$v_p^{\cdot,L} = \frac{(1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p}{1-\delta\,(1-q^T)} \geq \underline{v}_p. \tag{3.3}$$

Otherwise, the principal would prefer to terminate the relation once the agent is informed.[27] For the principal to have incentives to stay active, the agent cannot provide effort below a certain value. For that reason, in Section II, we assumed that $\underline{v}_p = \pi_p^L$. Intuitively once informed, if the agent decides to supply low effort, there is a minimum effort level that keeps the principal interested in participating.

With this restriction on the agent's effort, the principal's participation condition $v_p^{0,1} > \underline{v}_p$ is sufficient. When $v_p^{0,1} = \underline{v}_p$, the principal will not trade with the agent. This is the case when the agent's dominant strategy is to provide low effort in both informational states, i.e. $v_p^{L,L} = \underline{v}_p$. The principal does not contract with an agent who has no incentives to provide high effort at least during the uninformed state.[28]

The Agent Participation is always guaranteed by the assumption $\pi_a^L > \pi_a^H > \underline{v}_a$. In any scenario the agent guarantees at least $\underline{v}_a$.

Putting together the principal's and the agent's participation conditions, we obtain the constraints on the agent binary effort set

$$\frac{w+\underline{v}_p}{b} = e^L < e^H < \frac{w-\underline{v}_a}{c}, \tag{3.4}$$

with e^L endogenously determined and e^H exogenous but constrained. The former value is higher when the principal has a high outside option, the market compensation is high or the productivity of the labor is low. Also intuitive is the upper bound on high effort, which depends positively on the market wage but negatively on the agent's outside option and the cost of effort.

[27] Termination of the relation after an informative signal can be part of an equilibrium. Here, we want to rule out such a possibility.

[28] We thus disallow strategic behavior from the low effort agents, so as to provide high effort in just one period of the relation, which would push the principal's payoff above $\underline{v}_p$ in expected terms. The principal participation decision is based on the agent's incentives, which the principal knows because δ and other relevant parameters are common knowledge. We want to improve over the stage game payoff $\underline{v}_p$.

Equilibrium Behavior - Given that high effort is supplied in the uninformed state, let δ^+ denote the discount threshold above which the agent prefers to supply high effort once informed that the monitor type is θ^T. Given the principal's strategic limitation, δ^- denotes the discount threshold below which the agent prefers to provide low effort in any informational state.

Proposition 4 *When $\delta \in [\delta^-, \delta^+)$, the agent chooses $\{e^0, e^1\} = \left\{e^H, e^L\right\}$. The interval $[\delta^-, \delta^+) \subseteq [0, 1)$ and is nonempty if*

$$q^T \left(\pi_a^H - \underline{v}_a\right) > p^T \left(\pi_a^L - \underline{v}_a\right) \tag{3.5}$$

and

$$s \left(\pi_a^H - \underline{v}_a\right) > r \left(\pi_a^L - \underline{v}_a\right). \tag{3.6}$$

Independently of the informational state, when $\delta \in [0, \delta^-)$ the agent provides low effort, while if $\delta \in [\delta^+, 1)$ the agent provides high effort.

The agent provides low effort in the informed state if $\delta < \delta^+$. Given this behavior, it is a dominant strategy to supply high effort in the uninformed state if $\delta \geq \delta^-$. Condition (3.5) and (3.6) establish that the intersection of these intervals is nonempty, while condition (3.5) alone guarantees that $\delta^+ < 1$.

Since $\pi_a^L > \pi_a^H > \underline{v}_a$, condition (3.5) states that it is not enough for q^T to be larger than p^T, it has to be sufficiently large. If that is the case, for a sufficiently patient agent, the market compensation is sufficient to sustain high effort.

Similarly, inequality (3.6) requires that the likelihood of an output observation inside the informative region to be much greater in the case of low effort. This condition is easier to interpret when added to condition (3.5). In this case we have $q^S \left(\pi_a^H - \underline{v}_a\right) > p^S \left(\pi_a^L - \underline{v}_a\right)$. Low effort accelerates the potential revelation of the monitor type, but at the cost of a higher likelihood of punishment. In other words, since a tolerant type occurs only with probability $(1 - \beta) \in (0, 1)$, the acceleration of the revelation process turns into an acceleration of the punishment event, independently of the true monitor type. This explains why the agent chooses $\{e^0, e^1\} = \left\{e^H, e^L\right\}$ for $\delta \in [\delta^-, \delta^+)$, but also why $\{e^0, e^1\} = \left\{e^L, e^H\right\}$ is always a dominated strategy.

Proposition 4 identifies three distinct potential behaviors depending on how the agent discounts the future. When $\delta \in [\delta^+, 1)$, the agent provides high effort, ignoring any signal regarding the monitor type. The agent prefers to suffer the extra disutility $c\left(e^{H}-e^{L}\right)$ imposed by high effort in every future period of the repeated game rather than to provide low effort and increasing the probability of punishment from p^{T} to q^{T}. The reference payoff $v_{a}^{H,H}$ is dominant and the principal obtains the payoff $v_{p}^{H,H}>\underline{v}_{p}$. In the opposite direction, if $\delta\in[0,\delta^{-})$, the agent incentives are not satisfied and the principal prefers not to participate since $v_{p}^{L,L}=\underline{v}_{p}$.

The interesting possibility occurs for "intermediate" impatience levels, i.e. $\delta\in[\delta^{-},\delta^{+})$. The agent prefers to provide high effort while uninformed, reducing the punishment likelihood, waiting to find if the monitor type is tolerant, in which case he deviates to low effort and obtains larger expected gains. There is a prize for the agent if the informed state occurs. In this case the principal obtains $v_{p}^{H,L}>\underline{v}_{p}$.

The Reference Payoffs - Since there is the potential of mistaken punishments on the equilibrium path, there is no way to achieve full efficient payoffs. To perform payoff comparisons we define the reference payoffs, i.e. the payoffs that attain the highest possible aggregate surplus for a given δ.

Definition 5 *The reference payoff is the value that the principal and the agent would obtain if the agent provides high effort independently of the informational state, i.e.*

$$v_{i}^{H,H}=\frac{(1-\delta)\,\pi_{i}^{H}+\delta p^{E}\underline{v}_{i}}{D}+\delta\,(1-\beta)\,\frac{r}{D}\frac{(1-\delta)\,\pi_{i}^{H}+\delta p^{T}\underline{v}_{i}}{1-\delta\,(1-p^{T})}. \tag{3.7}$$

for $i\in\{a,p\}$.

The reference payoffs require high effort always to be chosen by the agent, Which is why the outcome is less inefficient. Instead of attempting to eliminate this inefficiency we acknowledge it,[29] our goal is to minimize the principal's payoff losses due to the revelation of the monitor type.

[29] A full efficient solution would require $v_{a}^{H,H}=\pi_{a}^{H}$ and $v_{p}^{H,H}=\pi_{p}^{H}$. For that to be possible, the signals had to be perfectly informative in case of high effort, i.e. $p^{T}=p^{S}=0$. Another way to obtain such a result is to employ punishments based on transfers of value between the principal and the agent, see Footnote 23. We choose $p^{T}>0$ and $p^{S}>0$, to deal with the possibility that nature and/or noisy information may disturb the decision process.

It is important to stress, that $v_i^{H,H}$ is an equilibrium payoff only when $\delta \in [\delta^+, 1)$. For that reason, for any $\delta < \delta^+$, the value $v_i^{H,H}$ only plays the role of a reference measure. Nonetheless, the sum $v_a^{H,H} + v_p^{H,H} \equiv W$ achieves the largest total surplus.[30]

In the present paper we focus on replacement strategies. However, compensation incentives can sustain the surplus W for impatience levels below δ^+. This is why compensation incentives are equivalent to transfers of value from the principal to the agent, leaving the total surplus unchanged. Later, we will see that there are perverse effects associated with replacement strategies other than the replacement costs.

Payoff Comparisons - The agent observation of the realized output cannot harm him in payoff terms. Moreover, independently of the information that the principal might hold and given the strategic limitations that we imposed on her behavior, we should expect her payoff to be penalized by the agent's strategic behavior. The following result establishes the relation between the reference payoffs of Definition 5 and the payoffs $v_i^{H,L}$ of Lemma 2 for the interesting case $\delta \in [\delta^-, \delta^+)$.

Corollary 6 *For $\delta \in [\delta^-, \delta^+) \subseteq [0, 1)$ we have;*
(i) $v_a^{H,L} > v_a^{H,H}$,
(ii) $v_p^{H,L} < v_p^{H,H}$.
(iii) Under condition (3.5), $W > v_a^{H,L} + v_p^{H,L}$ for all $\delta \in [0, 1)$.

When the agent learns that the monitor type is tolerant and $\delta \in [\delta^-, \delta^+)$, the principal suffers payoff losses. Since the parties in the relation have opposing interests, the result obtained in part (i) justifies the result of part (ii). An extra gain for the agent implies a loss for the principal (not necessarily equivalent).

Part (iii) establishes that high effort in both informational states is surplus superior for any discount level. The extra gains obtained by the agent by deviating to low effort in the informed state, part (i), does not compensate for the payoff losses incurred by the principal, part (ii). This is the case because the gains that the agent obtain are due to a reduction in effort, affecting the generated surplus in an adverse way.

[30] Expression (3.7) can be obtained along the same lines as in the Proof of Lemma 2. Alternatively, in expression (3.1), simply replace respectively q^T and π_i^L for p^T and π_i^H.

IV. THE PRINCIPAL RESPONSE - SEARCHING FOR SOLUTIONS

In the previous section we have intentionally limited the principal's strategic role to isolate the monitor type revelation effect. We now take a step further towards more realistic scenarios. In addition to the monitoring activities, in an attempt to correct the losses on the principal's payoff, we consider that the principal can strengthen the agent's incentives by costly replace the monitor. Every time the monitor is replaced any learning that the agent has acquired is lost, restoring the uncertainty about the monitor type. The agent is then forced to revise his effort choice.

In the previous section , the information structure was unimportant for the principal, with costly replacement depending on whether she observes the realized output or not leads, to different problems. We consider two potential situations:

(i) Public information with costly replacement, Section V.

(iii) Agent's private information with costly replacement, Section VI.

Before exploring these information structures, we consider a solution that has been suggested in Holmström (1982) and Cripps, Mailath and Samuelson (2004), which calls for replacing the monitor in every period. This goal is to establish an upper bound on the principal's payoffs with replacement strategies in the discounting region $[\delta^*, \delta^+)$.

A. The Trivial Solution - Costless Replacement

We now discuss the case where there are no replacement costs and the principal is free to substitute the monitors with any desired frequency. The costless replacement case is just a particular case of more general costly structures. Nonetheless, since it corresponds to the opposing limit situation discussed in the previous Section, it deserves exclusive treatment in the sense that it formalizes the principal's best case scenario outside the discount region $[\delta^+, 1)$. Because of the replacement flexibility, we call the costless scenario, the trivial solution.

Proposition 7 *Independently of the monitoring informational structure and supposing that there are no replacement costs, when $\delta \in [\delta^*, \delta^+)$, the best strategy for the principal is to replace the monitor in every period and for the agent to supply high effort.*

Conditions (3.5) and (3.6) guarantee that $[\delta^*, \delta^+)$ is nonempty and $[\delta^-, \delta^+) \subset [\delta^*, \delta^+)$.

The successive replacement of the monitor disrupts the learning process. The agent cannot profit from the information acquired in one period because the monitor constantly changes. He is never able to update his prior beliefs.

When $\delta \geq \delta^*$, the agent supplies high effort in all periods of the infinitely repeated game. Below δ^*, since the agent incentives are not satisfied, the principal's optimal strategy is to not participate. This scenario is similar to one where the monitor type is known to be θ^E.

It is worth noticing that the repeated replacement of the monitor has the positive effect of sustaining high effort in the discount region $[\delta^*, \delta^-)$. Until now this was not possible (see Proposition 4).

Denote v_i^{H*} as the payoff that the principal and the agent obtain under permanent costless replacement. Such a payoff can be sustained in the discount region $[\delta^*, 1)$. However, if $\delta \in [\delta^+, 1)$, the market compensation is enough to sustain incentives for high effort in any informational state. In this case the principal should never replace the monitor.

Even though the permanent replacement solution is able to discipline the agent, it surprisingly achieves payoffs below the reference measure.

Proposition 8 *Independently of the monitoring informational structure and supposing that there are no replacement costs, we have $v_i^{H,H} > v_i^{H*}$ for all $\delta \in (0, 1)$ and $i \in \{a, p\}$.*

This result highlights the negative side of replacement strategies. To understand this result, notice that when incentives for high effort are met, the tolerant monitor is preferred by both the agent and the principal. The reason is that this type incurs less often in mistaken punishments. The permanent replacement of the monitor provides the agent with incentives for high effort for all $\delta \in [\delta^*, 1)$. However, the parties cannot benefit from the extra gains associated with the potential revelation of a tolerant monitor. For that reason, the payoffs under permanent replacement are bounded away from the reference payoffs.

