Executive Compensation, Monitoring and Collusion in Boards of Directors*

Sylvain Bourjade

Université de Toulouse, TBS

s.bourjade@tbs-education.fr

Laurent Germain

Université de Toulouse, TBS

l.germain@tbs-education.fr

July 2015

Abstract

This paper studies the intensity with which a board of directors should monitor a CEO in order to maximize firm's performance in a framework with asymmetric information, collusion and uncertainty about the optimal projects for the firm. We derive the optimal incentive compensation contract of the CEO and characterize conditions under which these incentive compensation contracts are sufficient in order to induce the CEO not to mis-behave irrespective of the intensity of monitoring. In addition, we provide empirical predictions about the relationship between the intensity of monitoring, incentive compensation and firms' characteristics.

Keywords: Incentive Compensation, Collusion, Corporate Governance, Asymmetric Information, Uncertainty.

JEL Classification: D81, D82, G34.

*We thank Gilles Chemla, Francesca Cornelli, Mara Faccio, Daniel Ferreira, Urs Peyer, Charu Raheja, Silvia Rossetto, David Thesmar and participants to the Toulouse Business School Corporate Governance workshop, the EEA-ESEM 2011 conference in Oslo, the AFSE 2012 conference in Paris, the EARIE 2012 conference in Rome and to the CASS Business School, EM Lyon and Rouen Business School seminars for helpful comments. All remaining errors are ours.

1 Introduction

Collusion between members of boards of directors and top executives represents a major problem for the corporate governance of firms. Indeed, in most of the recent corporate governance scandals, a significant proportion of board members[1] proved to stay always loyal to their CEO. An example of such a behavior is highlighted in The Boston Globe (January 6, 2007): "Despite his failure to increase the value of Home Depot's stock, chief executive officer Robert Nardelli left the company this week with a $210 million farewell package, the result of an agreement he negotiated with the board of directors in 2000. Across America, a culture of collusion between board members and prospective CEOs inflates executive pay and needs to be checked by greater shareholder involvement." Those "collusive" directors (some of them referred to as the "Bernie's Boys" for Worldcom) vote in favor of the CEO's propositions and allow her to receive, for example, generous bonuses, severance packages and golden retirement pensions. In many of these cases of "bad governance," one of the main issues is either an explicit or an implicit collusion between the directors of the board and the CEO.

This paper investigates how shareholders can design incentive compensation contracts and choose the intensity with which the board of directors should monitor the top executives in order to maximize firm's performance. The Sabarnes-Oxley Act, the NYSE and the NASDAQ regulations in the US require that independent directors, who are supposed to supervise more efficiently firms' top executives, play a more important role in the boards of directors. To understand the efficiency of such requirements, we examine the optimal intensity with which the board of directors monitors top executives from a shareholders perspective in the presence of collusion between the top executives and the board of directors. In addition, we analyze how incentive compensation contracts can help shareholders to solve these agency problems.

[1] Vinci Group, Worldcom or Home Depot, among others.

In our setting, the top executive (CEO hereafter) has to choose between two projects where the first (second) project is better for shareholders when the economic context is bad (good). The CEO's ability to undertake projects (High or Low) is unknown to the shareholders. The project's type and the CEO's ability are her private information. Hence, this is a two-dimensional adverse selection model. To overcome the technical difficulties in solving such a model, we propose to rewrite it as a function of only one parameter that completely captures the dependence of the CEO's program on the two adverse selection dimensions.

Selecting one of the projects yields a private benefit to the CEO. This private benefit can be thought of as her utility from deriving various advantages such as perks, or building empires. To limit the CEO's discretion, shareholders have the opportunity to choose the intensity with which the CEO will be monitored by the board[2].

In addition, we allow for the possibility of collusion between the board and the CEO. The CEO can propose a monetary or a non monetary transfer (such as future salary increases, perks, insurance to stay in the board,...) to some of the directors to induce them not to reveal to shareholders that she has made a bad decision for the firm. Consequently, the collection of information from the CEO by shareholders is more difficult and more costly because collusion reduces the toughness of monitoring by directors.

Monitoring of the CEO by the board of directors influences the CEO's behavior.[3] The lower the intensity of monitoring, the more likely is the board to engage in collusion with the CEO, but also the more precise the board's information is. These two effects are due, for example, to his relationships with the CEO (degree of confidence, local networking, bargaining power) and his executive role in the firm for instance.[4]

Within this framework, we derive the optimal compensation contract of the CEO that

[2]See, Faleye, Hoitash and Hoitash (2011) or Ferreira, Ferreira and Raposo (2011).

[3]In our model, even though the Board may report information about the type of the project that has been advised by the CEO, we focus on his monitoring role.

[4]See, Ferreira, Ferreira and Raposo (2011), for instance.

consists of a fixed part and a variable part. Our results are as follows. In the benchmark the case of no board of directors (equivalent to no CEO's monitoring by the directors), we show that the variable part of the CEO's wage is higher for a high ability CEO than for a low ability CEO. Second, in the case where we allow shareholders to recruit a board of directors in order to monitor the CEO, assuming no collusion, the board behaves as a perfectly honest board. The contract takes the same form as the one with no monitoring, that is, no informational rent for a low ability CEO and a positive informational rent for a high ability CEO. Those informational rents correspond to the surplus a CEO can extract from the shareholders thanks to her informational advantage. However, the informational rents are lower in this case than when there is no monitoring from the board. Consequently, it is less costly for shareholders to obtain information from the CEO when the board monitors her. This enables us to characterize a threshold wage such that if the board's wage is lower than this threshold, recruiting a board of directors to monitor the CEO is always beneficial for the shareholders.

Third, we allow for collusion between the board and the CEO and show that the optimal contract is collusion proof: it is optimal for the shareholders to offer a contract preventing collusion to emerge. The optimal contract is designed such that shareholders have to concede to the CEO the same informational rents as in the presence of a perfectly honest board. However, they also have to ensure that the coalition Board-CEO does not collude which is costly in terms of informational rents. In addition, we prove that there exists a degree of independence of the board above which it is not profitable for the coalition Board-CEO to engage in collusion. In this case, shareholders do not have to care about preventing collusion when designing the optimal contract. The board behaves as a perfectly honest board.

To our knowledge, our paper is the first theoretical model to consider the explicit collusion between the board of directors and the CEO. However, collusion has received a large attention in the Mechanism Design literature. The seminal paper of Tirole (1986) studies a three-tier organization with a principal, a supervisor and an agent in a moral hazard framework.[5] In Tirole (1986), the agent and the supervisor can collude. Tirole (1986) derives the optimal collusion-proof contract. We also address this problem, but in an adverse selection framework, and show that the optimal contract is also collusion-proof. Faure-Grimaud, Laffont and Martimort (2003) also study, in an adverse selection model, the optimal design of organization and the value of delegation when the supervisor and the agent can collude against the principal. Our paper differs from theirs in at least three respects. First, we allow the CEO (the agent in their framework) to choose between different investment projects. Second, in our model, collusion is impacted by the shareholders' (the principal) choice of the intensity of monitoring (proportion of independent directors). Third, we study the optimal composition of boards of directors (supervisor) in the presence of collusion.

Moreover, we derive the optimal intensity with which the board should monitor the CEO. This intensity corresponds to the most efficient incentive compensation contract that allows shareholders to pay the lowest informational rents to the CEO. Contrary to the usual idea that an optimal board should be strongly monitored, we find that shareholders may prefer to select a low intensity of monitoring as incentive compensation contracts allow them to collect truthful information from the CEO without monitoring her. Indeed, when designing the optimal compensation contract, shareholders face a trade-off between the information that they can extract from the board and the costs from both extracting it and avoiding collusion. We characterize conditions under which incentive compensation contracts are sufficient to induce the CEO not to misbehave irrespective of the intensity of monitoring. These conditions are as follows: the benefits of choosing the best project for shareholders should be relatively low and the intensity of monitoring necessary to have a perfectly honest board should be high enough. However, when these benefits are higher or when the intensity of monitoring necessary to have a perfectly honest board is low enough, it is optimal to select

[5] Other related papers are Kofman and Lawarrée (1993) and Khalil and Lawarrée (1995) in an audit framework.

a high intensity of monitoring. In this case, the shareholders should not care about collusion because collusion is not profitable for such boards.

When the benefits of choosing the best project are high, it is of great importance to monitor the CEO and to induce her to choose the best project for the shareholders. In this case, shareholders should select a high intensity of monitoring. The problem is less acute when these benefits decrease. This is consistent with empirical results. For example, Demsetz and Lehn (1985) and Ferreira, Ferreira and Raposo (2011) show that more monitoring is needed in more complex firms. Finally, we provide empirical predictions about the relationship between the intensity of monitoring, incentive compensation and firms' characteristics.

There is a large literature in corporate governance about the composition of the boards of directors (Boone, Field, Karpoff and Raheja, 2007, Dahya and McConnell, 2007, Harris and Raviv, 2006, Linck, Netter and Yang, 2008, Raheja, 2005), the relationship between the structure of boards of directors and the CEO compensation (Chhaochharia and Grinstein, 2009) as well as the monitoring role (Hermalin and Weisbach, 1998, or Cornelli, Kominek and Ljungqvist, 2010) and the advisory role of the boards of directors (Adams and Ferreira, 2007). Nevertheless, the problem of potential collusion between the CEO and the board has received little attention.[6]

Our paper is related to Adams and Ferreira (2007). In their model, there is a continuum of projects but the projects do not differ in their probability of success. The CEO is reluctant to transmit information to the board of directors because of the board's monitoring role. The composition of the board of directors influences the behavior of the CEO as the more independent is the board of directors, the more the CEO is monitored and the less the CEO is inclined to share information with the board. We find similar results: it is optimal for the shareholders to choose a board who is reluctant to monitor the CEO. However, the forces driving our result are different from theirs. First, Adams and Ferreira model

[6] For reviews of the Corporate Governance literature, see, Adams, Hermalin and Weisbach (2010), Bebchuk and Weisbach (2010),or Tirole (2001).

information transmission between the board and the CEO as a cheap-talk game while we use a mechanism design framework that allows us to derive the optimal incentive compensation contracts. Second, in their article, when the board's independence level is low, there is a low probability for the CEO to lose control. This makes revelation of information less costly for him and implies that choosing such a board may be optimal for the shareholders. In our paper, shareholders select a board with a low level of CEO's monitoring because incentive compensation contracts, even without any monitoring from the board of directors, are sufficient to collect the optimal amount of information to make an efficient investment decision. Finally, they do not explicitly model collusion between the CEO and the board members.

Another related paper is Hermalin and Weisbach (1998) who analyze the role of independent directors in boards and the intensity of monitoring. They show that a bad CEO is more likely to be replaced when the board is independent. Therefore, independent directors are means for controlling the performance of the firm and a threat for bad CEOs. We also examine the optimal intensity of monitoring. However, we focus on board's monitoring of projects's probability of success and not on board's monitoring of the CEO's ability. Moreover, we characterize the optimal incentive compensation for the CEO and the role of collusion on the board's monitoring role. Hermalin (2005) studies the decision of hiring an internal versus an external CEO. The model he develops determines whether it is optimal to keep an existing CEO or to replace him at a certain cost, however, less is known about the external CEO. Even if we do not address directly the question of the replacement of the CEO, monitoring of the CEO by the directors can entail a high fine for him which may be interpreted as his dismissal.

Raheja (2006) studies the question of the optimal composition and the ideal size of boards of directors. In the model, the optimal board structure is determined by the trade-off between insiders' incentives to reveal their private information and the outsiders' costs to verify projects. We also derive the optimal intensity of monitoring of the CEO, which may be interpreted by the board's composition in Raheja's model, taking into account the collusive behavior of the CEO and the directors.

Kumar and Sivaramakrishnan (2008) also examine the effect of incentive compensation contracts on the board's monitoring intensity and firm performance. However, they do not consider collusion between the board and the CEO and its resulting effects on the optimal monitoring intensity and incentive compensation contracts.

The article is organized as follows. Section 2 describes the model. Section 3 analyzes the benchmark case of no monitoring of the CEO by the board while section 4 introduces board monitoring. Section 5 studies the impact of collusion on our results. The optimal intensity with which the board monitors the CEO is characterized in section 6. Finally, section 7 concludes.

2 The Model

2.1 The CEO and the Projects of the Company

A firm can undertake a project which yields an uncertain payoff. The firm is run for the shareholders by a CEO, i.e. the CEO's task is to select the project that will be undertaken by the firm.

The CEO's ability to succeed in the projects may be either low, $\beta = \beta_L$, with probability (γ) or high, $\beta = \beta_H$ with probability $(1 - \gamma)$. As $\beta_L$corresponds to a low CEO's ability and β_H to a high ability, we have $\beta_H \geq \beta_L$.

We assume that the firm can undertake two projects where the first (second) project is better for shareholders when the economic context is bad (good). The implementation of those projects initially require a fixed investment I by the firm's shareholders. The characteristics of those projects are the following:

(i) Project 1 either succeeds, that is, yields verifiable income $R > 0$ or fails, that is, yields no income. The probability of success is denoted by (q_1). Moreover, this project may have a low probability of success, that is, $q_1 = p_L\beta_i$ with probability (ν) or may have a high probability of success $q_1 = p_H\beta_i$ with probability $(1-\nu)$ where $\beta_i \in \{\beta_H; \beta_L\}$ is the CEO's ability to succeed in the projects.