In other words, when $\delta \in [\delta^+, 1)$, the market compensation is sufficient to provide incentives. The principal should not replace the monitor, she can thus benefit from the potential revelation of a tolerant monitor and obtain $v_p^{H,H}$.[31]

[31] Care should be taken when comparing v_i^{H*} with $v_i^{H,H}$. The latter is a reference value and can only be sustained in the discount region $[\delta^+, 1)$, otherwise incentives collapse. The former can be sustained in

However, below δ^+ the agent incentives for high effort in the informed state collapse. The potential benefit from the revelation of a tolerant monitor disappear. The principal is better off replacing the monitor in every period, keeping the uncertainty about the monitor type always alive.

The results obtained until now lead us to the following conclusion about the principal's chances of recovering the losses incurred due to the revelation of a tolerant monitor when employing replacement strategies.

Corollary 9 *Independently of the monitoring structure, with replacement strategies we have the following bounds on players' payoffs:*

(i) $v_i = v_i^{H,H}$ *for* $i \in \{a, p\}$ *, when* $\delta \in [\delta^+, 1)$ *.*

(ii) $v_p \in \left[v_p^{H,L}, v_p^{H*}\right]$ *and* $v_a \in \left[v_a^{H*}, v_a^{H,L}\right]$ *, when* $\delta \in [\delta^-, \delta^+)$ *.*

(iii) $v_p \in \left[\underline{v}_p, v_p^{H*}\right]$ *and* $v_a \in \underline{v}_a \cup \left[v_a^{H*}, v_a^{H,L}\right]$ *, when* $\delta \in [\delta^*, \delta^-)$ *.*

In general terms in any information structure, the principal and the agent cannot obtain a payoff above v_p^{H*} and $v_a^{H,L}$ respectively, the exception is the case $\delta \in [\delta^+, 1)$. The agent cannot get less than v_a^{H*} unless the costs required to keep the agent with incentives are so high that the principal prefers not to trade. This situation only occurs in the region $[\delta^*, \delta^-)$ because the principal's choices are constrained by the need to provide the agent with incentives at least in the uninformed state. Otherwise, for $\delta \in [\delta^-, \delta^+)$, the principal can always guarantee at least $v_p^{H,L}$ by never replacing the monitor.

The result reinforces the message of Proposition 8; we cannot rely on replacement strategies to reach the most efficient outcome. However, these strategies can provide a partial solution to the monitor revelation problem in the discounting region $[\delta^*, \delta^+)$. For that reason they are worth studying.

Finally, we acknowledge the difficulty of motivating situations with permanent replacement. Nonetheless, the trivial solution seems to fit with the behavior of some managers in specific situations. To be more concrete, consider a monitor that in some days presents a good mood , similar to a type θ^T behavior, while on others days he presents a bad mood,

---

the interval $[\delta^*, 1)$. Strictly speaking, these payoffs can only be compared for $\delta \in [\delta^+, 1)$. In the region $[\delta^*, \delta^+)$ the value v_i^{H*} can be sustained, while $v_i^{H,H}$ is just a reference, see Propositions 4 and 7.

similar to a type θ^S behavior. The subordinates are then unable to identify his state on a particular day. In this case, the manager is using a behavioral strategy, randomizing over the mood θ^S with probability β and the mood θ^T with probability $1-\beta$ in each period. While intuition may support the existence of strategic behavior of this kind in principal-agent relations, further research should verify the validity of such an assertion.

V. PUBLIC MONITORING WITH COSTLY REPLACEMENT

In most economic problems, the assumption of free replacement is hard to sustain. We now consider a more interesting scenario where the principal pays a fixed cost $k \geq 0$ every time she decides to replace the existing monitor. These are organizational costs due to adaptation, learning, and/or mandatory firing costs.

The realized output is publicly observed by both the principal and the agent. This information structure captures situations where the principal delegates the monitoring task but, at the same time, keeps track of the realized output. We look at the monitor task as being more complex than simply observing output realizations, he also provides support and assists the agent. Without the monitor no trade would be possible.

Alternatively, we can think that the monitor sends the principal a report at the end of each period with the realized output and the action taken , i.e. the monitor's personal interpretation of the observed signal.

If the true monitor type is tolerant, the principal learns it at the same time as the agent. For that reason it is not rational for the principal to replace the monitor before the occurrence of a revealing signal.

A replacement strategy is then a decision to substitute the existing monitor, $n = 0, 1, 2, ...$ periods after a revealing signal has been observed. Intuitively, when $n = 0$ no learning is possible, the monitor is replaced as soon as an informative signal is observed. This is the highest (rational) replacement frequency and consequently the one with highest total cost. As these costs increase, it might be better for the principal to choose $n = 1$, i.e. to substitute the monitor one period after a revealing signal is observed. The total costs decrease due to a decrease in the replacement frequency, but if $\delta \in [\delta^*, \delta^+)$, the agent can benefit during one period from learning that the monitor is tolerant. The extreme case is when the monitor is never substituted, i.e. $n \to \infty$. In this case, the principal pays no replacement costs but the agent benefits from the potential revelation of the monitor for the rest of the relation.

The principal faces a trade off between frequent replacement, i.e. right after the monitor revelation, with larger costs but no learning, or less intensive replacement with lower costs but allowing potential deviations from high effort.

Denote $v_{i,k,n}^{0,1}$ as the expected normalized value of the relation for $i \in \{a, p\}$ when the principal replaces the monitor $n \in \mathbb{N}_0 \equiv \{0, 1, 2, ...\}$ periods after a revealing signal, paying the cost $k \geq 0$ per replacement, and the agent effort choice in each informational state is $\{e^0, e^1\}$.

Lemma 10 *Suppose that the monitoring is public, $k \geq 0$, the principal replaces the monitor $n \in \mathbb{N}_0$ periods after a revealing signal, and the agent chooses $\{e^0, e^1\} = \{e^H, e^L\}$. The infinitely repeated normalized expected payoff for player $i \in \{a, p\}$ is*

$$v_{i,k,n}^{H,L} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i + \delta(1-\beta)\,r\left((1-\delta)\,\pi_i^L + \delta q^T \underline{v}_i\right)\frac{1-\delta^n\left(1-q^T\right)^n}{1-\delta(1-q^T)} - k(1-\delta)\,D}{1-\delta\left(1-p^E-(1-\beta)\,r\left(1-\delta^n\left(1-q^T\right)^n\right)\right)} + k(1-\delta)\,, \tag{5.1}$$

where $D \equiv 1-\delta\left(1-p^E-(1-\beta)\,r\right)$, and with $k = 0$ when $i = a$.

The replacement cost k is exclusively incurred by the principal, for that reason when $i = a$ we must set $k = 0$.

The expression (5.1) incorporates the agent's optimal strategic behavior (best response) for a given principal replacement choice n. Such a construction reduces the computation of the sequential equilibrium of the infinitely repeated game to an optimization problem from the principal's point of view.

The following properties of expression (5.1) are worth to notice. When $n \to \infty$ we obtain the expression $v_i^{H,L}$ of Lemma 2, i.e.

$$v_{i,k,n}^{H,L} \to v_i^{H,L} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i}{D} + \delta(1-\beta)\,\frac{r}{D}\,\frac{(1-\delta)\,\pi_i^L + \delta q^T \underline{v}_i}{1-\delta\left(1-q^T\right)}. \tag{5.2}$$

This is the no replacement case. Similarly when $n = 0$, we obtain

$$v_{i,k,0}^{H,L} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i - k\delta(1-\delta)(1-\beta)\,r}{1-\delta\left(1-p^E\right)}, \tag{5.3}$$

which for $k = 0$ equals expression v_i^{H*}. Notice that with public monitoring and costless replacement, the case where the monitor is replaced after a revealing signal and the case where the monitor is replaced in every period are equivalent. In either situation the agent cannot benefit from learning.

Denote $v_{i,k,n}^{H,H}$ as the value of the relation when, in expression (5.1), we replace π_i^L and q^T by π_i^H and p^T respectively, and denote $v_{i,k,n}^{L,L}$ as the case where π_i^H, r and p^E are replaced by π_i^L, s and q^E respectively. Similar to Section III, since replacement costs do not enter into the agent's payoff function, the agent's behavior is determined by the discount thresholds that solve $v_{a,0,n}^{L,L} = v_{a,0,n}^{H,L}$ and $v_{a,0,n}^{H,L} = v_{a,0,n}^{H,H}$ which we denote by δ_n^- and δ_n^+ respectively.

When the principal replaces the monitor $n \in \mathbb{N}_0$ periods after a revealing signal, if $\delta \geq \delta_n^+$ the agent provides high effort independently of the informational state. On the other hand, for $\delta < \delta_n^-$, low effort is chosen independently of the informational state. Between these two discounting regions, i.e. $\delta \in [\delta_n^-, \delta_n^+)$, high effort is chosen in the uninformed state and low effort is chosen for n periods, while the tolerant monitor is not replaced after it has been revealed.

In order to conciliate the principal replacement choices with the agent incentives, it is important to understand how δ_n^- and δ_n^+ change with n. We start with the latter.

In the neighborhood of δ_n^+, the agent has a dominant strategy to supply high effort while uninformed. The question is, how does his behavior in the informed state change with n. The larger n is, the more it increases the agent's gains in the informed state; since the monitor is tolerant, the relation is expected to last for a larger number of periods. Low effort in this state increases the immediate expected gains but reduces the life expectancy of the relation due to a higher exposure to punishment. As n gets larger the latter effect becomes more important than the former, favoring high effort behavior. Consequently, we must have $\delta_{n+1}^+ \leq \delta_n^+$ for all $n \in \mathbb{N}$, or more generally

$$\delta^+ = \delta_\infty^+ \leq \ldots\ldots \leq \delta_{n+1}^+ \leq \delta_n^+ \leq \ldots \leq \delta_1^+. \tag{5.4}$$

The cut-off value δ_n^+ is always above δ^+ for all $n \in \mathbb{N}$.[32] This observation supports the results obtained in the previous Sections; when the agent discounts more than δ^+, the principal is better off never replacing the monitor.

[32] Notice also that when $n = 0$, there is no informed state; for that reason δ_0^+ is not defined.

To understand the behavior of δ_n^- with respect to n, we start by noticing that in the neighborhood of δ_n^-, low effort in the informed state is a dominant strategy for the agent. What is not clear is how the agent's behavior in the uninformed state changes with n. A deviation to low effort in the uninformed state increases the likelihood of punishment, but it also accelerates the potential revelation of a tolerant monitor, because $s > r$. The latter effect is stronger the larger n is due to the larger "revelation prize" in the informed state. The former effect is not affected by variations in n. Consequently, a decrease in n reduces the importance of the latter effect by reducing the "revelation prize", which makes for impatient agents with less incentives to deviate in the uninformed state. Then, we must have $\delta_n^- \leq \delta_{n+1}^-$ for all $n \in \mathbb{N}_0$, i.e.

$$\delta^* = \delta_0^- \leq ... \leq \delta_n^- \leq \delta_{n+1}^- \leq ... \leq \delta_\infty^- = \delta^-. \tag{5.5}$$

The two extreme values of this sequence can be derived using the limit cases (5.2) and (5.3) for $i = a$ and $k = 0$, or as in Sections III and IV respectively. For $\delta < \delta^*$, the payoffs of the first periods of the relation become more important and, for that reason, low effort becomes a dominant strategy, disregarding the increased punishment likelihood.

As a summary of the preceding discussion, we have the following relation between sets $[\delta^-, \delta^+) \subseteq [\delta_n^-, \delta_n^+) \subseteq [\delta^*, 1)$ for all $n \in \mathbb{N}_0$.[33]

Lemma 10 characterizes the agent's strategic behavior as a function of the principal replacement choice n. We now need to find the principal's optimal replacement strategy that maximizes her payoff constrained by the associated costs and the agent's incentives, which vary with δ and n.

Denote k^{PCn} as the cut-off cost value below which the principal's participation is guaranteed when she replaces the monitor n periods after the observation of a revealing public signal and $\delta \in [\delta^*, \delta^-)$. Let k^{cut} be a reference threshold cost that solves $v_{p,k,n}^{H,L} = v_{p,k,n+1}^{H,L}$. Unconstrained by any incentives, when $k < k^{cut}$, $n = 0$ is optimal because $v_{p,k,n}^{H,L} > v_{p,k,n+1}^{H,L}$ for all n. Otherwise $n \to \infty$ is the optimal choice. In the discount region $[\delta^-, \delta^+)$ the value $k^{cut} = k^{RC}$ is the replacement threshold above which the principal prefers to never replace

[33] When n gets large we obtain higher order polynomials. The problem becomes untractable and the discount thresholds δ_n^- and δ_n^+ have to be computed numerically. However, since we know how δ_n^- and δ_n^+ vary with n, this is enough for our proposes.

the monitor. Recall that when the agent discounts on this region, the principal always has the option of never replacing the monitor, thus securing a payoff of $v_p^{H,L} > \underline{v}_p$.

Proposition 11 *Suppose that the information is public and $k \geq 0$, the principal's best strategy:*

(i) When $\delta \in [\delta^-, \delta^+)$, is to choose $n = 0$ for $k \in [0, k^{cut})$ and $n \to \infty$ for $k \in [k^{cut}, \infty)$. Where

$$k^{cut} \equiv \frac{\pi_p^H - \underline{v}_p}{1 - \delta(1 - p^T - r)}. \tag{5.6}$$

(ii) When $\delta \in \left[\delta_n^-, \delta_{n+1}^-\right) \subset [\delta^*, \delta^-)$, is to choose $n = 0$ for $k \in [0, k^{cut})$, n for $k \in \left[k^{cut}, k^{PCn}\right)$, and no trade otherwise. Where

$$k^{PCn} \equiv \frac{\pi_p^H - \underline{v}_p}{\delta(1-\beta) r \delta^n (1 - q^T)^n}. \tag{5.7}$$

When the replacement costs are sufficiently low, the principal's best strategy is to replace the monitor after a revealing signal, not allowing the agent to learn. This is true providing $k < k^{cut}$ and $\delta \in [\delta^*, \delta^+)$. The choice $n = 0$ guarantees high effort in both informational states. On the other hand, if replacement costs are large, i.e. $k \geq k^{cut}$ and $\delta \in [\delta^-, \delta^+)$, it is better to never replace the monitor, i.e. $n \to \infty$. However, when $\delta \in \left[\delta_n^-, \delta_{n+1}^-\right) \subset [\delta^*, \delta^-)$ and $k^{cut} \leq k < k^{PCn}$, in order to keep the agent with incentives for high effort in the uninformed state, the principal must replace the monitor n periods after a revealing signal. In other words, to keep the agent with incentives in the uninformed state the principal must allow the agent to benefit from learning during n periods. However, such a demand might be too costly for the principal, i.e. $k \geq k^{PCn}$, in which case she prefers to not trade and get the outside option.

Recall that the optimal behavior for $\delta \in [\delta^+, 1)$ by Proposition 4 is to choose $n \to \infty$.

From (5.6) we can see that a choice $n = 0$ is favoured; when the difference between the principal's stage game gains and the outside option is larger, when the agent is more patient or the punishment and learning probabilities are lower. The effect of these variables is intuitive and not surprising. The value of k^{PCn} in (5.7) is affected in the same way by these variables, but also increases; the larger n becomes, the smaller the proportion of tolerant monitors becomes or larger the punishment probability of a tolerant monitor in the case of low effort.

In the limit, as $n \to \infty$, we obtain expression (3.1) as a particular case of (5.1). This choice cannot improve the principal's payoff. The strategic limited setting of Section III provides a lower bound on the payoffs that the principal can obtain, i.e. $v_p^{H,L}$ for $\delta \in [\delta^-, \delta^+)$ and $\underline{v}_p$ for $\delta \in [\delta^*, \delta^-)$. When the optimal choice takes a finite number of periods and the principal wants to trade, it must be because the replacement costs are sufficiently small and allows for payoff improvements. The following result formalizes this intuition.

Corollary 12 *Suppose that the information is public and $k \geq 0$. If $\delta \in [\delta^-, \delta^+)$ and the optimal choice is $n = 0$ the principal improves her payoffs w.r.t. $v_p^{H,L}$. If $\delta \in [\delta^*, \delta^-)$, the optimal choice n is finite and there is trade, then the principal improves her payoffs w.r.t. $\underline{v}_p$.*

The principal's payoff losses associated with the revelation of the monitor type can be partially recovered when the replacement costs are not too high. These strategies are particularly powerful under public monitoring because the principal enjoys a great replacement precision. However, these strategies require some extra destruction of value due to the replacement costs;[34] consequently, the principal's expected payoff is bounded from v_p^{H*} in the interval $[\delta^*, \delta^+)$, see Corollary 9. The higher the replacement costs, further down the principal's payoff is pushed.

Figure 2 illustrates the value of $v_{p,k,n}^{H,L}$ for the cases where $k = 1$ and $k = 30$ when $\delta \in [\delta^-, \delta^+)$. Since $k^{cut} = 20.24$, in the former case, $n = 0$ is optimal, while in the latter $n \to \infty$ is the optimal replacement choice. Notice also, how both functions converge to $v_p^{H,L}$ as $n \to \infty$. The value of v_p^{H*} and the reference value $v_p^{H,H}$ are also shown.

A. Public Information with Compensation Incentives

We complete this section by discussing the potential of compensation schemes as an instrument to strengthen the agent's incentives for high effort.

It is common in the incentives literature, under the usual constraints, to allow the principal to freely set the compensation. Translated to our setting, the principal would choose a compensation in the informed state and a compensation in the uninformed state. This

[34] Recall that there is also destruction of value due to mistaken punishments on the equilibrium path. Value burned due to replacement cost represents an extra layer in terms of loss in value.

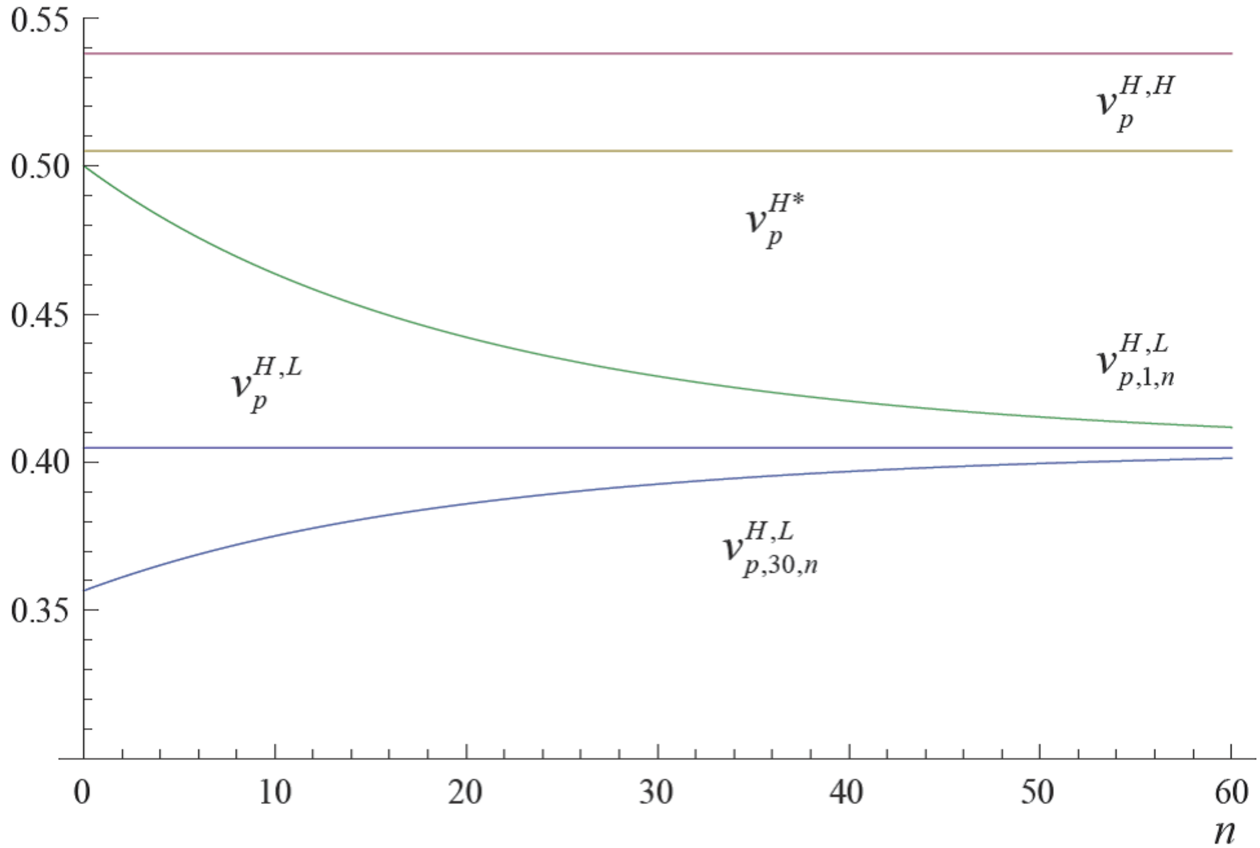

FIG. 2: Principal's payoffs under public monitoring for different replacement costs.

case is particularly penalizing for the agent. Since the incentives for high effort in both informational states are related, the principal would find it optimal to offer a low compensation $w^0 < w$ in the uninformed state, retaining a larger fraction of the surplus in the initial periods ("exploitation" state) of the relation, in exchange for a higher compensation $w^1 > w$ after a revealing signal ("reward" state).[35]

Since players discount the future, the initial larger gains become more relevant; for that reason we might observe not only a payoff improvement above the value v_p^{H*}, but also above the reference value $v_p^{H,H}$. Since the total surplus remains constant, such improvements are made at the agent's expense. (see Footnote 37 below)

[35] It represents an alternative explanation that justifies why young employees tend to work harder in the first period of their careers, relaxing later. See Medoff and Abraham (1981) for some early empirical evidence supporting this fact. Paradoxically it is also in the early stages of an individual career that compensations tend to be lower. See also Gibbons and Murphy (1992).

For this mechanism to work as described, the wage has to be sufficiently flexible. Situations like this one might require some intervention in favor of the agent. Wage restrictions, as we shall discuss, may provide a partial remedy to the problem.

Consider a more realistic scenario where compensations below the market value w cannot be offered.[36] The lower bound restriction w limits the principal's exploitation of the agent in the uninformed state.

To be more concrete, suppose that both the agent and the principal discount the future according to some $\delta \in [\delta^-, \delta^+)$. In this case the principal can offer a higher compensation in the informed state $w^1 > w$, moving δ^+ down towards δ, providing the agent with incentives for high effort in this state.[37] More impatient agents would require a higher w^1 but there is a bound on the compensation that the principal can offer, i.e. $be^H - w^1 \geq \underline{v}_p$, otherwise the principal would prefer not to trade. However, she will still be able to sustain high effort in the uninformed state by offering $w^0 \geq w$, pushing δ^- down towards δ.

The increasing compensation path $w^0 \leq w^1$ has empirical support. During a worker's career the salaries tend to increase above the market reference value. In our setting such an event is caused by the perverse effect associated with the agent's learning.

When high effort is always played in equilibrium, we obtain the same total surplus as in the reference scenario. Compensation schemes as described are transfers of value from the principal to the agent. For that reason, it is not clear whether compensation schemes are payoff superior from the principal's perspective to replacement strategies, in particular when the agent is sufficiently impatient.[38] Improvements over the reference value are impossible for the principal, it is the agent that benefits from the higher compensation. Nonetheless, improvements over the payoff v_p^{H*} are possible when the agent is sufficiently patient.

---

[36] See Taylor (1980) for early empirical observations about wage stickiness. Wage stickiness is particularly strong, even in periods of recession, see Bewley (1999).

[37] There is a second effect. Higher compensation in the informed state helps on the provision of incentives in the uninformed state, since the potential losses associated with an increase in the likelihood of punishment due to a deviation in this stage becomes more important. Consequently, the principal has higher freedom to offer a lower compensation in the uninformed state $w^0 < w$, increasing δ^- up to δ. However, this effect will not be available since the principal is restricted to $w^0 \geq w$. Notice also that if players are very patient, i.e. $\delta \geq \delta^+$, we observe $w^1 \leq w$. Independently of the informational state, the more patient players are, the more surplus the principal appropriates.

[38] A clear answer to this questions depends not only on δ but also on the cost k associated with the replacement of the monitor.

In a multiagent situation, replacement strategies have an amplifying effect. Typically, the same monitor interacts with multiple agents and a specific noisy signal of each agent effort is observed.[39] In this case, once one of the agents has learned that the monitor is tolerant, the principal either decides to increase his compensation or allows him to benefit from his learning. The latter solution leads to relaxation in his performance. The former case leads to a situation where one individual with the same average performance and the same qualifications is paid more than his colleagues. This situation might bring discontentment and a sense of unfairness among the other workers. Moreover, there is also the risk that the monitor type becomes common knowledge due to communication between workers. In order to reestablish the incentives, the principal has to increase the compensation to all the other workers, this solution might be extremely expensive.

Another possibility would be to suggest the replacement of that specific agent or even the whole group. The latter solution seems extremely expensive. In the former case, firing without justification is usually more costly than firing when there is evidence of low performance which can be shown in a court of law. Moreover, a sense of injustice may also emerge among the group members.

The preceding example involves behavioral considerations that go beyond the scope of the present paper. However, it might rationalize, for example, why in some sporting activities, after a sequence of bad results, it is the coach that is replaced not the whole team.