(ii) In the same way, Project 2 either succeeds, that is, yields verifiable income $R > 0$ or fails, that is, yields no income. The probability of success is denoted by (q_2). Moreover, this project may have a low probability of success, that is, $q_2 = (p_L - \varepsilon)\beta_i$ with probability (ν) or may have a high probability of success $q_2 = (p_H + \varepsilon)\beta_i$ with probability $(1-\nu)$ where $\beta_i \in \{\beta_H; \beta_L\}$ is the CEO's ability to succeed in the projects.

The success and the failure of both projects are assumed to be perfectly correlated i.e. (ν) represents the probability that the economic context is bad for the type of projects considered by the firm while (ε) represents the increase in the probability of success when the best project is selected: Project 1 (Project 2) when the economic context is bad (good). ε can therefore be interpreted as the value of choosing the best project for shareholders in terms of probability of success.

As the Net Present Value of Project 2 has to be at least higher than the NPV of the other project, we have:

$$(\nu(p_L - \varepsilon)\beta_i + (1-\nu)(p_H + \varepsilon)\beta_i) R - I \geq (\nu p_L\beta_i + (1-\nu) p_H\beta_i) R - I.$$

This is equivalent to:

$$\nu \leq \frac{1}{2}.$$

The CEO perfectly knows both her ability's type and the probability of success of the projects. However, shareholders only know their prior probability distributions.

The CEO may therefore send signals $\sigma_{i,j}$, with $i, j \in \{L, H\}$, to shareholders about her type and the project she advises to select[7].

$$\begin{cases} \sigma_{L,L} = (\beta = \beta_L, \text{Project 1}) & \Rightarrow q_2^L = (p_L - \varepsilon)\,\beta_L \text{ and } q_1^L = p_L \beta_L \\ \sigma_{L,H} = (\beta = \beta_L, \text{Project 2}) & \Rightarrow q_2^H = (p_H + \varepsilon)\,\beta_L \text{ and } q_1^H = p_H \beta_L \\ \sigma_{H,L} = (\beta = \beta_H, \text{Project 1}) & \Rightarrow q_2^L = (p_L - \varepsilon)\,\beta_H \text{ and } q_1^L = p_L \beta_H \\ \sigma_{H,H} = (\beta = \beta_H, \text{Project 2}) & \Rightarrow q_2^H = (p_H + \varepsilon)\,\beta_H \text{ and } q_1^H = p_H \beta_H \end{cases}$$

The CEO's compensation is composed by a fixed part $\alpha_{i,j}$ and a variable part $\mu_{i,j}\pi$ that depends on the profits from the project (π) where $i \in \{L, H\}$ corresponds to the CEO's signal about her ability (called hereafter the CEO's type) and $j \in \{L, H\}$ corresponds to the CEO's signal about the probability of success of the project (called hereafter the state of nature). We assume that CEO's compensation is designed by the firm's shareholders or equivalently by a compensation committee whose members' objectives are in line with shareholders' ones[8]. This assumption can be justified by the existence of "say on pay" rules, as in the US and in Europe, that allow shareholders to actively vote on how much top executives should be compensated.

When Project 2 is selected, the CEO receives a private benefit B which represents her private compensation for inducing the shareholders to choose her preferred project.

The CEO's reservation wage is w.

[7] The assumption that the CEO has private information both on the state and the ability of the CEO is necessary for our results to hold. As the CEO is risk neutral and contracting takes place ex-ante, the optimal incentive compensation contract with one dimensional private information would implement the first best outcome, see, Laffont and Martimort (2002).

[8] Bizjack, Lemmon, and Naveen (2008) document that most firms in the US use a compensation committee that relies on recommendations from outside consultants, peer groups and competitive benchmarking in order to structure the CEO's compensation. They show that total compensation is usually anchored to the peer group.

2.2 The Board of Directors

Shareholders also have the opportunity to hire a board. Even though the board may report information about the type of the project that has been advised by the CEO, the main role of the board is to monitor the information communicated by the CEO.

The monitoring of the CEO by the board of directors is endogenous, in the sense that shareholders choose it. The lower is this intensity of monitoring, the more precise the board's information is, but also the more likely the board is to engage in collusion with the CEO. A low level of monitoring by the board, because of close relationships with the CEO, as well as a high degree of confidence between both parties, their repeated interaction, or local networking for instance, induces the CEO to share more information about the projects with the board[9] but also makes collusion more easily enforceable. In practice, increasing the intensity with which the board monitors the CEO would correspond to an increase in the number of independent directors, a limited number of interlocked directors and mandates held by each director, a separation of the role of Chairman and CEO, an increase in the power and in the independence of the main committees, or an improvement of the internal audit process.

We model monitoring of the CEO by the board of directors by a variable $\tau \in [\tau_{\min}, +\infty]$, with $\tau_{\min} \geq 1$, that also represents the transaction costs of collusion. As in Tirole (1992), we assume that there are transaction costs associated with collusion. Consequently, when the board gets a collusive side payment of x from the top executives, this costs more than x to the Top Executives. The transaction costs capture (i) how collusion is difficult to sustain (for instance the costs of organizing collusion or the exposition to legal sanctions), or (ii) how collusion is accepted in this society (in terms of damages to directors' reputation if they collude), or (iii) the psychological costs of colluding (that is, the inherent aversion of more independent boards to collusion) or (iv) the fact that non-monetary transfers may not have

[9] See, Adams and Ferreira (2007).

the same value as monetary ones (a $10,000 trip paid by top executives to a member of the board may only have a $7,000 value for the latter).

When the intensity of monitoring, τ, increases, the amount of information held by a board decreases while his willingness to engage in collusion decreases[10].

Let $\xi(\tau) = \frac{1}{\tau}$ be the probability that a board with an intensity of monitoring τ collects the true information about the state of nature and therefore finds that the CEO has sent the wrong signal if it is really the case. We assume that, when τ increases, board monitoring increases and board members are less prone to collusion. However, as they have less information about the firm, their probability of knowing the truth is lower. We also assume that the CEO incurs a fine F when the board reveals to the shareholders that she has announced that the project has a high probability of success while it is a project with a low probability of success, that is, the case in which she gets the bonus B.

Our model could also be seen as a model with heterogeneity among members of the board in which shareholders can choose the proportion of independent directors and executive directors. Executive directors would have access to a more accurate information about the projects' type than independent ones. However, they are also less likely to monitor the CEO and more prone to engage in collusion with her because of close relationships. Assuming that independent directors always reveal the truth to shareholders and that executive directors are always induced to collude, then, $\xi(\tau)$, the probability that the board finds that the CEO has sent the wrong signal if it is really the case, represents the proportion of independent directors while $(1 - \xi(\tau))$ represents the proportion of executive directors. Hence, we can interpret the results of the model in terms of the optimal proportion of independent and executives directors in boards.

We are interested in determining the value of the intensity of monitoring τ such that the board is completely honest and never accepts to engage in collusion with the CEO (this

[10] τ can also be interpreted as the shareholders' willingness to increase the degree of toughness and also the enforceability of the Corporate Governance regulations and the laws against collusion. The tougher the laws, the more difficult it is for the coalition Board-CEO to engage in collusion.

however means that he has a less precise information about the state of nature). These boards would therefore perfectly represent shareholders.

When they engage in collusion, we assume that the CEO shares the collusive profits with the board and as in Tirole (1992) that the board has all the bargaining power.

As it is usually the case in practice, the board's wage is the total amount of the directors' fees which is constant and equals to w_0. As in Hermalin and Weisbach (1998), Hermalin (2005) or Adams and Ferreira (2007), we assume that the objectives of board members can be aggregated. This implies that the board behaves as if he were a single agent.

2.3 Multidimensional Screening Model

This model is a multidimensional screening model. Solving this kind of model is usually very complex (see, Rochet and Chone, 1998). However, the structure of the model allows us to reduce this problem's complexity. As the CEO's program can be specified as a function of only one parameter, $\theta_{i,j}$, that captures the effects of both adverse selection variables, we can rewrite the model as a usual four types unidimensional screening model. In this case, $\theta_{i,j}$ is defined in the following way:

$$\begin{cases} \theta_{L,L} = p_L \beta_L \\ \theta_{L,H} = (p_H + \varepsilon)\, \beta_L \\ \theta_{H,L} = p_L \beta_H \\ \theta_{H,H} = (p_H + \varepsilon)\, \beta_H \end{cases}$$

Moreover, we assume that $(p_L - \varepsilon)\, \beta_H \geq (p_H + \varepsilon)\, \beta_L$, that is, a high ability CEO undertaking a project with a low probability of success is more likely to succeed than a low ability CEO undertaking a project with a high probability of success. This assumption highlights the positive role of the CEO in her management of projects.

Denote the firm's profits $\pi(\theta_{i,j}) = \theta_{i,j} R - I$. The shareholders maximize their expected profits:

$$W = \nu\gamma\left[\left(1-\mu_{L,L}\right)\pi^{1}\left(\theta_{L,L}\right)-\alpha_{L,L}\right]+(1-\nu)\gamma\left[\left(1-\mu_{L,H}\right)\pi^{2}\left(\theta_{L,H}\right)-\alpha_{L,H}\right]$$
$$+\nu(1-\gamma)\left[\left(1-\mu_{H,L}\right)\pi^{1}\left(\theta_{H,L}\right)-\alpha_{H,L}\right]+(1-\nu)(1-\gamma)\left[\left(1-\mu_{H,H}\right)\pi^{2}\left(\theta_{H,H}\right)-\alpha_{H,H}\right]$$

We are now able to derive the optimal incentive compensation contracts for the CEO depending on the board's ability to monitor her and the opportunity for the board/CEO coalition to collude

3 No CEO's Monitoring by the Board

As a benchmark case, we first characterize the optimal incentive compensation contracts for the CEO when the board of directors is not able to monitor the CEO. Shareholders maximize their expected profits under the usual Participation and Incentive constraints. PC_{ij} is the Participation constraint of a CEO with ability $i \in \{L, H\}$ when the state of nature is $j \in \{L, H\}$. The Participation constraints ensure that the CEO will earn at least her reservation wage w. $IC_{ij\to kl}$ is the Incentive constraint of a CEO who reveals that her ability is $k \in \{L, H\}$ and the state $l \in \{L, H\}$ while her true ability is i and the true probability of success of the best project is j. The Incentive constraints ensure that the CEO earns a higher wage revealing the truth than lying to the shareholders. Through this process, shareholders induce the CEO to reveal the truth. Those constraints are stated here:

$$\alpha_{i,j}+\mu_{i,j}\pi\left(\theta_{i,j}\right)\geq w,\ \forall i,j\in\{H,L\} \qquad (PC_{ij})$$

$$\alpha_{i,j}+\mu_{i,j}\pi\left(\theta_{i,j}\right)\geq \alpha_{kj}+\mu_{kj}\pi\left(\theta_{i,j}\right),\ \forall i,j,k\in\{H,L\} \qquad (IC_{ij\to kj})$$

$$\alpha_{i,H} + \mu_{i,H}\pi\left(\theta_{i,H}\right) \geq \alpha_{kL} + \mu_{kL}\left[q_H^1 \beta_k R - I\right], \ \forall i,k \in \{H,L\} \qquad (IC_{iH\to kL})$$

$$\alpha_{i,L} + \mu_{i,L}\pi\left(\theta_{i,L}\right) \geq \alpha_{kH} + \mu_{kH}\left[q_L^2 \beta_k R - I\right] + B, \ \forall i,k \in \{H,L\} \qquad (IC_{iL\to kH})$$

Moreover, the Spence Mirrlees condition has to be satisfied, that is:

$$\mu_{HH} \geq \mu_{HL} \geq \mu_{LH} \geq \mu_{LL,}$$

and by assumption, we know that the following condition is satisfied:

$$\left(p_L - \varepsilon\right)\beta_H - \left(p_H + \varepsilon\right)\beta_L \geq 0 \qquad (1)$$

Then we can characterize the optimal incentive compensation contract when there is no monitoring from the board in the firm's organization. This is stated in the following Proposition:

Proposition 1 *When they do not induce the board of directors to monitor the CEO, shareholders should offer the following incentive compensation contract, U_{ij} with $i,j \in [L,H]$, to a CEO:*

$$\begin{aligned}
U_{LL} &= w \\
U_{LH} &= w \\
U_{HL} &= w + \frac{B\left(p_L - \varepsilon\right)\Delta\beta}{\beta_L\left[\Delta p + 2\varepsilon\right]} \\
U_{HH} &= \begin{cases} w + \frac{B(p_L-\varepsilon)p_H(\Delta\beta)^2}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H - p_H\beta_L]} \ \mathit{if}\ \varepsilon \leq \varepsilon_{nb} = \frac{\beta_L\Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} \\ w + \frac{B\Delta\beta(p_H+\varepsilon)}{\beta_L[\Delta p+2\varepsilon]} \ \mathit{if}\ \varepsilon \geq \varepsilon_{nb} \end{cases}
\end{aligned}$$

Moreover, the shareholders' expected profits are

$$W_{NM}=\begin{cases} E(\pi)-w-\frac{(1-\gamma)B(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p+2\varepsilon)}\left[\frac{-\nu\beta_H\Delta p+p_H\Delta\beta}{p_L\beta_H-p_H\beta_L}\right] & \text{if } \varepsilon\leq\varepsilon_{nb} \\ E(\pi)-w-(1-\gamma)\left(p_H+\varepsilon-\nu\Delta p-2\nu\varepsilon\right)\frac{B\Delta\beta}{\beta_L(\Delta p+2\varepsilon)} & \text{if } \varepsilon\geq\varepsilon_{nb} \end{cases}$$

A low ability CEO does not receive any incentive compensation whatever the type of project she advises to select. However, when her signal pushes shareholders to select the project with the highest volatility (Project 2), she receives a variable wage while she only gets a fixed wage when shareholders are induced to select Project 1.