VI. AGENT'S PRIVATE MONITORING WITH COSTLY REPLACEMENT

We now consider the possibility that the agent and the monitor observe the realized output while the principal does not.

This situation is typical in large corporations where the management and the monitoring functions are separated. The principal only observes an aggregate measure of the full output produced by a particular department or by the whole firm. This measure includes the contributions of all the individuals involved in the production process. The principal cannot

[39] This setting is distinct from the one studied in the Seminal work of Holmström (1982b) with moral-hazard-in-teams, where only the team aggregate effort measure is observed. Here, a signal of each agent effort is observed.

disentangle the output of a particular agent from that of the other individuals.[40] It is then the responsibility of lower rank managers (the monitor in our case) to take decisions about a particular individual. The monitor has an informational advantage, for that reason the principal fully delegates the monitoring task.[41] The monitor then reports to the principal when a relevant event has occurred, i.e. an interpretation of low effort.

Since the principal is informed about a punishment event we are able to keep the recursive structure of the problem. To be precise, the setting of this Section is one of private monitoring with communication.[42]

Unlike in the public monitoring case, since the principal does not observe the noisy measure of the agent effort, she misses the monitor's revelation process. To be more concrete, suppose the true monitor type is tolerant and the agent receives a revealing signal and consequently supplies low effort. The principal does not know in which moment in time (or even if) this revelation has occurred. Unlike in Section V, this "reference moment" is not available.

Nonetheless, the principal knows the model and all the associated parameters; moreover, she knows what payoffs are due in case of low and high effort. Given her knowledge about the whole problem, the principal has to design a replacement scheme that is optimal given her "blind" position.

For that reason, her replacement choices are always limited, either because they are premature, in the sense that the agent was still uninformed incurring in an unnecessary cost, or because they are late, in the sense that the monitor type was already revealed and the agent is providing low effort. These imprecisions weaken the effectiveness of replacement

---

[40] If the principal were able to observe the payoffs associated with the effort choices of a particular agent, she would be able to infer the agent noise signal by looking at her own payoff.

[41] That does not mean that the monitor's work is not object of monitoring. We can assume that a higher hierarchy monitor verifies if the monitor is performing his work according to the standards defined by the corporation. This issue is related with the firm's organization design and boundaries. We refer the reader to Rahman (2009), which suggests some interesting answers to the question; who monitors the monitor?

[42] Kandori (2002) presents a description of the challenges associated with private monitoring. Early folk-theorems for private monitoring with communication were obtained by Compte (1998) and Kandori and Matsushima (1998). In our setting we allow for mistaken punishments and we do not consider transfers of value among the players, for that reason we are always bounded far from full efficiency. More recently, Obara (2009) and Zheng (2008) relaxed some of the assumptions of the early contributions.

strategies in private monitoring contexts.[43] However, that does not imply that improvements over the strategic restricted payoff $v_p^{H,L}$ are not possible.

As mentioned in the introductory Section, auditing companies and financial supervision authorities experience a problem with a similar information structure. Auditing companies (the principal) rotate the external auditors (the monitor) on a regular basis.[44] This practice attempts to eliminate what is known in the accounting/auditing jargon as the "familiarity threat" between the client (the agent) and the auditor.[45] Through repeated interaction, the auditor reveals professional and personal characteristics to the client. Learning issues of this kind favor the occurrence of strategic behaviour from the client.[46]

The auditing company is hardly aware of these facts but knows that they are likely to occur. Given the information structure and the agent's strategic behavior; with which frequency should the auditing company rotate its external auditors? This Section provides an answer to this question.

Now, n is the number of periods a given monitor stays in charge, after being hired and after the first signal realization. ($n+1$ is the actual number of periods that the monitor is employed) For example $n=0$ means that the monitor is replaced every period, i.e. is employed for a single period, while $n=1$ means that the monitor is replaced every second period and so on. Notice the difference in the interpretation of n with respect to the public monitor case of Section V.

Denote $\widetilde{v}_{i,k,n}^{0,1}$ as the expected normalized value of the relation for $i \in \{a,p\}$ when the

[43] The same weakness would be present in any other incentives scheme, the difficulty is in the information structure.

[44] The Sarbanes-Oxley Act (enacted on July 30, 2002) in Section 203, requires the lead audit partner and audit review partner (or concurring reviewer) to be rotated every five years on public company audits as well as on audits of issuers.

[45] Most of the literature on incentives studies this practice as a remedy to the breakout of collusive arrangements. Tirole (1986) points out that when the monitoring task is delegated to a third party, problems related with the monitor's conflicting interests might arise. See Laffont and Tirole (1991) and Kofman and Lawarree (1993) for further developments and extensions. In this paper we are not so concerned with delegation effects of this kind.

[46] The "familiarity threat" may also be caused by collusion between the agent and the monitor, but even in this case some prior learning has to occur. A rational dishonest client would not bribe an external auditor without a prior observation of his personal character traits, otherwise he could place himself in a worse situation.

principal replaces the monitor $n \in \mathbb{N}_0$ periods after have hiring him, paying the replacement cost $k \geq 0$ every time, and the agent effort choice in each informational state is $\{e^0, e^1\}$.

Lemma 13 *Suppose that the information is the agent's private monitoring, $k \geq 0$, the principal replaces the monitor $n \in \mathbb{N}_0$ periods after hiring him and the agent chooses $\{e^0, e^1\} = \{e^H, e^L\}$. The infinitely repeated normalized expected payoff for player $i \in \{a, p\}$ is*

$$\widetilde{v}^{H,L}_{i,k,n} = \frac{\Pi_i^H \frac{1-\delta^{n+1}z^{n+1}}{1-\delta z} + \delta y \Pi_i^L \frac{(z-x)(1-\delta^n z^n) - x(z^n - x^n)\delta^n(1-\delta z)}{(z-x)(1-\delta x)(1-\delta z)} - k(1-\delta)}{1 - \delta^{n+1}\frac{(z-x)z^{n+1} + y(z^{n+1} - x^{n+1})}{z-x}} + k(1-\delta), \tag{6.1}$$

where $x \equiv 1 - q^T$, $y \equiv (1-\beta) r$, $z \equiv 1 - p^T - r$, $\Pi_i^H \equiv (1-\delta)\pi_i^H + \delta p^E \underline{v}_i$ and $\Pi_i^L \equiv (1-\delta)\pi_i^L + \delta q^T \underline{v}_i$. When $i = a$ we have $k = 0$.

Expression (6.1) has the following asymptotic properties. When we let $n \to \infty$ we obtain $v_i^{H,L}$ as expression (5.2) in the previous Section. While, if $n = 0$, we obtain

$$\widetilde{v}^{H,L}_{i,k,0} = \frac{(1-\delta)\pi_i^H + \delta p^E \underline{v}_i - k\delta(1-\delta)\left(1 - p^E\right)}{1 - \delta\left(1 - p^E\right)}. \tag{6.2}$$

When $k = 0$ we get the expression v_i^{H*}.

Since $1 - p^E > (1-\beta) r$, the principal's payoff in (6.2) is smaller than in (5.3). The difference in payoffs reflects the loss in precision of replacement strategies under private monitoring. In other words, in Section V, the replacement of the monitor was only an issue after the arrival of a revealing signal which was publicly observed, while under the agent's private monitoring, replacement is a possibility from the first period of the relation.

As in Section V, we focus our attention on the discounting interval $[\delta^*, \delta^+)$. In particular, in discounting region $[\delta^*, \delta^-)$, we have a sequence of δ_n^- ordered as in (5.5) of Section V (for the intuition, the reader is referred to the discussion in that Section). However, under private monitoring δ_n^- is smaller. This is because the interpretation of n under private and public monitoring are different. With public monitoring, n is the number of periods that the tolerant monitor remains in charge after a revealing signal, while under private monitoring, n is the number of periods that a given monitor stays in charge after the first signal. Consequently, for the same n, replacement is more frequent under private than under public monitoring, justifying the difference of δ_n^- under the different information structures.[47]

[47] On the other hand, δ_n^+ under private monitoring is larger than δ_n^+ under public monitoring. Which is consistent with the observation made about δ_n^-. An equal ordered sequence of δ_n^+ as in (5.4) is obtained.

To better distinguish between both information structures, suppose that a revealing signal occurs at time $t \geq 1$. When the signals are public, the monitor stays in charge for a total of $t+n$ repetitions of the stage game. While, if monitoring is private, the monitor stays in charge for a total of $1+n$ repetitions of the stage game, and we might have $t \gtreqless n+1$.

Rotation of the monitor under public information is more accurate but the replacement cycle $t+n$ is stochastic. This is the case because the reference (revelation) period t is random, unknown ex-ante but observed ex-post. Replacement strategies under private monitoring are less accurate, since t is random and not known by the principal even ex-post. For that reason, the replacement cycle n is defined and known ex-ante.

Denote $k^{n,n+1}$ as the cut-off point which, infinitesimally below n, is an optimal choice for the principal and, infinitesimally above $n+1$, is optimal. It is a transition threshold between replacement choices. For fixed k, when $k^{n-1,n} \leq k \leq k^{n,n+1}$, n is an unconstrained optimal choice. Let $k^{\infty} \equiv \lim_{n\to\infty} k^{n,n+1}$ be the value above which the principal's optimal unconstrained choice is to never replace the monitor. We say unconstrained because we are not considering any incentives constraint.

For $\delta \in [\delta^-, \delta^+)$, let k^{RCn} denote the per replacement cost threshold below which the principal's optimal choice can be n. When $k^{RC\infty} \leq k$, the choice $n \to \infty$ must be optimal. Similarly, for $\delta \in [\delta^*, \delta^-)$, let k^{PCn} denote the participation condition below which n can be optimal. To keep the text clean, the functional form of each of these objects can be found in the Proof of the following result.

Proposition 14 *Suppose that the information is the agent's private monitoring and $k \geq 0$. Let τ denote the principal optimal choice.*

(i) For $\delta \in [\delta^-, \delta^+)$: If $k < k^{\infty}$, then $\tau = \inf\left\{n, n+1, ... : k < k^{RC\tau}\right\}$. Otherwise, i.e. $k^{\infty} = k^{RC\infty} \leq k$, $\tau \to \infty$.

(ii) For $\delta \in \left[\delta^-_m, \delta^-_{m+1}\right) \subseteq [\delta^*, \delta^-)$ and $n \leq m$: If $k < k^{\infty}$, then $\tau = \inf\left\{n, n+1, ..., m : k < k^{PC\tau}\right\}$, while if $k^{\infty} \leq k < k^{PCm}$, then $\tau = m$. Otherwise, i.e. $k \geq k^{PCm}$, there is no trade.

(iii) For $\delta \in \left[\delta^-_m, \delta^-_{m+1}\right) \subseteq [\delta^*, \delta^-)$ and $n > m$: If $k < k^{PCm}$, then $\tau = m$. Otherwise, i.e. $k \geq k^{PCm}$, there is no trade.

The principal participation and consequent employment of replacement strategies is guaranteed when $\widetilde{v}_{p,k,n}^{H,L} > \underline{v}_p$ for $\delta \in [\delta^*, \delta^-)$ and $\widetilde{v}_{p,k,n}^{H,L} > v_p^{H,L}$ for $\delta \in [\delta^-, \delta^+)$. The agent incentives for high effort in the uninformed state are guaranteed for $\delta \geq \delta_n^-$. When we join together these restrictions with the unconstrained optimal choice we obtain the equilibrium strategies for the agent and the principal.

The agent is rational and knows the replacement cycle. Consequently, when the principal replaces the monitor, the agent also shifts from low to high effort. Then he waits for the occurrence of a potential revealing signal before the next monitor replacement, in order to enjoy the remaining period providing low effort. The principal's strategy takes into account this strategic behavior.

In the discounting region $[\delta^-, \delta^+)$ the principal's best replacement strategy is no longer an exclusive choice between $n = 0$ and $n \to \infty$, as we found in Proposition 11. Now the equilibrium is more sensitive to the value k. In fact any n can be optimal, for that reason high effort in any informational state is only possible for $k < k^{0,1}$.[48] The reason is the trade-off between more frequent replacement, more costly but less likely to allow the agent to learn, and less frequent replacement, cheaper but with a higher probability that the principal's payoffs will be penalized by the revelation of a tolerant monitor.

When $\delta \in \left[\delta_m^-, \delta_{m+1}^-\right) \subseteq [\delta^*, \delta^-)$ and the unconstrained optimal choice $n \leq m$, the principal can choose n if $k < k^{PCn}$. Otherwise, since k^{PCn} is strictly increasing with n, she has to move way from the optimum in the direction of m. However, it might still happen that $k \geq k^{PCm}$, in which case the principal should not trade because replacement costs are too high. A choice above m is cheaper but does not provide the agent with incentives.

When $\delta \in \left[\delta_m^-, \delta_{m+1}^-\right) \subseteq [\delta^*, \delta^-)$ and the unconstrained optimal choice $n > m$, the principal must increase the replacement frequency, moving way from the optimum, down towards m to provide the agent with incentives in the uninformed state. However, if $k \geq k^{PCm}$ the provision of incentives is too costly and the principal should not enter into the relation.

When $n \to \infty$ we obtain expression (5.2) as a particular case of (6.1). It is better to allow the agent to provide low effort in the informed state than to make any costly replacement,

[48] It can be shown that $k^{0,1}$ is a very small positive number.

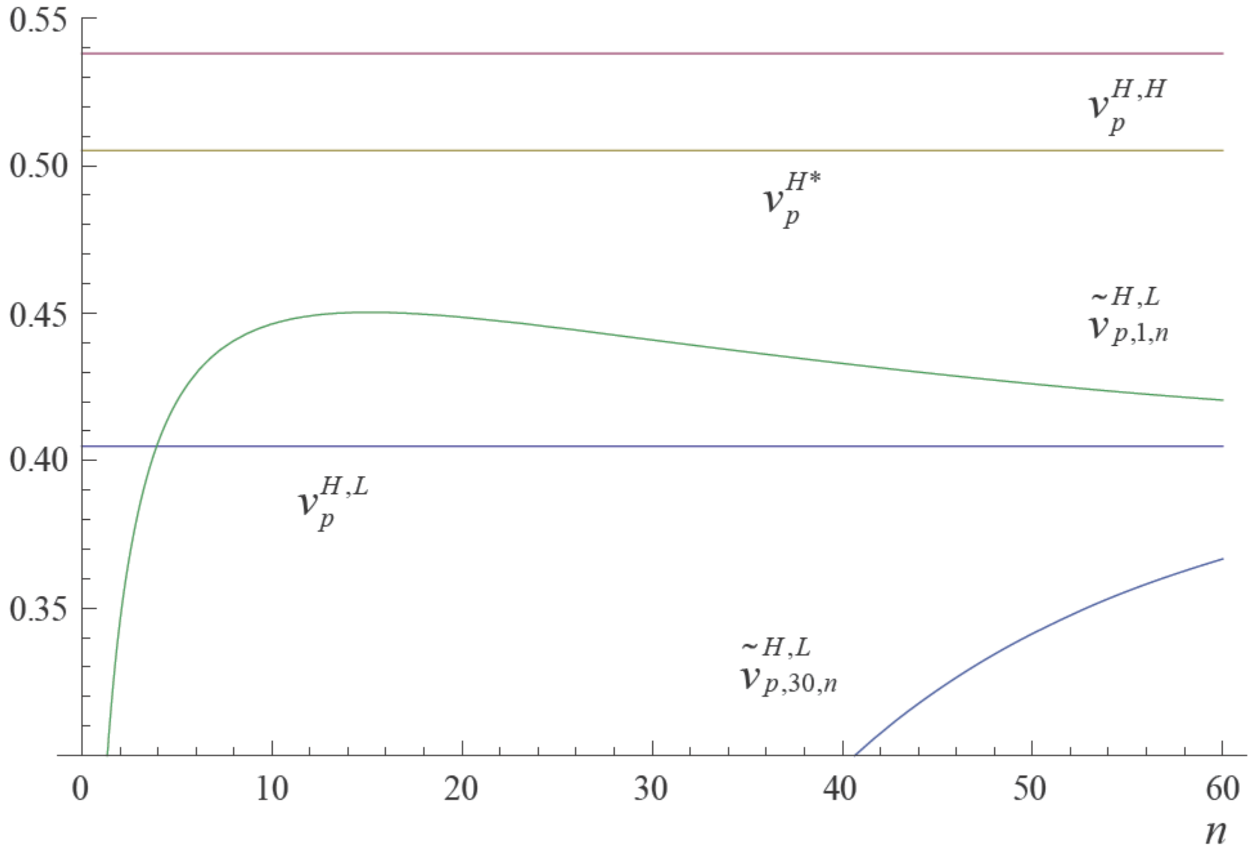

FIG. 3: Principal's payoffs under private monitoring for different replacement costs.

see Corollary 9. Nonetheless, for sufficiently low replacement costs, it is possible to make a payoff improvement using replacement strategies even without observing the realized output. The following result is in everything similar to Corollary 12.

Corollary 15 *Any equilibrium of Proposition 14 with a finite choice n and trade, improves the principal's payoff* $\widetilde{v}_{p,k,n}^{H,L}$ *over the payoffs associated with Proposition 4, i.e.* $v_p^{H,L}$ *for* $\delta \in [\delta^-, \delta^+)$ *and* $\underline{v}_p$ *for* $\delta \in [\delta^*, \delta^-)$.

For $\delta \in [\delta^*, \delta^-)$ we require $\widetilde{v}_{p,k,n}^{H,L} > \underline{v}_p$, which is the participation or a replacement constraint that can only be satisfied if n is finite. When $n \to \infty$ no trade is an equilibrium. For $\delta \in [\delta^-, \delta^+)$ we require that $\widetilde{v}_{p,k,n}^{H,L} > v_p^{H,L}$, which is only possible if replacement strategies are employed.[49] If $n \to \infty$ we have $\widetilde{v}_{p,k,n}^{H,L} \to v_p^{H,L}$, i.e. the worst case scenario for the principal,

[49] For that reason we call it a replacement constraint. Since $v_p^{H,L}$ is always larger than $\underline{v}_p$ we cannot talk about a participation constraint in the strict sense.

that she can always guarantee by not replacing the monitor.

Figure 3 illustrates the value of $\widetilde{v}^{H,L}_{p,k,n}$ for the cases where $k = 1$ and $k = 30$ when $\delta \in [\delta^-, \delta^+)$. In the former case $k^{13,14} = 0.93 < k^{14,15} = 1.04$ and $k^{RC14} = 2.89$, then $n = 14$ is the optimal choice. Since $k^\infty = 20.24$ when $k = 30$, $n \to \infty$ is optimal. Both functions converge to $v^{H,L}_p$ as $n \to \infty$.

A. Agent's Private Information with Compensation Incentives

We now comment on the possibility of compensation based incentives. As in Subsection V A we assume that the agent will not accept to work for less than the market wage. The principal is allowed to raise the compensation if she considers it convenient but cannot decrease it.

To understand how compensation incentives can be used in a private monitor setting, consider the following strategy. During the first $n \in \mathbb{N}_0$ periods of the relation, the principal pays the compensation $w^0 \geq w$. In the following periods, she switches to the compensation $w^1 \geq w$, with $w^0 < w^1$. The idea is to offer a lower compensation sufficient to provides the agent with incentives for high effort in the early periods, where it is less likely that the monitor type has been revealed to be tolerant, and then adjust to a higher compensation when it is more likely that the monitor has been revealed.

Now, together with the uninformed and the informed states of the relation, we have the low compensation stage and the high compensation stage. A strategy for the principal is a choice of w^0, w^1 and the stage separating period n. Again we might have $n = 0$, i.e. the relation starts in the high compensation stage, or $n \to \infty$, i.e. the relation remains in the low compensation stage forever, or an optimal intermediate choice of n, i.e. the relation passes through both stages.

The informational disadvantage of the principal with respect to the agent will necessarily reflects in a lower payoff for the former when compared with the case where the realized output is publicly monitored. Again, we expect mixed superiority of replacement strategies with respect to compensation incentives. The latter must be stronger when the agent is more patient and/or replacement costs are sufficiently large.

The two stage compensation scheme discussed here is similar to the existing one in the public sector. There is a distance, not only physical but also in monitoring terms, between the central authority and the lower hierarchical levels. The performance evaluation and the functioning of the associated public office is usually based on general reports. For that reason promotions, measured in compensation benefits, are usually independent of the performance but rather depend on years accumulated in service.

VII. FINAL COMMENTS

In many economic situations of interest managers have the necessity to delegate some of their tasks - in this model, the monitoring activities, which are crucial for the regular functioning and expansion of their businesses. The degree of delegation in some sense determines the subsequent information structure. Full delegation leads to an information structure similar to the agent private monitoring case discussed in this paper. Partial delegation is closer to public monitoring information structures. The present paper provides some results, about the optimal strategic behavior from the principal's perspective, for dealing with the negative effects associated with the revelation of specific organizational aspect which might be the object of adverse strategic behavior. We choose the monitor type to be the unknown piece of information that a potentially strategic agent might take advantage of once informed, but the spectrum of situations with similar characteristics is larger.

The principal usually has more freedom in the choice of the incentives schemes, in this paper we focus on replacement strategies. As mentioned before, the revelation of the player's type through repeated interaction is not a new finding. However, the way such a problem is modeled in this paper is novel.

Many questions are left open. For example, a clearer connection with the existing theories in multiple and common agency, renegotiation-proof, incomplete contacts, private evaluations, information sharing, etc. Compensation incentives were discussed but not formalized.

We also did not cover all potential information structures, for example the possibility of the realized output being exclusively observed by the principal (and/or the monitor) or when the principal holds prior private information about the monitor type. These cases capture situations where the principal has access to relevant information that for some reason, intentionally or not, it is blocked to the agent. We expect the principal's informational advantage to help her in achieving higher payoffs. In the former possibility, the principal can be more efficient in her replacement choices, in particular if the tolerant type is preferred.

However, the principal cannot replace the monitor successively until a tolerant type appears because that behavior would reveal the monitor type to the agent. There is here a trade off between replacement costs and payoff gains with a tolerant monitor. An optimal strategy for the principal must require some degree of randomization between replacement choices.

The contrast between the agent and monitor replacement is also an interesting point. Another possibility is to allow the monitor to play a strategic role or even to remove him and consider a strategic principal with an unknown type to the agent. This lead us to dynamic incentive problems of incomplete information, typically harder to handle but very rich in strategic terms.

Also interesting, but from a different perspective, is the introduction of new ingredients into the problem, and more empirical and experimental work on the subject are the next steps towards a better understanding of this type of revelation problems. Such research should also provide us with recommendations on how we could implement the proposed solutions in our organizations.

REFERENCES

1. Abreu, D., D. Pearce and E. Stacchetti (1986). "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory, 39, 251-269.

2. Abreu, D., D. Pearce and E. Stacchetti (1990). “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring.” Econometrica, 58, 1041-1063.

3. Alchian, A. and H. Demsetz (1972). "Production, Information Costs, and Economic Organization." American Economic Review, 62, 777-795.

4. Baker, G., R. Gibbons and K. J. Murphy (1994). "Subjective Performance Measures in Optimal Incentive Contracts." Quarterly Journal of Economics, 109, 1125-1156.

5. Baker, G., R. Gibbons and K. J. Murphy (2002). "Relational Contracts and the Theory of the Firm." Quarterly Journal of Economics, 117, 39–84.

6. Bewley, T. F. (1999) Why Wages Don’t Fall During a Recession. Harvard University Press.

7. Bolton P., M. Dewatripont (2005) Contract Theory. MIT Press, Cambridge, MA.

8. Bull, C. (1987). "The Existence of Self-Enforcing Implicit Contracts." Quarterly Journal of Economics, 102, 147–159.

9. Cole H., J. Dow, W. B. English (1995). "Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt." International Economic Review, 36, 365-385.

10. Compte, O. (1998). "Communication in Repeated Games with Imperfect Private Monitoring," Econometrica, 66, 597-626.

11. Cripps, M. W., J. C. Ely, G. J. Mailath, and L. Samuelson (2008). "Common Learning." Econometrica, 76, 909-933.

12. Cripps, M. W., G. J. Mailath, and L. Samuelson (2007). "Disappearing Private Reputations in Long-Run Relationships." Journal of Economic Theory, 134, 287–316.

13. Cripps, M. W., G. J. Mailath, and L. Samuelson (2004). "Imperfect Monitoring and Impermanent Reputations." Econometrica, 72, 407–432.

14. Ferguson, T. (1989). "Who solved the secretary problem?" Statistical science, 4, 282-296.

15. Fuchs, W. (2007). "Contracting with Repeated Moral Hazard and Private Evaluations," American Economic Review, 97, 1432-1448.

16. Fudenberg, D., D. Levine (1994). "Efficiency and Observability with Long-Run and Short-Run Players." Journal of Economic Theory, 62, 103–135.

17. Fudenberg, D., D. Levine and E. Maskin (1994). "The Folk Theorem with Imperfect Public Information." Econometrica, 62, 997-1040.

18. Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

19. Fudenberg, D. and J. Tirole (1991) Game Theory, MIT Press, Cambridge, MA.

20. Fudenberg, D. and J. Tirole (1986), "A "Signal-Jamming" Theory of Predation." Rand Journal of Economics, 17, 366–376.

21. Freeman, P. (1983). "The secretary problem and its extensions: A review." International Statistical Review, 51, 189-206.

22. Gibbons, R. and K. J. Murphy (1992). "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." The Journal of Political Economy, 100, 468-505.

23. Gossner, O. and T. Tomala (2009). "Repeated Games with Complete Information", in Meyers, Robert (Ed.) Encyclopedia of Complexity and Systems Science, pp 7616–7630. Springer New York.