A high ability CEO receives an informational rent which is higher when her signal induces shareholders to select Project 2 than when shareholders are induced to select Project 1. Moreover, the variable part of her wage is higher when Project 2 is selected than when it is Project 1. But, in any case, the variable part of a high ability CEO is higher that the one of a low ability CEO.

Finally, this highest informational rent when Project 2 is selected takes different forms depending on ε, that is, the value of choosing the best project for shareholders in terms of probability of success. Indeed, in order to induce the CEO to reveal the truth, the variable part of her wage has to be set as high as possible. When choosing the best project has a relatively low value for shareholders, a high ability CEO only has low incentives to lie about her type. However, when this value (ε) increases, the relative weight of the CEO's ability in $\theta_{i,j}$ decreases. This reduces the CEO's loss when lying about her ability. It is then necessary to ensure that she will not misreport her ability. This is made by increasing a high ability CEO's variable wage when this value exceeds some threshold.

4 Board Monitoring

In this section, we assume that the board has the ability to monitor the CEO but that collusion is not achievable between the two parties[11]. When shareholders hire a board, the CEO may incur a loss F when the board finds that she has announced that the Project has a high probability of success while it is a low probability of success project, that is, the case in which she has the bonus B. The Participation and Incentive constraints are:

$$\alpha_{i,j} + \mu_{i,j}\pi\left(\theta_{i,j}\right) \geq w,\ \forall i, j \in \{H, L\} \tag{PC_{ij}}$$

$$\alpha_{i,j} + \mu_{i,j}\pi\left(\theta_{i,j}\right) \geq \alpha_{kj} + \mu_{kj}\pi\left(\theta_{i,j}\right),\ \forall i, j, k \in \{H, L\} \tag{$IC_{ij\to kj}$}$$

$$\alpha_{i,H} + \mu_{i,H}\pi\left(\theta_{i,H}\right) \geq \alpha_{kL} + \mu_{kL}\left[q_H^1\beta_k R - I\right],\ \forall i, k \in \{H, L\} \tag{$IC_{iH\to kL}$}$$

$$\alpha_{i,L} + \mu_{i,L}\pi\left(\theta_{i,L}\right) \geq \left(1 - \xi(\tau)\right)\left\{\alpha_{kH} + \mu_{kH}\left[q_L^2\beta_k R - I\right] + B\right\} + \xi(\tau)\left(w - F\right),\ \forall i, k \in \{H, L\} \tag{$IC_{iL\to kH}$}$$

In addition, we assume that the CEO faces a limited liability constraint, that is, even if the board finds that the CEO has sent the wrong signal, she cannot get less than her reservation wage plus a fixed amount, K that would be paid if not caught. K may represent the minimal compensation written in the CEO's labor contract. This gives:

$$\left(1 - \xi(\tau)\right)\{w + B\} + \xi(\tau)\left(w - F\right) \geq w + \left(1 - \xi(\tau)\right)K \tag{LL}$$

$$\Leftrightarrow B \geq \frac{\xi(\tau)}{\left(1 - \xi(\tau)\right)}F + K$$

[11] We examine the case of collusion in the next section.

The optimal incentive compensation contract when there is board monitoring and when collusion is not achievable is characterized in the following Proposition:

Proposition 2 *When the board of directors is able to monitor the CEO and when collusion is not achievable, shareholders should offer the following incentive compensation contract, U_{ij} with $i,j \in [L,H]$, to a CEO:*

$$
\begin{aligned}
U_{LL} &= w \\
U_{LH} &= w \\
U_{HL} &= w + (1-\xi(\tau))\,(p_L-\varepsilon)\,\Delta\beta\left(\frac{\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}\right) \\
U_{HH} &= \begin{cases} w+\frac{(p_L-\varepsilon)p_H(\Delta\beta)^2[(1-\xi(\tau))B-\xi(\tau)F]}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H-p_H\beta_L]} & \text{if } \varepsilon \leq \varepsilon_{ib} = \frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L} \\ w+\frac{(p_H+\varepsilon)\Delta\beta\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L[(p_H+\varepsilon)-(p_L-\varepsilon)]} & \text{if } \varepsilon \geq \varepsilon_{ib} \end{cases}
\end{aligned}
$$

Moreover, the shareholders' expected profits are

$$
W_M = \begin{cases} \begin{array}{c} E(\pi)-w-w_0 \\ -(1-\gamma)(1-\xi(\tau))(p_L-\varepsilon)\Delta\beta\frac{\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p+2\varepsilon)}\left[\begin{array}{c} \nu \\ +(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H} \end{array}\right] \end{array} & \text{if } \varepsilon \leq \varepsilon_{ib} \\ \begin{array}{c} E(\pi)-w-w_0 \\ -(1-\gamma)\Delta\beta\frac{\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p+2\varepsilon)}\left[\begin{array}{c} \nu(1-\xi(\tau))(p_L-\varepsilon) \\ +(1-\nu)(p_H+\varepsilon) \end{array}\right] \end{array} & \text{if } \varepsilon \geq \varepsilon_{ib} \end{cases}
$$

In this case, the optimal contract has the same form than without monitoring, that is, no incentive compensation for a low ability CEO and a positive rent for a high ability CEO which is higher when Project 2 is selected following her advice. However, it is worth to notice that the informational rents extracted by a CEO when there is a monitoring board of directors having no possibility to collude are lower than when there is no monitoring irrespective of the CEO's type.

Moreover, we can prove that $\varepsilon_{ib} \leq \varepsilon_{nb}$. Indeed,

$$\varepsilon_{ib} - \varepsilon_{nb} = \frac{\beta_L \Delta p - \xi p_L \Delta\beta}{(1-\xi)\,\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} - \frac{\beta_L \Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} \leq 0,$$

$$\iff 2\xi\Delta\beta\left[p_H\beta_L - p_L\beta_H\right] \leq 0.$$

Thus, we can conclude that if the board's wage is low enough, inducing the board to monitor the CEO is always beneficial for the shareholders when collusion is not achievable, that is, $W_M \geq W_{NM}$ for all $w_0 \leq \widetilde{w_0}$.

Corollary 1 *There exists a board's wage $\widetilde{w_0}$ such that for all $w_0 \leq \widetilde{w_0}$, inducing the board to monitor the CEO is always beneficial for the shareholders when collusion is not achievable.*

To sum up, in this section we show that when collusion is not achievable, it is optimal for the shareholders to choose the intensity of monitoring as low as possible. If the board has no incentives to hide the information he has gathered, it is in the shareholders' interest to choose the board with the most precise information which corresponds to the lowest intensity of monitoring.

Corollary 2 *When collusion is not achievable, it is optimal for the shareholders to select the intensity of monitoring as low as possible.*

This directly follows from the fact that W_M is decreasing in τ.

5 Collusive Board

In this section, we examine a framework in which the CEO and the board of directors may collude when this is profitable for them. We assume that all bargaining power is allocated to the board of directors.

In the following inequalities, w_L is the income of a board that announces that the project has a low probability of success, w_H is the income of a board that announces that the project has a high probability of success, $w_\emptyset$ is the income of a board that announces that it has no information regarding the project probability of success, w_0 is the income of a board when collusion cannot emerge as in the previous section.

The following constraints ensure that the Board-CEO coalition receives more when telling the truth than colluding.

$$\gamma\left[U_{LL}-w+w_L\right]+(1-\gamma)\left[U_{HL}-w+w_L\right]\geq\gamma\left[\frac{U_{LH}-w}{\tau}+w_\emptyset\right]+(1-\gamma)\left[\frac{U_{HH}-w}{\tau}+w_\emptyset\right]$$
$$\Leftrightarrow w_L\geq\gamma\left[\frac{U_{LH}-w}{\tau}-(U_{LL}-w)\right]+(1-\gamma)\left[\frac{U_{HH}-w}{\tau}-(U_{HL}-w)\right]+w_\emptyset$$

$$\gamma\left[U_{LH}-w+w_H\right]+(1-\gamma)\left[U_{HH}-w+w_H\right]\geq\gamma\left[\frac{U_{LL}-w}{\tau}+w_\emptyset\right]+(1-\gamma)\left[\frac{U_{HL}-w}{\tau}+w_\emptyset\right]$$
$$\Leftrightarrow w_H\geq\gamma\left[\frac{U_{LL}-w}{\tau}-(U_{LH}-w)\right]+(1-\gamma)\left[\frac{U_{HL}-w}{\tau}-(U_{HH}-w)\right]+w_\emptyset$$

Since we have $U_{LL}\leq U_{LH}\leq U_{HL}\leq U_{HH}$ and $\tau\geq\tau_{\min}\geq 1$, necessarily

$$\gamma\left[\frac{U_{LH}-w}{\tau}-(U_{LL}-w)\right]+(1-\gamma)\left[\frac{U_{HH}-w}{\tau}-(U_{HL}-w)\right]\leq 0$$

.

We then have 4 constraints to satisfy:

$$\begin{aligned}
w_L &\geq \gamma\left[\frac{U_{LH}-w}{\tau}-(U_{LL}-w)\right]+(1-\gamma)\left[\frac{U_{HH}-w}{\tau}-(U_{HL}-w)\right]+w_\emptyset \ (1)\\
w_H &\geq \gamma\left[\frac{U_{LH}-w}{\tau}-(U_{LL}-w)\right]+(1-\gamma)\left[\frac{U_{HH}-w}{\tau}-(U_{HL}-w)\right]+w_\emptyset \ (2)\\
w_L &\geq w_0 \ (3)\\
w_H &\geq w_0 \ (4)
\end{aligned}$$

Next, we examine when it is in the shareholders' interest to avoid collusion between the board and the CEO. Avoiding collusion is costly because shareholders have to pay higher wages to the board in order to induce him to reveal the gathered information. If those informational rents are too high, it may be optimal for the board to let collusion happen.

5.1 Collusion-Proof Contract

We first analyze a situation in which shareholders want to ensure that collusion in the board is avoided. The only case they have to take into account is when the board tells that there is a low probability of success (the board is more likely to lie when the project is of a low probability of success; there is no point in lying when it is of a high probability of success). Consequently, we always have $w_L \geq w_H$. Shareholders can try to use w_L to pay the board into revealing the truth: if they set w_L high enough, collusion might be avoided. The shareholders' expected profits have the following form:

$$\begin{aligned} W_{CP} &= E(\pi) - \gamma\nu U_{LL} - \gamma(1-\nu)U_{LH} - (1-\gamma)\nu U_{HL} - (1-\gamma)(1-\nu)U_{HH} \\ &\quad -\nu\xi(\tau)w_L - (1-\nu)\xi(\tau)w_H - (1-\xi(\tau))w_0 \end{aligned}$$

In that case, the constraint on w_L is binding. Since they want to maximize their income, shareholders set $w_H = w_\emptyset = w_0$ (because w_0 is the lowest wage of the board).

$$\begin{aligned} w_L &= \gamma\left[\frac{U_{LH}-w}{\tau} - (U_{LL}-w)\right] + (1-\gamma)\left[\frac{U_{HH}-w}{\tau} - (U_{HL}-w)\right] + w_\emptyset \\ &= \gamma\left[\frac{U_{LH}-U_{LL}}{\tau} + \frac{1-\tau}{\tau}(U_{LL}-w)\right] + (1-\gamma)\left[\frac{U_{HH}-U_{HL}}{\tau} + \frac{1-\tau}{\tau}(U_{HL}-w)\right] + w_\emptyset \\ w_H &= w_\emptyset = w_0 \end{aligned}$$