24. Holmström, B. (1982a). "Managerial Incentive Problems - A Dynamic Perspective", in Essays in Economics and Management in Honor of Lars Wahlbeck. Helsinki: Swedish School of Economics. (See also Review of Economic Studies, (January 1999), 66, 169-182)

25. Holmström, B. (1982b). "Moral Hazard in Teams.", Bell Journal of Economics, 10, 74-91.

26. Kandori, M. (2002). "Introduction to Repeated Games with Private Monitoring," Journal of Economic Theory, 102, 1-15.

27. Kandori, M and H. Matsushima (1998) "Private Observation, Communication and Collusion," Econometrica, 66, 627-652.

28. Kennan J. and R. Wilson (1989). "Strategic Bargaining Models and Interpretation of Strike Data," Journal of Applied Econometrics, 4, Supplement: Special Issue on Topics in Applied Econometrics, S87-S130.

29. Klein B., and K. B. Leffler (1981). "The Role of Market Forces in Assuring Contractual Performance." The Journal of Political Economy, 89, 615-641.

30. Kofman, F. and J. Lawarree (1993), "Collusion in Hierarchical Agency", Econometrica, 61, 629-656.

31. Laffont, J. and J. Tirole (1991), "The Politics of Government Decision-Making: A Theory of Regulatory Capture", Quarterly Journal of Economics, 106, 1089-1127.

32. Leibenstein, H. (1987) Inside the Firm: The Inefficiencies of Hierarchy, Harvard University Press, Cambridge, MA.

33. Levin J. (2003). "Relational Incentive Contracts." American Economic Review, 93, 835-857.

34. MacLeod W. B., (2003). "Optimal contracting with subjective evaluation." American Economic Review, 93(1), 216-240.

35. MacLeod W. B., and J. M. Malcomson (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." Econometrica, 57, 447-480.

36. Mailath, G. and L. Samuelson (2001), "Who Wants a Good Reputation?" The Review of Economic Studies, 68, 415-441.

37. Mailath, G. and L. Samuelson (2006) Repeated Games and Reputations: Long-run Relationships. Oxford University Press, New York.

38. Medoff, J. and K. Abraham (1980). "Experience, Performance and Earnings", Quarterly Journal of Economics, 95, 703-736.

39. Mirman, L., L. Samuelson, and A. Urbano (1993). "Duopoly Signal Jamming," Economic Theory, Springer, 3(1), 129-49.

40. Obara, I. (2009). "Folk Theorem with Communication," Journal of Economic Theory, 144, 120-134.

41. Phelan, C. (2006). "Public trust and government betrayal," Journal of Economic Theory, 130, 27-43.

42. Radner, R. (1985). "Repeated Principal-Agent Games with Discounting," Econometrica, 53, 1173–1198.

43. Rahman, D. (2009). "But Who will Monitor the Monitor?," mimeo.

44. Renault, J. (2009). "Repeated Games with Incomplete Information", in Meyers, Robert (Ed.) Encyclopedia of Complexity and Systems Science, pp 7630-7651. Springer New York.

45. Rubinstein, A. (1979b). “An Optimal Conviction Policy for Offenses That May Have Been Committed byAccident,” in Applied Game Theory, ed. by S. J. Brams, A. Schotter, and G. Schwodiauer, 406–413. Physical-Verlag, Würzburg.

46. Rubinstein, A., and M. E. Yaari (1983). “Repeated Insurance Contracts and Moral Hazard,” Journal of Economic Theory, 30, 74–97.

47. Salanié, B. (2005) The Economics of Contracts: A Primer, 2nd Edition. MIT Press, Cambridge, MA.

48. Sannikov, Y. (2007). “Games with Imperfectly Observable Actions in Continuous Time,” Econometrica, 75, 1285–1329.

49. Shapiro, C. and J. Stiglitz (1984) "Equilibrium Unemployment as a Worker Discipline Device," American Economic Review, 74(3), 433-444.

50. Stiglitz, J. (1974). "Incentives and Risk Sharing in Sharecropping," The Review of Economic Studies 41 (2), 219-255.

51. Taylor, J. (1980). "Aggregate Dynamics and Staggered Contracts." Journal of Political Economy, 1, 1-23.

52. Tirole, J. (1986). "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations", Journal of Law, Economics and Organization, 2, 181-214.

53. Zheng, B. (2008). “Approximate Efficiency in Repeated Games with Correlated Private Signals,” Games and Economic Behavior, 63, 406-416.

APPENDIX A: PROOFS OF LEMMAS, PROPOSITIONS AND COROLLARIES

The proof of the various Lemmas makes extensive use of the dynamic programming methods developed by Abreu, Pearce and Stacchetti (1986, 1990).

Proof of Lemma 2. Consider player $i = a$. Suppose that in the beginning of the game, i.e. state 0, the agent selects an effort e^H, receiving an expected payoff π_a^H. In the end of the first period, if a low output is observed, the game enters in the punishment stage with expected payoff $\underline{v}_a / (1-\delta)$. This event occur with probability p^E. If a high output is observed, it might be uninformative in which case the value of the game for the agent associated with action e^H is $v_a^{H,L} / (1-\delta)$. Notice that we have a recursive pattern here, this case is equivalent to a repetition of the initial game one period later. This event occur with probability $1 - p^E - (1-\beta) r$.

With probability $(1-\beta) r$ the signal might be revealing, in which case the agent adjust his effort accordingly to e^L, obtaining an expected payoff of π_a^L for the following period. In this case there is no more learning, and we have a simple recursive structure. With probability q^T the agent is punished in the following period and with the remaining probability he obtains the value $v_a^{\cdot\cdot,L} / (1-\delta)$.

Formally we have two recursive patterns, they are respectively

$$v_a^{H,L} = (1-\delta)\,\pi_a^H + \delta\left[p^E \underline{v}_a + (1-\beta)\, r v_a^{\cdot\cdot,L} + \left(1 - p^E - (1-\beta)\, r\right) v_a^{H,L}\right],$$

and

$$v_a^{\cdot\cdot,L} = (1-\delta)\,\pi_a^L + \delta\left[q^T \underline{v}_a + \left(1-q^T\right) v_a^{\cdot\cdot,L}\right],$$

which can be solved for $v_a^{H,L}$ to obtain (3.1). Reasoning in a similar way we obtain expression (3.1) for the principal. Employing the substitution suggested in the text we obtain the payoffs $v_i^{L,L}$ and $v_i^{H,H}$ for $i \in \{a, p\}$. ■

Proof of Proposition 4. First we search for the condition $\delta < \delta^+$. After observing an informative signal the agent chooses e^L if $v_a^{\cdot\cdot,H} < v_a^{\cdot\cdot,L}$. Where

$$v_a^{\cdot\cdot,H} \equiv \frac{(1-\delta)\,\pi_a^H + \delta p^T \underline{v}_a}{1-\delta\left(1-p^T\right)} \text{ and } v_a^{\cdot\cdot,L} \equiv \frac{(1-\delta)\,\pi_a^L + \delta q^T \underline{v}_a}{1-\delta\left(1-q^T\right)}, \tag{A1}$$

with $v_a^{\cdot\cdot,H}$ being the value of the infinitely repeated state 1 subgame for the informed agent when he chooses e^H and $v_a^{\cdot\cdot,L}$ has a similar interpretation but with the agent choosing e^L. Rearranging for δ we obtain

$$\delta < \frac{\pi_a^L - \pi_a^H}{\left(\pi_a^L - \pi_a^H\right) + q^T\left(\pi_a^H - \underline{v}_a\right) - p^T\left(\pi_a^L - \underline{v}_a\right)} \equiv \delta^+. \tag{A2}$$

Since $\pi_a^L > \pi_a^H$ and if $q^T\left(\pi_a^H - \underline{v}_a\right) > p^T\left(\pi_a^L - \underline{v}_a\right)$, i.e. condition (3.5) holds, we have $0 < \delta^+ < 1$.

The agent infinitely repeated game expected payoff when he supplies e^H while uninformed and e^L in the state 1 is $v_a^{H,L}$, and given by (3.1), the infinitely repeated game expected payoff when supplying e^L in any informational state is denoted and given by

$$v_a^{L,L} \equiv \frac{(1-\delta)\,\pi_a^L + \delta q^E \underline{v}_a}{D^L} + \delta\,(1-\beta)\,\frac{s}{D^L}\,\frac{(1-\delta)\,\pi_a^L + \delta q^T \underline{v}_a}{1-\delta\,(1-q^T)}, \tag{A3}$$

where $D^L \equiv 1-\delta\left(1-q^E-(1-\beta)\,s\right)$. Solving $v_a^{H,L} \geq v_a^{L,L}$ for δ with equality we obtain the expression for δ^-. To guarantee the nonemptiness of the interval $[\delta^-, \delta^+)$, i.e. $\delta^+ > \delta^-$, we plug δ^+ in $v_a^{H,L}$ and $v_a^{L,L}$. After some algebra simplifications the inequality relation $v_a^{H,L} > v_a^{L,L}$ becomes

$$\frac{r\beta\left(\pi_a^H-\pi_a^L\right)\left(q^T\left(\pi_a^H-\underline{v}_a\right)-p^T\left(\pi_a^L-\underline{v}_a\right)\right)}{r\left(\pi_a^L-\pi_a^H\right)-p^T\left(\pi_a^H-\underline{v}_a\right)+q^T\left(\pi_a^H-\underline{v}_a\right)} > \frac{\beta s\left(\pi_a^H-\pi_a^L\right)\left(q^T\left(\pi_a^H-\underline{v}_a\right)-p^T\left(\pi_a^L-\underline{v}_a\right)\right)}{s\left(\pi_a^L-\pi_a^H\right)-p^T\left(\pi_a^L-\underline{v}_a\right)+q^T\left(\pi_a^L-\underline{v}_a\right)},$$

further manipulations lead us to

$$\left(\frac{s}{\left(\pi_a^L-\pi_a^H\right)s+\left(\pi_a^L-\underline{v}_a\right)\left(q^T-p^T\right)} - \frac{r}{\left(\pi_a^L-\pi_a^H\right)r+\left(\pi_a^H-\underline{v}_a\right)\left(q^T-p^T\right)}\right)$$
$$\times\beta\left(\pi_a^L-\pi_a^H\right)\left(q^T\left(\pi_a^H-\underline{v}_a\right)-p^T\left(\pi_a^L-\underline{v}_a\right)\right) > 0.$$

From which we require conditions (3.5) and (3.6) to hold simultaneously. Since $\pi_a^L > \pi_a^H > \underline{v}_a$, it is easy to show that $\delta^- > 0$. Putting all together conditions (3.5) and (3.6) guarantee that $0 < \delta^- < \delta^+ < 1$. ■

Proof of Corollary 6. The reference payoff $v_i^{H,H}$ is given by (3.7). The expression for $v_i^{H,L}$ is given by (3.1) in Lemma 2.

(i) The conditions of Proposition 4 establish that $v_a^{H,L} > v_a^{H,H}$ for $\delta < \delta^+$ where δ^+ solves $v_a^{H,L} = v_a^{H,H}$.

(ii) Solving $v_p^{H,L} \leq v_p^{H,H}$ for δ and using the fact that $\pi_p^L = \underline{v}_p$ we obtain

$$\delta \leq \frac{1}{1+q^T} \equiv \delta^{ref}. \tag{A4}$$

It enough to show that $\delta^+ \leq \delta^{ref}$; it lead us to the following condition

$$0 \leq \left(\pi_a^L-\underline{v}_a\right)\left(q^T-p^T\right),$$

which is satisfied with strict inequality, since the right-hand side is strictly positive, i.e. $\pi_a^L > \underline{v}_a$ and $q^T > p^T$. Then $v_p^{H,L} < v_p^{H,H}$ for all $\delta < \delta^+$.