Notice that there exists τ_0 such that $w_L \geq w_0 \Longleftrightarrow \tau \leq \tau_0$. This means that for $\tau \geq \tau_0$, engaging in collusion is not beneficial for the Board-CEO coalition and the optimal contract is the same as with a perfectly honest board. Actually, when $\tau \geq \tau_0$, the board will not collude no matter what happens. Shareholders do not need to induce the board to tell the truth because he will do it anyway. So, in this case, we have:

$$w_L = w_H = w_0$$

We now characterize τ_0 :

$$\begin{aligned} w_L \geq w_0 &\iff \tau\left[\gamma\left(U_{LL}-w\right)+(1-\gamma)\left(U_{HL}-w\right)\right] \leq \gamma\left(U_{LH}-w\right)+(1-\gamma)\left(U_{HH}-w\right) \\ &\iff \tau \leq \frac{U_{HH}-w}{U_{HL}-w} = \begin{cases} \frac{1}{1-\xi(\tau_0)}\frac{p_H+\varepsilon}{p_L-\varepsilon} \text{ if } \varepsilon \geq \varepsilon_{ib} \\ \frac{p_H \Delta\beta}{p_L\beta_H - \beta_L p_H} \text{ if } \varepsilon \leq \varepsilon_{ib} \end{cases} \end{aligned}$$

And then, as $\xi(\tau) = \frac{1}{\tau}$:

$$\tau_0 = \begin{cases} 1 + \frac{p_H+\varepsilon}{p_L-\varepsilon} \text{ if } \varepsilon \geq \varepsilon_{ib} \\ \frac{p_H \Delta\beta}{p_L\beta_H - \beta_L p_H} \text{ if } \varepsilon \leq \varepsilon_{ib} \end{cases}$$

However, for an intensity of monitoring in the interval $[\tau_{\min}, \tau_0]$, since shareholders have paid enough to avoid collusion, the CEO's rents are the same as in the board Monitoring section:

$$\begin{aligned} U_{LL} &= w \\ U_{LH} &= w \\ U_{HL} &= w + (1-\xi(\tau))\left(p_L-\varepsilon\right)\Delta\beta\left(\frac{\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[\left(p_H+\varepsilon\right)-\left(p_L-\varepsilon\right)\right]}\right) \\ U_{HH} &= \begin{cases} w + \frac{(p_L-\varepsilon)p_H(\Delta\beta)^2[(1-\xi(\tau))B-\xi(\tau)F]}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H-p_H\beta_L]} \text{ if } \varepsilon \leq \varepsilon_{ib} \\ w + \frac{(p_H+\varepsilon)\Delta\beta\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L[(p_H+\varepsilon)-(p_L-\varepsilon)]} \text{ if } \varepsilon \geq \varepsilon_{ib} \end{cases} \end{aligned}$$

This is stated in the following Proposition.

Proposition 3 *Assume that collusion between the board of directors and the CEO is achievable.*

- *In the optimal collusion proof contract, shareholders should offer the same incentive compensation contract to a CEO as in the presence of a monitoring board. In this case, the shareholders' expected profits are*

$$W_{CP} = \begin{cases} E(\pi) - w - w_0 \\ -(1-\gamma)(1-\xi(\tau))(p_L-\varepsilon)\Delta\beta \frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \begin{bmatrix} \nu(1-\xi(\tau)) \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{bmatrix} \\ \textit{if } \varepsilon \leq \varepsilon_{ib} \\ E(\pi) - w - w_0 - (1-\gamma)\Delta\beta \frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \begin{bmatrix} \nu(1-\xi(\tau))^2(p_L-\varepsilon) \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})(p_H+\varepsilon) \end{bmatrix} \\ \textit{if } \varepsilon \geq \varepsilon_{ib} \end{cases}$$

- *Moreover, there exists τ_0 such that for boards of directors with an intensity of monitoring $\tau \geq \tau_0$, it is not beneficial to engage in collusion.*

The second part of this Proposition means that for some intensity of monitoring for the board, it is so difficult for the Board-CEO coalition to engage in collusion that they prefer not to collude without any shareholders' intervention. For such boards, the shareholders should not care about collusion. Since these boards would perfectly represent shareholders.

5.2 Collusion Free Contract

We now characterize the optimal collusion free contract. In this case, shareholders would have to pay too much to avoid collusion. Hence, they decide to let it happen because avoiding collusion will be too costly for them in terms of informational rents paid to the board. The shareholders' expected profits have the following form:

$$\begin{aligned} W_{CF} &= E(\pi) - \gamma\nu U_{LL} - \gamma(1-\nu)U_{LH} - (1-\gamma)\nu U_{HL} - (1-\gamma)(1-\nu)U_{HH} \\ &\quad -\nu\xi(\tau)w_L - (1-\nu)\xi(\tau)w_H - (1-\xi(\tau))w_\emptyset \end{aligned}$$

It is optimal to set $w_L = w_0$. (1) and (2) do not need to be satisfied. We therefore have:

$$w_L = w_H = w_\emptyset = w_0$$

Since the board is collusive, shareholders should not trust what the board reports. Therefore, the CEO's rents are the same as in the No Monitoring case.

$$\begin{aligned} U_{LL} &= w \\ U_{LH} &= w \\ U_{HL} &= w + \frac{B(p_L-\varepsilon)\Delta\beta}{\beta_L[\Delta p + 2\varepsilon]} \\ U_{HH} &= \begin{cases} w + \frac{B(p_L-\varepsilon)p_H(\Delta\beta)^2}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H - p_H\beta_L]} \text{ if } \varepsilon \leq \varepsilon_{nb} \\ w + \frac{B\Delta\beta(p_H+\varepsilon)}{\beta_L[\Delta p+2\varepsilon]} \text{ if } \varepsilon \geq \varepsilon_{nb} \end{cases} \end{aligned}$$

Proposition 4 *Assume that collusion between the board of directors and the CEO is achievable. In the optimal collusion free contract, shareholders should offer the same incentive compensation contract to a CEO as without any CEO's monitoring from the board.*

In this case, the shareholders' expected profits are:

$$W_{CF} = \begin{cases} E(\pi) - w_0 - w - (1-\gamma)\left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right]\frac{B\Delta\beta(p_L-\varepsilon)}{\beta_L(\Delta p+2\varepsilon)} \; \textit{if } \varepsilon \leq \varepsilon_{nb} \\ E(\pi) - w_0 - w - (1-\gamma)\left[(p_H+\varepsilon) - \nu(\Delta p + 2\varepsilon)\right]\frac{B\Delta\beta}{\beta_L(\Delta p+2\varepsilon)} \; \textit{if } \varepsilon \geq \varepsilon_{nb} \end{cases}$$

5.3 Optimal Contract with Collusion

We now use the specified form for the probability that a board with an intensity of monitoring τ has gathered the true information about the state of nature, i.e. $\xi(\tau) = \frac{1}{\tau}$.

To find the optimal contract in presence of collusion, W_{CB}, we have to compare W_{CP} and W_{CF} and find which one is the highest conditional on τ. Indeed, the shareholders will choose to design the contract (Collusion Proof or Collusion Free) in order to maximize their objective. As $\varepsilon_{ib} \leq \varepsilon_{nb}$ we only have three cases:

1. $\varepsilon \leq \varepsilon_{ib}$

2. $\varepsilon_{ib} \leq \varepsilon \leq \varepsilon_{nb}$

3. $\varepsilon_{nb} \leq \varepsilon$

The following Proposition characterizes the optimal contract when collusion is achievable.

Proposition 5 *For all $\tau \in [\tau_{\min}, \tau_0]$, the optimal contract is the collusion proof contract for all ε.*

This allows us to state that the shareholders's welfare, W_{CB} that depends on τ is, for all $\tau \in [\tau_{\min}, \tau_0]$:

$$W_{CB}(\tau) = max(W_{CP}; W_{CF}) = W_{CP}(\tau)$$

This is an important result as it means that when collusion is achievable and is profitable for the Board-CEO coalition , it is always beneficial for the shareholders to offer a contract preventing collusion to emerge. Therefore, the optimal compensation contract deters any attempt of collusion between the CEO and the board members even though this is costly in terms of informational rents. However, the informational gains from monitoring always exceed the costs of those informational rents paid to the board in order to induce him to monitor efficiently and not to collude.

This result and the results shown in the previous sections allow us to characterize the optimal structure of the board of directors from the shareholders' perspective.

6 Optimal Structure of the Board

We are now able to find what is the optimal board's intensity of monitoring τ^* maximizing the piecewise continuous shareholders's welfare $W_{CB}(\tau)$.

The optimal intensity of monitoring is related to the incentive compensation offered to the CEO. Shareholders choose the intensity of monitoring that will allow them to pay the lowest informational rents to the CEO, that is, the most efficient incentive compensation contract.

We also have to consider corner solutions such as $\tau \in [1; \tau_0]$.

To be able to solve this problem, we assume that shareholders optimally set the penalty F. This implies that F has to be set as high as possible, such that the CEO limited liability constraint binds: $\left[B - \frac{1}{\tau-1}F\right] = K$.

To simplify the computations, we rewrite the intervals of discontinuity of $W_{CB}(\tau)$ in order to build them with respect to τ. This gives

$$\varepsilon \geq \varepsilon_{ib} = \frac{p_H \beta_L \Delta p - \frac{1}{\tau} p_H p_L \Delta\beta}{\frac{(\tau-1)}{\tau}\Delta\beta p_H + p_L \beta_H - p_H \beta_L} \Leftrightarrow \tau \leq \frac{\frac{\Delta\beta p_H}{[p_L\beta_H - p_H\beta_L]}}{\frac{\Delta\beta p_H}{[p_L\beta_H - p_H\beta_L]} - \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}} = \widehat{\tau}$$

Hence, when $\widehat{\tau} \geq \tau_0$ or $\left[\frac{\Delta\beta p_H}{[p_L\beta_H - p_H\beta_L]} - \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] \leq 0, \iff \varepsilon \ \geq \frac{p_H\beta_L\Delta p}{\Delta\beta p_H + [p_L\beta_H - p_H\beta_L]} = \widehat{\varepsilon},$

$$\varepsilon \geq \varepsilon_{ib} \text{ for all } \tau$$

and when $\widehat{\tau} \leq \tau_0$ and $\left[\frac{\Delta\beta p_H}{[p_L\beta_H - p_H\beta_L]} - \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] \geq 0, \Longleftrightarrow \varepsilon \leq \frac{p_H\beta_L\Delta p}{\Delta\beta p_H + [p_L\beta_H - p_H\beta_L]} = \widehat{\varepsilon},$

$$\begin{aligned} \varepsilon &\geq \varepsilon_{ib} \text{ for } \tau \leq \widehat{\tau}, \text{ and} \\ \varepsilon &\leq \varepsilon_{ib} \text{ for } \tau \geq \widehat{\tau} \end{aligned}$$

The shareholders have the following objective function[12]:

When $\varepsilon \leq \widehat{\varepsilon}$ and $\widehat{\tau} \leq \tau_0$

$$W_{CB}(\tau) = \begin{cases} \begin{array}{c} E(\pi) - w - w_0 \\ -(1-\gamma)\Delta\beta \frac{\left[B - \frac{1}{\tau-1}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\begin{array}{c} \nu(\frac{\tau-1}{\tau})^2(p_L - \varepsilon) \\ +(1-\nu+\frac{\nu}{\tau^2})(p_H+\varepsilon) \end{array} \right] \end{array} & \text{if } \tau \leq \widehat{\tau} \\ \begin{array}{c} E(\pi) - w - w_0 \\ -(1-\gamma)(\frac{\tau-1}{\tau})(p_L-\varepsilon)\Delta\beta \frac{\left[B - \frac{1}{\tau-1}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\begin{array}{c} \nu(\frac{\tau-1}{\tau}) \\ +(1-\nu+\frac{\nu}{\tau^2})\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{array} \right] \end{array} & \text{if } \widehat{\tau} \leq \tau \leq \tau_0 \\ \begin{array}{c} E(\pi) - w - w_0 \\ -(1-\gamma)(\frac{\tau-1}{\tau})(p_L-\varepsilon)\Delta\beta \frac{\left[B - \frac{1}{\tau-1}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right] \end{array} & \text{if } \tau \geq \tau_0 \end{cases}$$

When $\varepsilon \geq \widehat{\varepsilon}$, or $\widehat{\tau} \geq \tau_0$

$$W_{CB}(\tau) = \begin{cases} \begin{array}{c} E(\pi) - w - w_0 \\ -(1-\gamma)\Delta\beta \frac{\left[B - \frac{1}{\tau-1}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\begin{array}{c} \nu(\frac{\tau-1}{\tau})^2(p_L - \varepsilon) \\ +(1-\nu+\frac{\nu}{\tau^2})(p_H+\varepsilon) \end{array} \right] \end{array} & \text{if } \tau \leq \tau_0 \\ \begin{array}{c} E(\pi) - w - w_0 \\ -(1-\gamma)\Delta\beta \frac{\left[B - \frac{1}{\tau-1}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\nu(\frac{\tau-1}{\tau})(p_L-\varepsilon) + (1-\nu)(p_H+\varepsilon)\right] \end{array} & \text{if } \tau \geq \tau_0 \end{cases}$$

Recall that shareholders set the penalty F as high as possible, such that $\left[B - \frac{1}{\tau-1}F\right] = K$. The following Proposition summarizes our results:

[12] As $\widehat{\tau} \leq \tau_0$ $\forall \varepsilon \leq \widehat{\varepsilon}$.

Proposition 6 *When $\varepsilon \leq \widehat{\varepsilon}$, and $\tau_0 \geq \widehat{\tau}$, incentive compensation contracts are sufficient to extract the optimal amount of information from the CEO whatever the degree of monitoring. It is therefore optimal for the shareholders to select a board of directors with a low intensity of monitoring, that is, $\tau^* = \widehat{\tau}$ and to offer contracts avoiding collusion between the board and the CEO.*

In all other cases, in order to induce the CEO to reveal the optimal amount of information, shareholders have to select a board of directors with a high intensity of monitoring, $\tau^* = \tau_0$. In this case, the shareholders should not care about collusion because collusion is not profitable for such boards.

Contrary to the usual idea that the optimal board should strongly monitor the CEO, we find that incentive compensation contracts may be sufficient in order to induce the CEO to reveal the optimal amount of information. In such cases, it is in the shareholders' interest to select a board of directors with a low intensity of monitoring. However, the result is not due, as in Adams and Ferreira (2007), to the fact that the CEO is more prone to reveal information to a "friendly" board. Here, there is a trade-off between the information that shareholders may extract from the board and the costs of providing the right incentives to the CEO and avoiding collusion. The higher the τ, the more difficult to engage in collusion for the Board-CEO coalition , but the less information about the projects they have.

The intuition for this result is as follows. Shareholders should not care about hiring a board with a high intensity of monitoring: (i) when it is too costly to do so and (ii) when potential collusion between the CEO and the board has not a big impact on the firm's decision which is the case when choosing the best project has not a high value for shareholders and the intensity of monitoring necessary to have a perfectly honest board is too high. Collusion allows the CEO to undertake projects that may not be optimal for shareholders. This means that the lower the value of choosing the best project for shareholders, the lower the costs of collusion. Therefore, choosing a board with a low intensity of monitoring may be optimal for two reasons. First, choosing a higher intensity of monitoring leads to extract less information. Second, it would be too costly in terms of incentive compensation to choose a perfectly honest board (because τ_0 is high), choosing a board with a low intensity of monitoring is therefore optimal. In theses cases, deterring collusion is less important than gathering information about the projects.

However, in all other cases, that is, when choosing the best project is highly valuable for shareholders or when the intensity of monitoring necessary to have a perfectly honest board is low enough, it is optimal for shareholders to choose a board with a high intensity of monitoring as compensation contracts are not powerful enough to provide the optimal right incentives to the CEO. Consequently, the optimal structure is a perfectly honest board and the shareholders should not care about collusion because collusion is not profitable for such boards.

Put differently, the optimal structure is a board with a low intensity of monitoring when:

(i) choosing the best project is not very valuable for shareholders, and

(ii) the intensity of monitoring necessary to have a perfectly honest board is too high, that is, the loss of information about the projects that would be associated with the choice of a perfectly honest board would be too important.

6.1 Policy and Empirical Predictions

It is optimal for shareholders to select a board with a low intensity of monitoring for firms with a stable economic environment (for instance, in industries and sectors having achieved a high degree of maturity or in low risk industries such as building, transport, chemistry) and with which it is difficult to find efficient and absolutely independent directors (firms for which only executives are able to gather information about the projects for strategic reasons such as Investment Banking, Petroleum Industry, Aeronautics, Military sectors. . .).

These implications are consistent with empirical results that link firms' risk or complexity (that can be interpreted as the value of choosing the best project for shareholders) to the intensity of monitoring. For example, Demsetz and Lehn (1985) show that riskier environments should be associated with more monitoring. Ferreira, Ferreira and Raposo (2011) also find that more complex firms require more monitoring.

As firms relying on incentive based compensations schemes usually need a high intensity of monitoring, our result, stating that the intensity with which boards monitor CEOs should be higher in innovative industries (in which choosing the best project is highly valuable for shareholders), is therefore consistent with both Murphy (1999) who shows that incentive compensations are lower in regulated utilities than in other industries and Ittner, Lambert and Larcker (2003) or Murphy (2003) who find that stock-based compensation is more frequently used by new economy firms than by old economy firms.

One way to test empirically the results of the optimal intensity of monitoring and the CEO's incentive compensation, would be to use as a proxy of the intensity of monitoring one of the following: the number of non independent directors, the number of interlocked directors, the number of mandates held by each director, the power of the main committees, or the quality of the internal audit process. In particular, we could test whether across industries that differ in the level of complexity they face and in which it is difficult to hire informed directors not connected with the firm's top executives there are differences across the intensity of monitoring and the executive compensation schemes.

Moreover, the intensity of monitoring of the board, τ, can also be interpreted as the shareholders' willingness to increase the degree of toughness and enforceability of the Corporate Governance regulations as well as the laws against collusion. A proxy for τ would be the toughness of laws and regulations, the ownership concentration or the ownership structure. Therefore, it would be be interesting to test if the boards' structures of firms having the previous characteristics have changed in countries that have modified the Corporate Governance regulations (see, Cornelli, Kominek and Ljungqvist, 2012), or for firms in which we observe a modification of the ownership structure (see, Ferreira, Ferreira and Raposo, 2011).

7 Conclusion

In this paper, we analyze the effect of collusion between a board of directors and a CEO on the optimal intensity of monitoring. We also characterize the optimal incentive compensation contracts.

We show that when there is no CEO's monitoring by the directors the variable part of the wage is higher for a high ability CEO than for a low ability CEO. When we introduce board monitoring, still without collusion, the board behaves as a perfectly honest board and the optimal compensation contract takes the same form as without monitoring. Allowing for the possibility of collusion between the board and the CEO, we show that the optimal contract is collusion proof: it is always optimal for the shareholders to offer a contract preventing collusion to emerge. We also prove that there exists an intensity with which the board monitors the CEO above which it is not profitable for the Board-CEO coalition to engage in collusion. Such boards therefore behave as perfectly honest boards.

In addition, we derive the optimal intensity with which the board of directors monitors the CEO from the shareholders point of view. Contrary to the usual idea that an optimal board should strongly monitor, we find that incentive compensation contracts may be sufficient to induce the CEO to reveal the optimal amount of information irrespective of the degree of monitoring. In this case, it is in the shareholders' interest to choose a low intensity of monitoring of the CEO by the board of directors. More precisely, the optimal structure is a board with a low intensity of monitoring when the value of choosing the best project for shareholders is low, and the intensity of monitoring necessary to have a perfectly honest board is high enough. Finally, we provide practical and empirical implications of our model.

8 Appendix

Proof of Proposition 1. When there isn't any CEO's monitoring from the board, shareholders maximize their expected profits under the usual Participation and Incentive constraints. PC_{ij} is the Participation constraint of a CEO with ability $i \in \{H, L\}$ when the state of nature is $j \in \{H, L\}$. The Participation constraints ensure that the CEO will earn at least her reservation wage w. $IC_{ij \to kl}$ is the Incentive constraint of a CEO who reveals that her ability is $k \in \{H, L\}$ and the is $l \in \{H, L\}$ while her true ability is i and the true probability of success of the project is j. The Incentives constraints ensure that the CEO earns a higher wage revealing the truth than lying to the shareholders. Through this process, shareholders induce the CEO to reveal his real type. As usual in this kind of problem, the binding constraints are :

$$\alpha_{LL} + \mu_{LL} \left[p_L \beta_L R - I \right] = w \tag{PC_{LL}}$$

$$\alpha_{LH} + \mu_{LH} \left[\left(p_H + \varepsilon \right) \beta_L R - I \right] + B = \alpha_{LL} + \mu_{LL} \left[p_L \beta_L R - I \right] + \mu_{LL} R \beta_L \Delta p \tag{$IC_{LH \to LL}$}$$

$$\Leftrightarrow \alpha_{LH} + \mu_{LH} \left[\left(p_H + \varepsilon \right) \beta_L R - I \right] + B = w + \mu_{LL} R \beta_L \Delta p$$

$$\alpha_{HL} + \mu_{HL} \left[p_L \beta_H R - I \right] = \alpha_{LH} + \mu_{LH} \left[\left(p_H + \varepsilon \right) \beta_L R - I \right] + B + \mu_{LH} R \left[p_L \beta_H - p_H \beta_L - \varepsilon \left(\beta_L + \beta_H \right) \right] \tag{$IC_{HL \to LH}$}$$

$$\Leftrightarrow \alpha_{HL} + \mu_{HL} \left[p_L \beta_H R - I \right] = w + \mu_{LL} R \beta_L \Delta p + \mu_{LH} R \left[p_L \beta_H - p_H \beta_L - \varepsilon \left(\beta_L + \beta_H \right) \right] + B$$

$$\alpha_{HH}+\mu_{HH}\left[(p_H+\varepsilon)\beta_H R-I\right]+B=\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]+\mu_{HL}R\beta_H\Delta p \tag{$IC_{HH\to HL}$}$$

$$\Leftrightarrow \alpha_{HH}+\mu_{HH}\left[(p_H+\varepsilon)\beta_H R-I\right]+B=\left(\begin{array}{c} w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[\begin{array}{c} p_L\beta_H-p_H\beta_L \\ -\varepsilon\left(\beta_L+\beta_H\right)\end{array}\right] \\ +\mu_{HL}R\beta_H\Delta p+B \end{array}\right)$$

In order to minimize the CEO's informational rents, shareholders set μ_{LL}, μ_{LH} and μ_{HL} as low as possible while satisfying the other Incentive constraints. We now check what are the conditions due to the other Incentive constraints (and will check later that Participation constraints are satisfied). There is no constraint on μ_{LL}, we can therefore set:

$$\mu_{LL}=0$$

$$\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right]\geq\alpha_{LH}+\mu_{LH}\left[(p_L-\varepsilon)\beta_L R-I\right]+B= \tag{$IC_{LL\to LH}$}$$

$$\alpha_{LH}+\mu_{LH}\left[(p_H+\varepsilon)\beta_L R-I\right]-\mu_{LH}R\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]+B$$

$$\Leftrightarrow \mu_{LH}\geq\frac{B}{R\beta_L\left[\Delta p+2\varepsilon\right]}$$

and then

$$\mu_{LH}=\frac{B}{R\beta_L\left[\Delta p+2\varepsilon\right]}$$

$$\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]+B\geq\alpha_{LL}+\mu_{LL}\left[p_H\beta_H R-I\right]=w+\mu_{LL}R\left[p_H\beta_H-p_L\beta_L\right] \tag{$IC_{HH\rightarrow LL}$}$$

$$\Leftrightarrow\left[\begin{array}{c} w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[\begin{array}{c} p_L\beta_H-p_H\beta_L \\ -\varepsilon\left(\beta_L+\beta_H\right) \end{array}\right] \\ +\mu_{HL}R\beta_H\Delta p+B \end{array}\right]\geq w+\mu_{LL}R\left[p_H\beta_H-p_L\beta_L\right]$$

$$\Leftrightarrow\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+\mu_{HL}R\beta_H\Delta p+B\geq 0$$

which is satisfied, as $\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\geq 0$.

$$\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]\geq\alpha_{LL}+\mu_{LL}\left[p_L\beta_H R-I\right]=w+\mu_{LL}Rp_L\Delta\beta \tag{$IC_{HL\rightarrow LL}$}$$

$$\Leftrightarrow w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+B\geq w+\mu_{LL}Rp_L\Delta\beta$$

$$\Leftrightarrow\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+B\geq 0$$

As $\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\geq 0$, $(IC_{HL\rightarrow LL})$ is not binding.

$$\alpha_{LH}+\mu_{LH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]+B\geq\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]+B \tag{$IC_{LH\rightarrow HH}$}$$

$$=\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]+B-\mu_{HH}R\left(p_H+\varepsilon\right)\Delta\beta$$

$$\Leftrightarrow w+\mu_{LL}R\beta_L\Delta p\geq w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]$$

$$+\mu_{HL}R\beta_H\Delta p+B-\mu_{HH}R\left(p_H+\varepsilon\right)\Delta\beta$$

$$\Leftrightarrow\mu_{HH}\geq\frac{B}{R\beta_L\left[\Delta p+2\varepsilon\right]}=\mu_{LH}.$$

This is satisfied from the Spence Mirrlees condition.

$$\begin{aligned}\alpha_{LH}+\mu_{LH}\left[(p_H+\varepsilon)\beta_L R-I\right]+B \geq\ & \alpha_{HL}+\mu_{HL}\left[p_H\beta_L R-I\right] \\ &=\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]-\mu_{HL}R\left[p_L\beta_H-p_H\beta_L\right]\end{aligned} \tag{$IC_{LH\to HL}$}$$

$$\Leftrightarrow w+\mu_{LL}R\beta_L\Delta p \geq \left(\begin{array}{c} w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[\begin{array}{c} p_L\beta_H-p_H\beta_L \\ -\varepsilon\left(\beta_L+\beta_H\right)\end{array}\right] \\ -\mu_{HL}R\left[p_L\beta_H-p_H\beta_L\right]+B\end{array}\right)$$

$$\Leftrightarrow \mu_{HL} \geq \frac{B\left[p_L-\varepsilon\right]\Delta\beta}{R\beta_L\left[\Delta p+2\varepsilon\right]\left[p_L\beta_H-p_H\beta_L\right]}=\mu_{HL}^1$$

$$\begin{aligned}\alpha_{HH}+\mu_{HH}\left[(p_H+\varepsilon)\beta_H R-I\right] \geq\ & \alpha_{LH}+\mu_{LH}\left[(p_H+\varepsilon)\beta_H R-I\right] \\ &=w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left(p_H+\varepsilon\right)\Delta\beta\end{aligned} \tag{$IC_{HH\to LH}$}$$

$$\Leftrightarrow \left(\begin{array}{c} w+\mu_{LL}R\beta_L\Delta p+\mu_{HL}R\beta_H\Delta p+B \\ +\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\end{array}\right) \geq w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left(p_H+\varepsilon\right)\Delta\beta$$

$$\Leftrightarrow \mu_{HL} \geq \frac{B\Delta\beta}{R\Delta p\beta_H\beta_L}=\mu_{HL}^2$$

$$\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right] \geq \alpha_{HL}+\mu_{HL}\left[p_L\beta_L R-I\right]=\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]-\mu_{HL}Rp_L\Delta\beta \tag{$IC_{LL\to HL}$}$$

$$\Leftrightarrow w \geq w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]-\mu_{HL}Rp_L\Delta\beta+B$$

$$\mu_{HL} \geq \frac{B\left(p_L-\varepsilon\right)}{Rp_L\beta_L\left[\Delta p+2\varepsilon\right]}$$