(iii) Solving $v_a^{H,H} + v_p^{H,H} > v_a^{H,L} + v_p^{H,L}$ for δ we obtain three roots $\delta > 0$, $\delta < 1$ and

$$\delta > \frac{\left(\pi_p^H-\underline{v}_p\right)-\left(\pi_a^L-\pi_a^H\right)}{\left(\pi_a^H+\pi_p^H-\underline{v}_p\right)\left(1-q^T\right)-\pi_a^L\left(1-p^T\right)+\underline{v}_a\left(q^T-p^T\right)}.$$

The third root is larger than one if

$$q^T\left(\pi_a^H-\underline{v}_a\right)+q^T\left(\pi_p^H-\underline{v}_p\right)>p^T\left(\pi_a^L-\underline{v}_a\right),$$

which is guaranteed by condition (3.5). ■

Proof of Proposition 7. The problem is equivalent to a infinitely repeated game without learning and the monitor type known to be θ^E. In this case the players payoffs are simply

$$v_i^{H*}=\frac{(1-\delta)\,\pi_i^H+\delta p^E\underline{v}_i}{1-\delta\,(1-p^E)}\ \textit{and}\ v_i^{L*}=\frac{(1-\delta)\,\pi_i^L+\delta q^E\underline{v}_i}{1-\delta\,(1-q^E)},\tag{A5}$$

when the agent provides high and low effort respectively and $i\in\{a,p\}$. The recursive structure is simple; the agent supplies a given effort and with probability p^E or q^E he is punished or with the remaining probabilities the same pattern is repeated. Solving $v_a^{H*}\geq v_a^{L*}$ for δ we obtain

$$\delta\geq\frac{\pi_a^L-\pi_a^H}{\left(\pi_a^L-\pi_a^H\right)+\left(\pi_a^H-\underline{v}_a\right)q^E-\left(\pi_a^L-\underline{v}_a\right)p^E}\equiv\delta^*.$$

Then $\delta^*\in(0,1)$ if $q^E\left(\pi_a^H-\underline{v}_a\right)>p^E\left(\pi_a^L-\underline{v}_a\right)$, which is a general version of condition (3.5) for all $\beta\in(0,1)$. Since p^E and q^E depends linearly on β, and the sum of (3.5) and (3.6) gives $q^S\left(\pi_a^H-\underline{v}_a\right)>p^S\left(\pi_a^L-\underline{v}_a\right)$, the lowest difference $q^E\left(\pi_a^H-\underline{v}_a\right)-p^E\left(\pi_a^L-\underline{v}_a\right)$ must be reached at $\beta\to 0$ by condition (3.6). This difference is always positive. The value of δ^+ is given by (A2) and is larger than δ^* when (3.6) holds.

To show that $\delta^*<\delta^-$, just replace δ^* for δ in $v_a^{H,L}\leq v_a^{L,L}$ and rearrange to obtain

$$(1-\beta)\left(\pi_a^L-\pi_a^H\right)^2\left(s\left(\pi_a^H-\underline{v}_a\right)-r\left(\pi_a^L-\underline{v}_a\right)\right)\left(q^E\left(\pi_a^H-\underline{v}_a\right)-p^E\left(\pi_a^L-\underline{v}_a\right)\right)>0,$$

which, following the previous argument, is satisfied when both (3.5) and (3.6) hold. ■

Proof of Proposition 8. Manipulate the inequality $v_i^{H,H}>v_i^{H*}$, which the expressions are given by (3.7) and (A5) respectively, we obtain

$$r\,(1-\beta)\left(p^E-p^T\right)\left(\pi_i^H-\underline{v}_i\right)(1-\delta)\,\delta^2>0,$$

which is clearly larger than zero for all $\delta\in(0,1)$, since $p^E=p^T+\beta r>p^T$, $\pi_i^H>\underline{v}_i$ for $i\in\{a,p\}$ and $\beta\in(0,1)$. Participation is guaranteed if $v_i^{H*}\geq\underline{v}_i$. ■

Proof of Corollary 9. Lets start noticing that replacement cost are only incurred by the principal. (i) For $\delta\in[\delta^+,1)$ the market compensation is sufficient to keep the agent with incentives. The principal never replaces the monitor.

(ii) By Proposition 7, without replacement costs and with full freedom in the monitor replacement the principal cannot obtain more than v_p^{H*} then she must not be able to do more if we introduce replacement costs. In this case the agent obtain v_a^{H*} as his worst payoff. On the other hand if costs are too high and in consequence payoffs are expected to fall bellow $v_p^{H,L}$, then principal has always the option to never replace the monitor, as in Proposition 4, guaranteeing at least a payoff of $v_p^{H,L}$. In this case the agent benefit from the principal replacement passivity and obtain the payoff $v_a^{H,L}$.

(iii) Without replacement strategies both player would obtain $\underline{v}_p$ and $\underline{v}_a$. Costless replacement strategies expand the discount region to δ^*. If replacement strategies improve over $v_p^{L,L} = \underline{v}_p$ then the principal must obtain a payoff of at most v_p^{H*}. When replacement cost needed to provide the agent with incentives are too expensive, i.e. $v_p \leq \underline{v}_p$, the principal sticks to her outside option. The agent cannot get less than v_a^{H*} when the principal employs replacement strategies. Only if the principal finds optimal to not participate, in this case he obtains $\underline{v}_a$. Providing that the replacement costs are not too high the principal will employ replacement strategies. In order to keep the principal with incentives the agent has to provide high effort at least when uninformed, in the best scenario he would obtain $v_a^{H,L}$. ■

Proof of Lemma 10. Consider the case where the principal change the monitor as soon as an informative signal is observed, that is $v_{i,k,0}^{H,L}$. Using a similarly reasoning used to prove Lemma 2, we obtain the following recursive payoffs for the case where $i = p$,

$$v_{p,k,0}^{H,L} = (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,k,0}^{H,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,0}^{H,L}\right],$$

and

$$v_{p,0,0}^{H,L} \equiv (1-\delta)\,\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,k,0}^{H,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,0}^{H,L}\right].$$

The first expression $v_{p,k,0}^{H,L}$ is the value of the repeated game that starts with a costly replacement of the monitor. The agent starts providing high effort, then in the following period he is punished with probability p^E. A revealing signal occurs with probability $(1-\beta)\, r$, in which case we the principal replace immediately the monitor and we have a recursive structure. With probability the $1 - p^E - (1-\beta)\, r$ none of these events occur and the monitor is not replaced. The second expression $v_{p,0,0}^{H,L}$ is the value of the relation in this case, which has a recursive structure. Solving recursively these two expression for $v_{p,k,0}^{H,L}$ we obtain expression (5.1) for the case where $n = 0$. Notice that we are solving the recursion assuming that there is a replacement cost in the beginning of the game. This simplifies the recursion, to correct it in the end we add $k(1-\delta)$.

Consider now the case where the principal replace the monitor one period after a revealing signal is observed, in this case we have the following system of equations

$$\begin{aligned}
v_{p,k,1}^{H,L} &= (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta) r v_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1-\beta) r\right) v_{p,0,1}^{H,L}\right], \\
v_{p,0,1}^{H,L} &\equiv (1-\delta)\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta) r v_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1-\beta) r\right) v_{p,0,1}^{H,L}\right], \\
v_{p,0,1}^{\cdot,L} &\equiv (1-\delta)\pi_p^L + \delta\left[q^T \underline{v}_p + \left(1 - q^T\right) v_{p,k,1}^{H,L}\right].
\end{aligned}$$

Notice that now we have a third equation, that is due to the fact that the agent is allowed to enjoy the benefit of learning for one period. After we have solve for $v_{p,k,1}^{H,L}$ and added $k(1-\delta)$ we obtain expression (5.1) for the case where $n = 1$. For general n, in the previous system of equations replace the value $v_{p,k,1}^{H,L}$ for $v_{p,k,n}^{H,L}$, $v_{p,0,1}^{H,L}$ for $v_{p,0,n}^{H,L}$ and $v_{p,0,1}^{\cdot,L}$ for $v_{p,0,n}^{\cdot,L}$ in the two first equations and substitute the third equation by

$$v_{p,0,n}^{\cdot,L} = \left((1-\delta)\pi_p^L + \delta q^T \underline{v}_p\right)\sum_{l=0}^{n-1}\delta^l (1-q^T)^l + \delta^n \left(1-q^T\right)^n v_{p,k,n}^{H,L}. \tag{A6}$$

Solve the system for $v_{p,k,n}^{H,L}$, using the fact that $\sum_{l=0}^{n-1} x^l = (1-x^n)/(1-x)$ and adding the term $k(1-\delta)$, we obtain expression (5.1). A similar reasoning is done when $i = a$ with $k = 0$. ■

Proof of Proposition 11. First we show that $n = 0$ is an optimal choice when $v_{p,k,n}^{H,L} > v_{p,k,n+1}^{H,L}$, i.e. $k < k^{cut}$, and $n \to \infty$ is optimal if $v_{p,k,n}^{H,L} \leq v_{p,k,n+1}^{H,L}$, i.e. $k \geq k^{cut}$. The value k^{cut} is the value of k that solves $v_{p,k,n}^{H,L} = v_{p,k,n+1}^{H,L}$. Using (5.1) for n and $n+1$, setting $\pi_p^L = \underline{v}_p$, after some algebraic manipulation we obtain k^{cut} given in (5.6). Notice that k^{cut} is independent of n. Clearly $k^{cut} > 0$ since both the numerator and denominator are strictly positive. Then $v_{p,k,0}^{H,L}$ is the supremum of $\left\{v_{p,k,0}^{H,L}, v_{p,k,1}^{H,L}, ...\right\}$ when $k < k^{cut}$ and $v_{p,k,\infty}^{H,L}$ is the supremum when $k \geq k^{cut}$.

Lets first look at the case where $\delta \in [\delta^-, \delta^+)$. In order for $n = 0$ to be optimal the principal must obtain a payoff larger than $v_p^{H,L} = v_{p,k,\infty}^{H,L}$. This is not a participation constraint, but a condition for the monitor replacement. Solving $v_{p,k,n}^{H,L} > v_p^{H,L}$ for k, we again obtain (5.6). Then $n = 0$ is optimal if $k \in [0, k^{cut})$, otherwise $n \to \infty$ is optimal, i.e. for $k \geq k^{cut}$. The agent incentives are guaranteed by Proposition 4 and since $[\delta^-, \delta^+) = [\delta_\infty^-, \delta_\infty^+) \subset \left[\delta_0^-, \delta_0^+\right)$ by (5.4) and (5.5).

In order to find the principal optimal strategy for $\delta \in [\delta^*, \delta^-)$ both the principal and the agent must have incentives in participate. Recall that $\pi_p^L = \underline{v}_p$, imposes a lower bound on low effort, see (3.4). The principal participation for this discounting region is guaranteed if $v_{p,k,n}^{H,L} > v_p^{L,L} = \underline{v}_p$. Solving the inequality for k we obtain $k < k^{PCn}$, where k^{PCn} is given by (5.7). Moreover, k^{PCn} is a strictly positive and increasing function of n since $\delta\left(1 - q^T\right) < 1$ and $\pi_p^H > \underline{v}_p$, i.e. $0 < k^{PCn} < k^{PCn+1}$ for all $n \in \mathbb{N}_0$. To show that $k^{PCn} > k^{cut}$, notice that k^{PCn} reach its lowest value when $n = 0$. It is then the hardest to satisfy scenario, but in this case $1 - \delta\left(1 - p^T - \beta r\right) > 0$, implying that $k^{cut} < k^{PCn}$ for all $n \in \mathbb{N}_0$.

Recall that when $k < k^{cut}$, the principal participation is guaranteed for all n and $n = 0$ is the optimal replacement choice, this is also true for $\delta \in [\delta^*, \delta^-)$. The ordering of sequence (5.5) guarantees that the agent incentives for e^H in any uninformed state are satisfied. Consider now that $k \geq k^{cut}$, in this case the principal's optimal choice has to guarantee that the agent incentives to provide at least high effort in the uninformed state are satisfied. When $\delta \in \left[\delta_n^-, \delta_{n+1}^-\right) \subset [\delta^*, \delta^-)$ the principal optimal choice must be n. A choice of $n - 1$ is payoff inferior for the principal since $k \geq k^{cut}$, a choice $n + 1$ does not provide the agent with incentives. Finally the principal participation has also to be guaranteed, i.e. $k < k^{PCn}$, otherwise $v_{p,k,n}^{H,L} \leq \underline{v}_p$ and no trade is optimal. ■

Proof of Corollary 12. Since in Proposition 11 the principal participation constraints are on same time replacement constraints any finite optimal choice that allows for trade must allow a payoff improvement, the proof of this result can be found on the proof of Proposition 11. See also the proof of Corollary 9. ■

Proof of Lemma 13. The proof is identical to the proofs of Lemmas 2 and 10. We look first for $\widetilde{v}_{p,k,0}^{H,L}$, i.e. the case where the agent is substituted in every period. Let $i = p$, we obtain the following recursive payoff

$$\widetilde{v}_{p,k,0}^{H,L} = (1 - \delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1 - \beta)\, r \widetilde{v}_{p,k,0}^{H,L} + \left(1 - p^E - (1 - \beta)\, r\right) \widetilde{v}_{p,k,0}^{H,L}\right],$$

which can be solved for $\widetilde{v}_{p,k,0}^{H,L}$ and finally add $k(1 - \delta)$. Consider now the case where $n = 1$, in this case we obtain the following system of equations

$$\begin{aligned}
\widetilde{v}_{p,k,1}^{H,L} &= (1 - \delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1 - \beta)\, r \widetilde{v}_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1 - \beta)\, r\right) \widetilde{v}_{p,0,1}^{H,\cdot}\right], \\
\widetilde{v}_{p,0,1}^{\cdot,L} &= (1 - \delta)\, \pi_p^L + \delta q^T \underline{v}_p + \delta\left(1 - q^T\right) \widetilde{v}_{p,k,1}^{H,L}, \\
\widetilde{v}_{p,0,1}^{H,\cdot} &= (1 - \delta)\, \pi_p^H + \delta p^E \underline{v}_p + \delta\left(1 - p^E\right) \widetilde{v}_{p,k,1}^{H,L},
\end{aligned}$$

which is solved for $\widetilde{v}_{p,k,1}^{H,L}$ and finally adding $k(1-\delta)$. In this case the agent is substituted after two periods. The first equation is the value of the repeated game that starts with a costly selection of a monitor. The agent starts providing high effort. In the end of the first period he is punished with probability p^E, otherwise either a revealing signal occurs with probability $(1-\beta)r$ or he stays uninformed with probability $1-p^E-(1-\beta)r$. In the former case he enjoys the benefits from learning during on period after which he is replaced in case that is not punished. The repeated game restarts again with a costly replacement of the monitor. This is the second equation which is the value of the relation in the case where the agent gets informed and supplies with low effort for one period. The third expression has a similar interpretation as the second one, but for the case where the game remains in the uninformed state, the agent start by supplying high effort. If no punishment occurs, with probability $1-p^E$ the repeated game is restarted again.

Consider now the general case n, in this case we have $2n+1$ equations,

$$\begin{aligned}
\widetilde{v}_{p,k,n}^{H,L} &= (1-\delta)\left(\pi_p^H-k\right)+\delta\left[p^E\underline{v}_p+(1-\beta)r\widetilde{v}_{p,0,n}^{\cdot,L}+\left(1-p^E-(1-\beta)r\right)\widetilde{v}_{p,0,n}^{H,\cdot}\right],\\
\widetilde{v}_{p,0,n}^{\cdot,L} &= \left((1-\delta)\pi_p^L+\delta q^T\underline{v}_p\right)\sum_{k=0}^{n-1}\delta^k\left(1-q^T\right)^k+\delta^n\left(1-q^T\right)^n\widetilde{v}_{p,k,n}^{H,L},\\
\widetilde{v}_{p,0,n}^{H,\cdot} &= (1-\delta)\pi_p^H+\delta\left[p^E\underline{v}_p+(1-\beta)r\widetilde{v}_{p,0,n-1}^{\cdot,L}+\left(1-p^E-(1-\beta)r\right)\widetilde{v}_{p,0,n-1}^{H,\cdot}\right],\\
&\dots\\
\widetilde{v}_{p,0,2}^{\cdot,L} &= \widetilde{v}_p^{0,2\prime}=\left((1-\delta)\pi_p^L+\delta q^T\underline{v}_p\right)\left(1+\delta\left(1-q^T\right)\right)+\delta^2\left(1-q^T\right)^2\widetilde{v}_{p,k,n}^{H,L},\\
\widetilde{v}_{p,0,2}^{H,\cdot} &= \widetilde{v}_p^{0,2}=(1-\delta)\pi_p^H+\delta\left[p^E\underline{v}_p+(1-\beta)r\widetilde{v}_{p,0,1}^{\cdot,L}+\left(1-p^E-(1-\beta)r\right)\widetilde{v}_{p,0,1}^{H,\cdot}\right],\\
\widetilde{v}_{p,0,1}^{\cdot,L} &= (1-\delta)\pi_p^L+\delta q^T\underline{v}_p+\delta\left(1-q^T\right)\widetilde{v}_{p,k,n}^{H,L},\\
\widetilde{v}_{p,0,1}^{H,\cdot} &== (1-\delta)\pi_p^H+\delta p^E\underline{v}_p+\delta\left(1-p^E\right)\widetilde{v}_{p,k,n}^{H,L},
\end{aligned}$$

which can be solved recursively for $\widetilde{v}_{p,k,n}^{H,L}$ to obtain, after adding $k(1-\delta)$, the expression

$$\widetilde{v}_{p,k,n}^{H,L}=\frac{\left((1-\delta)\pi_p^H+\delta p^E\underline{v}_p\right)\sum_{k=0}^{n}\delta^kz^k+\left((1-\delta)\pi_p^L+\delta q^T\underline{v}_p\right)\sum_{r=1}^{n}yz^{r-1}\sum_{k=0}^{n-r}\delta^{k+r}x^k-k(1-\delta)}{1-\delta^{n+1}\left(z^{n+1}+\sum_{k=0}^{n}x^{n-k}yz^k\right)}+k(1-\delta), \tag{A7}$$

which equals expression (6.1) after all the summations have been solved. Similar reasoning is employed when $i=a$ with $k=0$. ■

When required to shorten in notation we use the following definitions; $x\equiv 1-q^T$, $y\equiv(1-\beta)r$ and $z\equiv 1-p^E-(1-\beta)r=1-p^T-r$.

Proof of Proposition 14. Suppose that $\widetilde{v}^{H,L}_{p,k,n}$ has a unique global maximum in $\mathbb{N}_0$. Then $\widetilde{v}^{H,L}_{p,k,n}$ given by (6.1) reach its maximum value at n if $\widetilde{v}^{H,L}_{p,k,n-1} \leq \widetilde{v}^{H,L}_{p,k,n}$ and $\widetilde{v}^{H,L}_{p,k,n} \geq \widetilde{v}^{H,L}_{p,k,n+1}$. Let $k^{n,n+1}$ be the value k that solves $\widetilde{v}^{H,L}_{p,k,n} = \widetilde{v}^{H,L}_{p,k,n+1}$ for $n = 0, 1, 2....$ Consider $\pi^L_p = \underline{v}_p$, after some algebraic manipulation we obtain

$$k^{n,n+1} = \frac{\left(\pi^H_p - \underline{v}_p\right) y \left[z^{n+1}(1-\delta z) - x^{n+1}(1-\delta x) + z^{n+1}x^{n+1}\delta^{n+1}(z-x)\delta\right]}{(1-\delta z)\left[z^{n+1}(1-\delta z)(z+y-x) - x^{n+1}y(1-\delta x)\right]},$$

Notice that $k^{-1,0} = 0$. The expression for $k^{n,n+1}$ is increasing in n if

$$y \leq \frac{(x-z)\left(1-\delta^{n+1}z^{n+1}\right)}{\delta^{n+1}\left(x^{n+1} - z^{n+1}\right)}. \tag{A8}$$

Independently of the relation between x and z, the right-hand side is monotonically increasing in n, converging to ∞ when $n \to \infty$ and to $(1-\delta z)/\delta$ when $n \to 0$. The latter it is the lowest value and for that reason the hardest to satisfy. Then substituting $y \equiv (1-\beta) r$ and $z \equiv 1 - p^T - r$ and solving for δ we obtain $\delta \leq 1/\left(1 - p^T - \beta r\right)$, which is always satisfied because its larger than 1. Then the optimal replacement choice n has to satisfy $\widetilde{v}^{H,L}_{p,k,n-1} \leq \widetilde{v}^{H,L}_{p,k,n}$ and $\widetilde{v}^{H,L}_{p,k,n} \geq \widetilde{v}^{H,L}_{p,k,n+1}$, or equivalently $k^{n-1,n} \leq k \leq k^{n,n+1}$.

Now, let $n \to \infty$ in the expression for $k^{n,n+1}$. If $x \geq z$, i.e. $r \geq q^T - p^T$, then we obtain

$$k^{\infty} = \frac{\pi^H_p - \underline{v}_p}{1-\delta z} = \frac{\pi^H_p - \underline{v}_p}{1-\delta\left(1-p^T-r\right)},$$

which is expression (5.6). When $x < z$, i.e. $r < q^T - p^T$, we obtain

$$k^{\infty} = \frac{\pi^H_p - \underline{v}_p}{1-\delta z}\frac{y}{z+y-x} = \frac{\pi^H_p - \underline{v}_p}{1-\delta\left(1-p^T-r\right)}\frac{(1-\beta) r}{q^T - p^T - \beta r},$$

where the second ratio on the right-hand side is smaller than one. Also, since $p^T + r = p^S < 1$, we have $k^{\infty} > 0$ always. In resume for any $\delta \in [\delta^*, \delta^+)$ and $k \in [0, k^{\infty})$, when unrestricted the principal prefers to replace the monitor after a finite number of periods, while if $k \geq k^{\infty}$ he must choose $n \to \infty$.

We need now to verify when the principal prefers to employ replacement strategies. For this $\widetilde{v}^{H,L}_{p,k,n} \geq \underline{v}_p$ when $\delta \in [\delta^*, \delta^-)$ and $\widetilde{v}^{H,L}_{p,k,n} \geq v^{H,L}_p$ when $\delta \in [\delta^-, \delta^+)$. Lets look first to the former case. Solve $\widetilde{v}^{H,L}_{p,k,n} = \underline{v}_p$ for k to obtain the participation condition in terms of replacement costs, denote it by k^{PCn} and is given by

$$k^{PCn} = \frac{\left(\pi^H_p - \underline{v}_p\right)(x-z)\left(1-\delta^{n+1}z^{n+1}\right)}{\left[\left(x^{n+1} - z^{n+1}\right)y + z^{n+1}(x-z)\right](1-\delta z)}.$$

Independently of the relation between x and z, the function k^{PCn} is strictly increases in n. Formally it is enough to show that $k^{PCn} < k^{PCn+1}$, that is

$$y > -\frac{(x-z)\, z^{n+1}\,(1-\delta z)}{(x^{n+1}-z^{n+1}) - \delta\,(x^{n+2}-z^{n+2}) + \delta\,(x-z)\,\delta^{n+1}x^{n+1}z^{n+1}}. \tag{A9}$$

The right-hand side is always negative, since $(x^{n+1}-z^{n+1}) > (x^{n+2}-z^{n+2})$ when $x > z$ and the reverse when $z > x$. The numerator and denominator always have the same signal. Then $k^{PC\infty}$ is the largest cost and takes the value ∞ when $n \to \infty$.

Consider now the case $\delta \in [\delta^*, \delta^-)$, and solve $\widetilde{v}^{H,L}_{p,k,n} = v^{H,L}_p$ for k to obtain the replacement condition

$$k^{RCn} = \frac{\left(\pi^H_p - \underline{v}_p\right) y\,(x^{n+1}-z^{n+1})}{\left[(x^{n+1}-z^{n+1})\, y + (x-z)\, z^{n+1}\right](1-\delta z)}.$$

Taking the limit $n \to \infty$ we found that $k^{RCn} \to k^{\infty}$ whether $x \geq z$ or $x < z$. Moreover k^{RCn} is strictly increasing with n. In order for $k^{RCn} > k^{n,n+1}$, the same condition (A8) must be strictly satisfied, which we have shown above to be always the case. Moreover, since $v^{H,L}_p > \underline{v}_p$ then $k^{PCn} > k^{RCn}$, consequently we have $k^{PCn} > k^{n,n+1}$ for all $n \in \mathbb{N}_0$.

Now we consider the agent incentives to place high effort in the uniformed state, given a replacement choice. (i) For $\delta \in [\delta^-, \delta^+)$, since for all n, $\delta^-_n \leq \delta^-$ the agent incentives to provide high effort while uninformed are always satisfied. The principal choose some finite n if $k \in [0, k^{\infty})$ and $n \to \infty$ otherwise. However if $k \geq k^{RCn}$, she must choose the smallest τ larger than n such that the inequality is reversed, i.e. $k < k^{RC\tau}$. The monitor has to be replaced less often in order to increase the replacement condition bound $k^{RC\tau}$. However, if such it is not possible, i.e. $k \geq k^{RC\infty} = k^{\infty}$, the principal should choose $n \to \infty$. Since $k^{PC\infty} = \infty$, participation is always guaranteed.

Now suppose that $\delta \in \left[\delta^-_m, \delta^-_{m+1}\right) \subseteq [\delta^*, \delta^-)$. Then we have two possibilities. (ii) The unconstrained optimal $n \leq m$, in this case the choice n is optimal if $k < k^{PCn} < k^{\infty}$, since the principal and agent incentives are met. However, if $k^{PCn} \leq k < k^{\infty}$, the principal has to choose the lowest $\tau \in \{n+1, ..., m\}$ such that $k < k^{PC\tau}$. The principal cannot choose a replacement frequency above m because in this case the agent would not have incentives to provide effort in the uninformed state. Similarly, when $k > k^{\infty}$, she must choose m if $k < k^{PCm}$. It might happen that $k \geq k^{PCm}$ then the principal should not participate because the replacement costs are too high and for any replacement frequency above m the agent has no incentives. (iii) Second, if the unconstrained optimal $n > m$, to keep the agent with incentives, it is optimal to choose m, but such is only the case if $k < k^{PCm}$. Otherwise, trade is not possible, because it is too costly for the principal to provide the agent with incentives. ■

Proof of Proposition 15. The argument is similar to the one used in the Proof of Corollary 12. In Proposition 14 the principal participation constraints are on same time replacement constraints. Then any finite optimal choice that allows for trade must allow a payoff improvement, the proof of this result can be found along the proof of Proposition 14. ■