This is always verified as $\frac{B(p_L-\varepsilon)}{R\beta_L[\Delta p+2\varepsilon]p_L} \leq \mu_{LH}$ and due to the Spence Mirrlees condition.

$$\begin{aligned}
\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right] \geq \alpha_{LH}+\mu_{LH}\left[\left(p_L-\varepsilon\right)\beta_L R-I\right]+B = \qquad (IC_{LL\rightarrow LH})\\
\alpha_{LH}+\mu_{LH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]-\mu_{LH}R\beta_L\left[\left(p_H+\varepsilon\right)-\left(p_L-\varepsilon\right)\right]+B\\
\Leftrightarrow w \geq w+\mu_{LL}R\beta_L\Delta p-\mu_{LH}R\beta_L\left[\Delta p+2\varepsilon\right]+B
\end{aligned}$$

$(IC_{LL\rightarrow LH})$ is thus not binding.

$$\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right] \geq \alpha_{HH}+\mu_{HH}\left[\left(p_L-\varepsilon\right)\beta_H R-I\right]+B \qquad (IC_{HL\rightarrow HH})$$

$$= w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+\mu_{HL}R\beta_H\Delta p-\mu_{HH}R\beta_H\left[\Delta p+2\varepsilon\right]+2B$$

$$\Leftrightarrow \mu_{HH} \geq \frac{B\left[p_L-\varepsilon\right]\Delta\beta}{R\beta_L\left[\Delta p+2\varepsilon\right]^2\left[p_L\beta_H-p_H\beta_L\right]}\Delta p+\frac{B}{R\beta_H\left[\Delta p+2\varepsilon\right]} = \mu_{HH}^1$$

$$\begin{aligned}
\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right] \geq \alpha_{HH}+\mu_{HH}\left[\left(p_L-\varepsilon\right)\beta_L R-I\right]+B = \qquad (IC_{LL\rightarrow HH})\\
\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]-\mu_{HH}R\left[\left(p_H+\varepsilon\right)\beta_H-\left(p_L-\varepsilon\right)\beta_L\right]+B
\end{aligned}$$

$$\Leftrightarrow \mu_{HH} \geq \begin{array}{c} \mu_{LH}\frac{\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]}{\left[\left(p_H\beta_H-p_L\beta_L\right)+\varepsilon\left(\beta_L+\beta_H\right)\right]} \\ +\mu_{HL}\frac{\beta_H\Delta p}{\left[\left(p_H\beta_H-p_L\beta_L\right)+\varepsilon\left(\beta_L+\beta_H\right)\right]} \\ +\frac{2B}{R\left[\left(p_H\beta_H-p_L\beta_L\right)+\varepsilon\left(\beta_L+\beta_H\right)\right]} \end{array} = \mu_{HH}^2$$

We therefore have:

$$\begin{aligned}
\mu_{LL} &= 0\\
\mu_{LH} &= \frac{B}{R\beta_L\left[\Delta p+2\varepsilon\right]}\\
\mu_{HL} &= \max\left\{\mu_{LH};\mu_{HL}^1;\mu_{HL}^2\right\}\\
\mu_{HH} &\geq \max\left\{\mu_{HL};\mu_{HH}^1;\mu_{HH}^2\right\}
\end{aligned}$$

We now have to show that $\mu_{HL} = \begin{cases} \mu_{HL}^1 \text{ if } \varepsilon \leq \frac{\beta_L \Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{nb} \\ \mu_{HL}^2 \text{ if } \varepsilon \geq \frac{\beta_L \Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{nb} \end{cases}$

We only have six cases:

1. $\mu_{LH} \leq \mu_{HL}^1 \leq \mu_{HL}^2 \Longleftrightarrow \mu_{HL} = \mu_{HL}^2$ if $\varepsilon\Delta\beta \leq \Delta p \beta_L \leq \varepsilon \left[\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L\right]$. Indeed, we have :

$$\begin{aligned} \mu_{LH} \leq \mu_{HL}^1 \leq \mu_{HL}^2 \quad &\Longleftrightarrow \quad \frac{1}{\Delta p + 2\varepsilon} \leq \frac{(p_L - \varepsilon)\Delta\beta}{(\Delta p + 2\varepsilon)(p_L\beta_H - p_H\beta_L)} \leq \frac{\Delta\beta}{\Delta p \beta_H} \\ &\Longleftrightarrow \begin{cases} p_L\beta_H - p_H\beta_L \leq (p_L - \varepsilon)\Delta\beta \\ (p_L - \varepsilon)\Delta p \beta_H \leq (\Delta p + 2\varepsilon)(p_L\beta_H - p_H\beta_L) \end{cases} \\ &\Longleftrightarrow \begin{cases} \varepsilon\Delta\beta \leq \Delta p \beta_L \\ \Delta p \beta_L \leq \varepsilon\left(\Delta\beta + \beta_H\frac{p_L}{p_H} - \beta_L\right) \end{cases} \end{aligned}$$

For the following cases (2, 3 and 4), we use the same inequalities to obtain.

2. $\mu_{HL}^1 \leq \mu_{LH} \leq \mu_{HL}^2 \Longleftrightarrow \mu_{HL} = \mu_{HL}^2$ if $\Delta p \beta_L \leq \varepsilon\Delta\beta$

3. $\mu_{LH} \leq \mu_{HL}^2 \leq \mu_{HL}^1 \Longleftrightarrow \mu_{HL} = \mu_{HL}^1$ if $\varepsilon\left[\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L\right] \leq \Delta p \beta_L \leq 2\varepsilon\Delta\beta$

4. $\mu_{HL}^2 \leq \mu_{LH} \leq \mu_{HL}^1 \Longleftrightarrow \mu_{HL} = \mu_{HL}^1$ if $\Delta p \beta_L \geq 2\varepsilon\Delta\beta$

5. $\mu_{HL}^2 \leq \mu_{HL}^1 \leq \mu_{LH} \Longleftrightarrow$ impossible. Indeed, we would eventually obtain

$$\varepsilon\Delta\beta \geq \Delta p \beta_L \geq \varepsilon\left(\Delta\beta + \beta_H\frac{p_L}{p_H} - \beta_L\right)$$

which is not possible because the last term is strictly superior to the first one.

6. $\mu_{HL}^1 \leq \mu_{HL}^2 \leq \mu_{LH} \Longleftrightarrow$ impossible

We therefore have the result of the lemma.

And then :

$$U_{LL} = \alpha_{LL} + \mu_{LL}\left[p_L\beta_L R - I\right] = w$$

$$U_{LH} = \alpha_{LH} + \mu_{LH}\left[(p_H + \varepsilon)\beta_L R - I\right] = w + \mu_{LL} R \beta_L \Delta p = w$$

$$U_{HL}=\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]=w+\frac{B\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]}{\beta_L\left[\Delta p+2\varepsilon\right]}+B$$
$$\Longleftrightarrow U_{HL}=w+\frac{B\left(p_L-\varepsilon\right)\Delta\beta}{\beta_L\left[\Delta p+2\varepsilon\right]}$$

Moreover, when $\varepsilon\leq\frac{\beta_L\Delta p}{\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{nb}$

$$U_{HH}=\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]=w+\frac{B\left(p_L-\varepsilon\right)\Delta\beta}{\beta_L\left[\Delta p+2\varepsilon\right]}+\frac{B\left[p_L-\varepsilon\right]\Delta\beta\beta_H\Delta p}{\beta_L\left[\Delta p+2\varepsilon\right]\left[p_L\beta_H-p_H\beta_L\right]}$$
$$\Longleftrightarrow U_{HH}=w+\frac{B\left(p_L-\varepsilon\right)p_H\left(\Delta\beta\right)^2}{\beta_L\left[\Delta p+2\varepsilon\right]\left[p_L\beta_H-p_H\beta_L\right]}$$

Moreover, when $\varepsilon\geq\frac{\beta_L\Delta p}{\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{nb}$

$$U_{HH}=\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]=w+\frac{B\left(p_L-\varepsilon\right)\Delta\beta}{\beta_L\left[\Delta p+2\varepsilon\right]}+\frac{B\Delta\beta}{\beta_L}$$
$$\Longleftrightarrow U_{HH}=w+\frac{B\Delta\beta\left(p_H+\varepsilon\right)}{\beta_L\left[\Delta p+2\varepsilon\right]}$$

To sum up, here are the CEO' informational rents when there is no monitoring:

$$\begin{aligned}
U_{LL} &= w\\
U_{LH} &= w\\
U_{HL} &= w+\frac{B\left(p_L-\varepsilon\right)\Delta\beta}{\beta_L\left[\Delta p+2\varepsilon\right]}\\
U_{HH} &= \begin{cases} w+\frac{B(p_L-\varepsilon)p_H(\Delta\beta)^2}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H-p_H\beta_L]} \text{ if } \varepsilon\leq\varepsilon_{nb}\\ w+\frac{B\Delta\beta(p_H+\varepsilon)}{\beta_L[\Delta p+2\varepsilon]} \text{ if } \varepsilon\geq\varepsilon_{nb}\end{cases}
\end{aligned}$$

We can verify now that we have

$$U_{LL}\leq U_{LH}\leq U_{HL}\leq U_{HH}$$

When $\varepsilon\leq\varepsilon_{nb}$, we need to see if $\frac{p_H\Delta\beta}{p_L\beta_H-p_H\beta_L}\geq 1$, which is true since $p_L\beta_H-p_H\beta_L=$

$p_H \Delta\beta - \beta_H \Delta p$. Subsequently, we have $U_{HL} \leq U_{HH}$. When $\varepsilon \geq \varepsilon_{nb}$, since $p_L - \varepsilon \leq p_H + \varepsilon$, we also have $U_{HL} \leq U_{HH}$.

Rewriting the shareholders' expected profits depending on those informational rents, when there is no monitoring, we have:

$$W_{NM} = E(\pi) - \gamma\nu U_{LL} - \gamma(1-\nu)U_{LH} - (1-\gamma)\nu U_{HL} - (1-\gamma)(1-\nu)U_{HH}$$

This gives, for $\varepsilon \leq \frac{\beta_L \Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{nb}$

$$W_{NM} = E(\pi) - w - \frac{(1-\gamma)B(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)}\left[\frac{-\nu\beta_H\Delta p + p_H\Delta\beta}{p_L\beta_H - p_H\beta_L}\right]$$

And for $\varepsilon \geq \frac{\beta_L \Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{nb}$

$$W_{NM} = E(\pi) - w - (1-\gamma)(p_H + \varepsilon - \nu\Delta p - 2\nu\varepsilon)\frac{B\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)}$$

■

Proof of Proposition 2. The binding constraints are:

$$\alpha_{LL} + \mu_{LL}[p_L\beta_L R - I] = w \qquad (PC_{LL})$$

$$\alpha_{LH} + \mu_{LH}[(p_H+\varepsilon)\beta_L R - I] + B = w + \mu_{LL}R\beta_L\Delta p \qquad (IC_{LH\to LL})$$

$$\alpha_{HL} + \mu_{HL}[p_L\beta_H R - I] = (1-\xi(\tau))\left\{\begin{array}{c} w + \mu_{LL}R\beta_L\Delta p \\ +\mu_{LH}R[p_L\beta_H - p_H\beta_L - \varepsilon(\beta_L+\beta_H)] \\ +B \end{array}\right\} + \xi(\tau)(w-F) \qquad (IC_{HL\to LH})$$

$$\alpha_{HH}+\mu_{HH}\left[(p_H+\varepsilon)\beta_H R-I\right]+B=\begin{array}{c}(1-\xi(\tau))\left\{\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\+\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\\+B\end{array}\right\}\\+\xi(\tau)(w-F)+\mu_{HL}R\beta_H\Delta p\end{array}\tag{$IC_{HH\to HL}$}$$

Again, in order to minimize the informational rents, shareholders will set μ_{LL}, μ_{LH} and μ_{HL} as low as possible while satisfying the other incentive constraints. We now check what are the conditions due to the other Incentive constraints (and will check later that Participation constraints are satisfied).

$$\mu_{LL}=0$$

$$\alpha_{HH}+\mu_{HH}\left[(p_H+\varepsilon)\beta_H R-I\right]+B\geq\alpha_{LL}+\mu_{LL}\left[p_H\beta_H R-I\right]\tag{$IC_{HH\to LL}$}$$

$$=w+\mu_{LL}R\left[p_H\beta_H-p_L\beta_L\right]=w$$

$$\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right]\geq(1-\xi(\tau))\left\{\alpha_{LH}+\mu_{LH}\left[(p_L-\varepsilon)\beta_L R-I\right]+B\right\}+\xi(\tau)(w-F)=\tag{$IC_{LL\to LH}$}$$

$$(1-\xi(\tau))\left\{\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\-\mu_{LH}R\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]+B\end{array}\right\}+\xi(\tau)(w-F)$$

$$\Leftrightarrow\mu_{LH}\geq\frac{B-\frac{\xi(\tau)}{(1-\xi(\tau))}F}{R\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}$$

As $\frac{\xi(\tau)}{(1-\xi(\tau))}F-B\leq0$, we have

$$\mu_{LH}=\frac{B-\frac{\xi(\tau)}{(1-\xi(\tau))}F}{R\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}$$

$$\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]\geq\alpha_{LL}+\mu_{LL}\left[p_L\beta_H R-I\right]=w+\mu_{LL}Rp_L\Delta\beta \qquad (IC_{HL\to LL})$$

$$\Leftrightarrow(1-\xi(\tau))\left\{\begin{array}{c}\mu_{LL}R\beta_L\Delta p\\ +\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+B\end{array}\right\}-\xi(\tau)F\geq\mu_{LL}Rp_L\Delta\beta$$

$$\Leftrightarrow\mu_{LH}\geq\frac{\frac{\xi(\tau)}{(1-\xi(\tau))}F-B}{R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]}$$

As $\frac{\xi(\tau)}{(1-\xi(\tau))}F-B\leq 0$, this is satisfied

$$\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]+B\geq\begin{array}{c}\alpha_{LH}+\mu_{LH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]+B\\ =w+\mu_{LL}R\beta_L\Delta p+\mu_{LH}R\left(p_H+\varepsilon\right)\Delta\beta\end{array}$$

$$\Leftrightarrow\left\{\begin{array}{c}(1-\xi(\tau))\left\{\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\ +\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\\ +B\end{array}\right\}\\ +\xi(\tau)\left(w-F\right)+\mu_{HL}R\beta_H\Delta p\end{array}\right\}\geq\left(\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\ +\mu_{LH}R\left(p_H+\varepsilon\right)\Delta\beta\end{array}\right)$$

$$\mu_{HL}\geq\frac{\left(p_H+\varepsilon\right)\Delta\beta-(1-\xi(\tau))\left(p_L-\varepsilon\right)\Delta\beta}{\beta_H\Delta p}\left(\frac{\left[B-\frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{R\beta_L\left[\left(p_H+\varepsilon\right)-\left(p_L-\varepsilon\right)\right]}\right)$$

$$\Leftrightarrow\mu_{HL}\geq\mu_{LH}\frac{\left[\left(p_H+\varepsilon\right)-(1-\xi(\tau))\left(p_L-\varepsilon\right)\right]\Delta\beta}{\beta_H\Delta p}=\mu_{HL}^1$$

$$\alpha_{LH}+\mu_{LH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]+B\geq\alpha_{HL}+\mu_{HL}\left[p_H\beta_L R-I\right] \qquad (IC_{LH\to HL})$$

$$=\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]-\mu_{HL}R\left[p_L\beta_H-p_H\beta_L\right]$$

$$\Leftrightarrow w+\mu_{LL}R\beta_L\Delta p\geq\left\{\begin{array}{c}(1-\xi(\tau))\left\{\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\ +\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]+B\end{array}\right\}\\ +\xi(\tau)\left(w-F\right)-\mu_{HL}R\left[p_L\beta_H-p_H\beta_L\right]\end{array}\right\}$$

$$\mu_{HL}\geq\frac{(1-\xi(\tau))\left(p_L-\varepsilon\right)\Delta\beta}{\left[p_L\beta_H-p_H\beta_L\right]}\mu_{LH}=\mu_{HL}^2$$

We can verify that

$$\mu_{HL}=\begin{cases}\mu_{HL}^{1}\text{ if }\varepsilon\geq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{ib}\\ \mu_{HL}^{2}\text{ if }\varepsilon\leq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{ib}\end{cases}$$

Indeed, we have :

$$\begin{aligned}\mu_{HL}=\mu_{HL}^{1}&\iff \mu_{HL}^{1}\geq\mu_{HL}^{2}\\ &\iff \frac{\left[\Delta p+2\varepsilon+\xi(\tau)\left(p_L-\varepsilon\right)\right]\Delta\beta}{\beta_H\Delta p}\geq\frac{\left(1-\xi(\tau)\right)\left(p_L-\varepsilon\right)\Delta\beta}{p_L\beta_H-p_H\beta_L}\\ &\iff \varepsilon\left[\begin{array}{c}\left(2-\xi(\tau)\right)\left(p_L\beta_H-p_H\beta_L\right)\\ +\left(1-\xi(\tau)\right)\beta_H\Delta p\end{array}\right]\geq\left[\begin{array}{c}\left(1-\xi(\tau)\right)p_L\beta_H\Delta p\\ -\left(\Delta p+\xi(\tau)p_L\right)\left(p_L\beta_H-p_H\beta_L\right)\end{array}\right]\\ &\iff \varepsilon\geq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{\left(1-\xi(\tau)\right)\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}\end{aligned}$$

Moreover, when $\varepsilon\geq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}$, one can easily check that :

$$\mu_{HL}=\mu_{HL}^{1}\geq\mu_{LH}$$

and when $\varepsilon\leq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}$

$$\mu_{HL}=\mu_{HL}^{2}\geq\mu_{LH}$$

$$\alpha_{LL} + \mu_{LL}\left[p_L\beta_L R - I\right] \geq \alpha_{HL} + \mu_{HL}\left[p_L\beta_L R - I\right] = \alpha_{HL} + \mu_{HL}\left[p_L\beta_H R - I\right] - \mu_{HL} R p_L \Delta\beta \tag{$IC_{LL\to HL}$}$$

$$\Leftrightarrow \mu_{HL} R p_L \Delta\beta \geq (1-\xi(\tau))\left\{\mu_{LH} R\left[p_L\beta_H - p_H\beta_L - \varepsilon\left(\beta_L + \beta_H\right)\right] + B\right\} - \xi(\tau)F$$

$$\Leftrightarrow \mu_{HL} \geq \frac{(1-\xi(\tau))\left(p_L - \varepsilon\right)\Delta\beta}{p_L\Delta\beta}\left(\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{R\beta_L\left[(p_H+\varepsilon) - (p_L-\varepsilon)\right]}\right)$$

$$\Leftrightarrow \mu_{HL} \geq \frac{(1-\xi(\tau))\left(p_L - \varepsilon\right)\Delta\beta}{p_L\Delta\beta}\mu_{LH}$$

Since $(1-\xi(\tau))\, p_L\Delta\beta \leq p_L\Delta\beta$ and since $\mu_{HH} \geq \mu_{LH}$, $IC_{LL\to HL}$ is also satisfied. Finally, we get

$$\begin{aligned} \mu_{HL} &= \max\left\{\mu^1_{HL}; \mu^2_{HL}\right\} \\ &= \begin{cases} \mu^1_{HL} \text{ if } \varepsilon \geq \frac{\beta_L\Delta p - \xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{ib} \\ \mu^2_{HL} \text{ if } \varepsilon \leq \frac{\beta_L\Delta p - \xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{ib} \end{cases} \end{aligned}$$

$$\alpha_{LL} + \mu_{LL}\left[p_L\beta_L R - I\right] \geq (1-\xi(\tau))\left\{\alpha_{HH} + \mu_{HH}\left[\left(p_L-\varepsilon\right)\beta_L R - I\right] + B\right\} + \xi(\tau)\left(w - F\right) = \tag{$IC_{LL\to HH}$}$$

$$(1-\xi(\tau))\left\{\alpha_{HH} + \mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R - I\right] - \mu_{HH} R \begin{bmatrix} \left(p_H+\varepsilon\right)\beta_H \\ -\left(p_L-\varepsilon\right)\beta_L \end{bmatrix} + B\right\} + \xi(\tau)\left(w-F\right)$$

$$\Leftrightarrow \mu_{HH} \geq \left\{\begin{array}{c} (1-\xi(\tau))\left\{\begin{array}{c} \mu_{LH} R \begin{bmatrix} p_L\beta_H - p_H\beta_L \\ -\varepsilon\left(\beta_L+\beta_H\right) \end{bmatrix} \\ +B \end{array}\right\} \\ -\xi(\tau)F + \mu_{HL} R\beta_H \Delta p \end{array}\right\}\left(\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{R\left[(p_H+\varepsilon)\beta_H - (p_L-\varepsilon)\beta_L\right]}\right)$$

$$= \mu^1_{HH}$$

$$\alpha_{HL}+\mu_{HL}\left[p_L\beta_H R-I\right]\geq(1-\xi(\tau))\left\{\alpha_{HH}+\mu_{HH}\left[\left(p_L-\varepsilon\right)\beta_H R-I\right]+B\right\}+\xi(\tau)\left(w-F\right) \tag{$IC_{HL\to HH}$}$$

$$=(1-\xi(\tau))\left\{\begin{array}{c}(1-\xi(\tau))\left\{\begin{array}{c}w+\mu_{LL}R\beta_L\Delta p\\ +\mu_{LH}R\left[\begin{array}{c}p_L\beta_H-p_H\beta_L\\ -\varepsilon\left(\beta_L+\beta_H\right)\end{array}\right]+B\end{array}\right\}\\ +\xi(\tau)\left(w-F\right)+\mu_{HL}R\beta_H\Delta p-\mu_{HH}R\beta_H\left[\Delta p+2\varepsilon\right]+B\end{array}\right\}+\xi(\tau)\left(w-F\right)$$

$$\Leftrightarrow\mu_{HH}\geq\frac{\begin{array}{c}(1-\xi(\tau))\,\mu_{HL}R\beta_H\Delta p-\xi(\tau)\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\\ +(1-\xi(\tau))\,B-\xi(\tau)F\end{array}}{R\beta_H\left[\Delta p+2\varepsilon\right]}=\mu_{HH}^2$$

$$\alpha_{LH}+\mu_{LH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]+B\geq\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_L R-I\right]+B \tag{$IC_{LH\to HH}$}$$

$$=\alpha_{HH}+\mu_{HH}\left[\left(p_H+\varepsilon\right)\beta_H R-I\right]+B-\mu_{HH}R\left(p_H+\varepsilon\right)\Delta\beta$$

$$\Leftrightarrow\mu_{HH}\geq\frac{\begin{array}{c}(1-\xi(\tau))\,\mu_{LH}R\left[p_L\beta_H-p_H\beta_L-\varepsilon\left(\beta_L+\beta_H\right)\right]\\ +\mu_{HL}R\beta_H\Delta p\end{array}+(1-\xi(\tau))\,B-\xi(\tau)F}{R\left(p_H+\varepsilon\right)\Delta\beta}=\mu_{HH}^3$$

We thus have:

$$\begin{aligned}
\mu_{LL}&=0\\
\mu_{LH}&=\frac{B-\frac{\xi(\tau)}{(1-\xi(\tau))}F}{R\beta_L\left[\left(p_H+\varepsilon\right)-\left(p_L-\varepsilon\right)\right]}\\
\mu_{HL}&=\left\{\begin{array}{l}\mu_{HL}^1\text{ if }\varepsilon\geq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{ib}\\ \mu_{HL}^2\text{ if }\varepsilon\leq\frac{\beta_L\Delta p-\xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta+\frac{p_L}{p_H}\beta_H-\beta_L}=\varepsilon_{ib}\end{array}\right.\\
\mu_{HH}&\geq\max\left\{\mu_{HL};\mu_{HH}^1;\mu_{HH}^2;\mu_{HH}^3\right\}
\end{aligned}$$

$$U_{LL}=\alpha_{LL}+\mu_{LL}\left[p_L\beta_L R-I\right]=w$$

$$U_{LH} = \alpha_{LH} + \mu_{LH}\left[(p_H+\varepsilon)\beta_L R - I\right] = w + \mu_{LL} R\beta_L \Delta p = w$$

$$\begin{aligned}
U_{HL} &= \alpha_{HL} + \mu_{HL}\left[p_L \beta_H R - I\right] \\
&= (1-\xi(\tau))\left\{\begin{array}{c} w + \mu_{LL} R\beta_L \Delta p \\ +\mu_{LH} R\left[p_L\beta_H - p_H\beta_L - \varepsilon\left(\beta_L+\beta_H\right)\right] + B \end{array}\right\} + \xi(\tau)(w-F) \\
&= w + (1-\xi(\tau))(p_L-\varepsilon)\Delta\beta\left(\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}\right)
\end{aligned}$$

$$U_{HH} = \alpha_{HH} + \mu_{HH}\left[(p_H+\varepsilon)\beta_H R - I\right] = \left[\begin{array}{c} (1-\xi(\tau))\left\{\begin{array}{c} w + \mu_{LL} R\beta_L \Delta p \\ +\mu_{LH} R\left[\begin{array}{c} p_L\beta_H - p_H\beta_L \\ -\varepsilon\left(\beta_L+\beta_H\right) \end{array}\right] \\ +B \end{array}\right\} \\ +\xi(\tau)(w-F) + \mu_{HL} R\beta_H \Delta p \end{array}\right]$$

$$= \left\{\begin{array}{c} w + \frac{(p_L-\varepsilon)p_H(\Delta\beta)^2\left[(1-\xi(\tau))B - \xi(\tau)F\right]}{\beta_L[\Delta p + 2\varepsilon][p_L\beta_H - p_H\beta_L]} \text{ if } \varepsilon \leq \frac{\beta_L\Delta p - \xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{ib} \\ w + \frac{(p_H+\varepsilon)\Delta\beta\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L[(p_H+\varepsilon)-(p_L-\varepsilon)]} \text{ if } \varepsilon \geq \frac{\beta_L\Delta p - \xi(\tau)p_L\Delta\beta}{(1-\xi(\tau))\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} = \varepsilon_{ib} \end{array}\right.$$

To sum up, here are the CEO utilities when there is a monitoring board:

$$\begin{aligned}
U_{LL} &= w \\
U_{LH} &= w \\
U_{HL} &= w + (1-\xi(\tau))(p_L-\varepsilon)\Delta\beta\left(\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}\right) \\
U_{HH} &= \left\{\begin{array}{c} w + \frac{(p_L-\varepsilon)p_H(\Delta\beta)^2\left[(1-\xi(\tau))B - \xi(\tau)F\right]}{\beta_L[\Delta p + 2\varepsilon][p_L\beta_H - p_H\beta_L]} \text{ if } \varepsilon \leq \varepsilon_{ib} \\ w + \frac{(p_H+\varepsilon)\Delta\beta\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L[(p_H+\varepsilon)-(p_L-\varepsilon)]} \text{ if } \varepsilon \geq \varepsilon_{ib} \end{array}\right.
\end{aligned}$$

We can verify now that we have

$$U_{LL} \leq U_{LH} \leq U_{HL} \leq U_{HH}$$

When $\varepsilon \leq \varepsilon_{ib}$, we need to see if $\frac{p_H \Delta\beta}{p_L\beta_H - p_H\beta_L} \geq 1$, which is true since $p_L\beta_H - p_H\beta_L = p_H\Delta\beta - \beta_H\Delta p$. Subsequently, we have $U_{HL} \leq U_{HH}$. When $\varepsilon \geq \varepsilon_{ib}$, since $(1-\xi(\tau))(p_L - \varepsilon) \leq p_H + \varepsilon$, we also have $U_{HL} \leq U_{HH}$.

One can remark that types (HL) and (HH) informational rents are lower with monitoring than without.

$$\begin{aligned}
U_{HLib} \leq U_{HLnb} &\iff (1-\xi(\tau))(p_L-\varepsilon)\Delta\beta\left(\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]}\right) \leq \frac{B(p_L-\varepsilon)\Delta\beta}{\beta_L[\Delta p + 2\varepsilon]} \\
&\iff (1-\xi(\tau))\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right] \leq B \text{ , which is true}
\end{aligned}$$

Moreover, we can prove that $\varepsilon_{ib} \leq \varepsilon_{nb}$. Indeed,

$$\begin{gathered}
\varepsilon_{ib} - \varepsilon_{nb} \leq 0 \\
\iff \frac{\beta_L\Delta p - \xi p_L\Delta\beta}{(1-\xi)\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} - \frac{\beta_L\Delta p}{\Delta\beta + \frac{p_L}{p_H}\beta_H - \beta_L} \leq 0 \\
\iff 2\xi\Delta\beta\left[p_H\beta_L - p_L\beta_H\right] \leq 0
\end{gathered}$$

which is true since we have $\beta_L p_H - p_L\beta_H \leq 0$.

This implies that we only have three possible cases to consider for U_{HH}

1. When $\varepsilon \leq \varepsilon_{ib}$

$$\begin{aligned}
U_{HHib} - U_{HHnb} &= \frac{(p_L-\varepsilon)p_H(\Delta\beta)^2\left[(1-\xi(\tau))B - \xi(\tau)F\right]}{\beta_L[\Delta p + 2\varepsilon][p_L\beta_H - p_H\beta_L]} - \frac{B(p_L-\varepsilon)p_H(\Delta\beta)^2}{\beta_L[\Delta p+2\varepsilon][p_L\beta_H - p_H\beta_L]} \\
sign(U_{HHib} - U_{HHnb}) &= sign(-\xi(\tau)(B+F)(p_L-\varepsilon)(\Delta\beta)^2) \leq 0
\end{aligned}$$

2. When $\varepsilon_{ib} \leq \varepsilon \leq \varepsilon_{nb}$

$$\begin{aligned} U_{HHib} - U_{HHnb} &= \frac{(p_H+\varepsilon)\,\Delta\beta\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]} - \frac{B\,(p_L-\varepsilon)\,p_H\,(\Delta\beta)^2}{\beta_L\left[\Delta p + 2\varepsilon\right]\left[p_L\beta_H - p_H\beta_L\right]} \\ sign(U_{HHib} - U_{HHnb}) &= sign\left[(p_H+\varepsilon)\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]\left[p_L\beta_H - p_H\beta_L\right] - B\,(p_L-\varepsilon)\,p_H\Delta\beta\right] \end{aligned}$$

Since $B \geq \left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]$ we need to prove that $(p_L-\varepsilon)\,p_H\Delta\beta \geq (p_H+\varepsilon)\,(p_L\beta_H - p_H\beta_L)$

$$\begin{aligned} (p_L-\varepsilon)\,p_H\Delta\beta - (p_H+\varepsilon)\,(p_L\beta_H - p_H\beta_L) &= \begin{bmatrix} p_Lp_H\Delta\beta - p_Hp_L\beta_H + p_Hp_H\beta_L \\ -\varepsilon\,(p_H\Delta\beta - \beta_Hp_L + p_H\beta_L) \end{bmatrix} \\ &= p_H\left[\beta_L\Delta p - \varepsilon\left(\Delta\beta - \frac{p_L}{p_H}\beta_H + \beta_L\right)\right] \end{aligned}$$

Since $\varepsilon \leq \varepsilon_{nb}$, we have $\beta_L\Delta p - \varepsilon\left(\Delta\beta - \frac{p_L}{p_H}\beta_H + \beta_L\right) \geq 0$

3. When $\varepsilon_{ib} \leq \varepsilon_{nb} \leq \varepsilon$

$$U_{HHib} - U_{HHnb} = \frac{(p_H+\varepsilon)\,\Delta\beta\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L\left[(p_H+\varepsilon)-(p_L-\varepsilon)\right]} - \frac{B\Delta\beta\,(p_H+\varepsilon)}{\beta_L\left[\Delta p + 2\varepsilon\right]}$$

Since $B \geq \left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]$ we have $U_{HHib} \leq U_{HHnb}$.

We can now calculate the income of the shareholders. There are two cases to consider. When $\varepsilon \leq \varepsilon_{ib}$,

$$W_M = E(\pi) - w - w_0 - (1-\gamma)(1-\xi(\tau))(p_L-\varepsilon)\Delta\beta\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)}\left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_Lp_H}\right]$$

When $\varepsilon \geq \varepsilon_{ib}$,

$$W_M = E(\pi) - w - w_0 - (1-\gamma)\Delta\beta\frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)}\left[\nu(1-\xi(\tau))(p_L-\varepsilon) + (1-\nu)(p_H+\varepsilon)\right]$$

■

Proof of Proposition 5. We have to find for which values of τ, the contract is collusion proof.

1. $\varepsilon \leq \varepsilon_{ib}$

$$\begin{aligned} W_{CP} - W_{CF} &= -(1-\gamma)(1-\xi(\tau))(p_L-\varepsilon)\Delta\beta \frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \begin{bmatrix} \nu(1-\xi(\tau)) \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{bmatrix} \\ &\quad +(1-\gamma)\left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right]\frac{B\Delta\beta(p_L-\varepsilon)}{\beta_L(\Delta p + 2\varepsilon)} \geq 0 \text{ for all } \tau \in [\tau_{\min}, \tau_0] \end{aligned}$$

with $\tau_0 = \frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}$ for $\varepsilon \leq \varepsilon_{ib}$. Indeed,

$$W_{CP} - W_{CF} = \frac{(1-\gamma)(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)} \left(\begin{array}{c} \left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right] B \\ -\left[(1-\xi(\tau))B - \xi(\tau)F\right] \begin{bmatrix} \nu(1-\xi(\tau)) \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{bmatrix} \end{array} \right)$$

As, we have $\xi(\tau)^{ED}(\tau) = \frac{1}{\tau}$, this gives

$$W_{CP} - W_{CF} = \frac{(1-\gamma)(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)} \left(\begin{array}{c} \left[\left(\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right)(B+F) + \nu B\right]\tau^2 \\ -\nu\left[B + F + B\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right]\tau \\ +\nu\,(B+F)\,\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{array} \right)$$

This polynomial in τ with a positive second degree term has 2 positive roots. If those roots are both lower than τ_0, then, $W_{CP} - W_{CF} \geq 0$ for all $\tau \in [\tau_{\min}, \tau_0]$. The lowest root is

$$\tau_1 = \frac{\nu\left[B+F+B\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right] - \sqrt{\begin{array}{c}\nu^2\left[B+F+B\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right]^2 \\ -4\nu\left(B+F\right)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\left[\begin{array}{c}\left(\begin{array}{c}\nu \\ +(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\end{array}\right)(B+F) \\ +\nu B\end{array}\right]\end{array}}}{2\left[\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)(B+F)+\nu B\right]}$$

We have

$$\begin{array}{c} \tau_1 \geq \tau_0 \\ \Longleftrightarrow \left(\begin{array}{c} 4\left(\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)^2\left[\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)(B+F)+\nu B\right]^2 \\ -4\nu\left(B+F+B\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\left[\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)(B+F)+\nu B\right] \\ -4\nu\left(B+F\right)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\left[\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)(B+F)+\nu B\right] \end{array}\right) \geq 0 \\ \Longleftrightarrow 4\left(\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)^2\left(\begin{array}{c}\left[\begin{array}{c}\left(\begin{array}{c}\nu \\ +(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\end{array}\right)(B+F) \\ +\nu B\end{array}\right] \\ * \\ \left[\left(\begin{array}{c}\nu \\ +(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\end{array}\right)(B+F)\right]\end{array}\right) \geq 0 \end{array}$$

which is true. $W_{CP} - W_{CF}$ is therefore positive for all $\tau \in [\tau_{\min}, \tau_0]$. The optimal contract is the collusion proof contract for $\varepsilon \leq \varepsilon_{ib}$.

2. $\varepsilon_{ib} \leq \varepsilon \leq \varepsilon_{nb}$

$$
\begin{aligned}
W_{CP} - W_{CF} &= -(1-\gamma)\Delta\beta \frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\nu(1-\xi(\tau))^2(p_L - \varepsilon) + (1 - \nu + \frac{\xi(\tau)\nu}{\tau})(p_H + \varepsilon)\right] \\
&\quad +(1-\gamma)\left[\nu + (1-\nu)\frac{p_H \Delta\beta}{p_L \beta_H - \beta_L p_H}\right] \frac{B\Delta\beta(p_L - \varepsilon)}{\beta_L(\Delta p + 2\varepsilon)} \geq 0 \text{ for all } \tau \in [\tau_{\min}, \tau_0]
\end{aligned}
$$

with $\tau_0 = \frac{p_H+\varepsilon}{p_L-\varepsilon} + 1$ if $\varepsilon \geq \varepsilon_{ib}$. Indeed,

$$
W_{CP} - W_{CF} = \frac{(1-\gamma)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)} \left(\begin{array}{c} \left[\begin{array}{c} \nu \\ +(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} \end{array}\right] B(p_L - \varepsilon) \\ + \left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right] \left[\begin{array}{c} \nu(1-\xi(\tau))^2(p_L - \varepsilon) \\ +(1 - \nu + \frac{\xi(\tau)\nu}{\tau})(p_H + \varepsilon) \end{array}\right] \end{array} \right)
$$

As, we have $\xi^{ED}(\tau) = \frac{1}{\tau}$, this gives

$$
W_{CP} - W_{CF} = \frac{(1-\gamma)\Delta\beta(p_L - \varepsilon)}{\beta_L(\Delta p + 2\varepsilon)(\tau - 1)} \left(\begin{array}{c} \left[(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} - (1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] B\tau^3 \\ \left[\begin{array}{c} -\left(\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right) B + 2\nu B \\ +\left(\nu + (1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)(B+F) \end{array}\right] \tau^2 \\ -\nu\left[B\left(1 + \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) + 2\nu(B+F)\right]\tau \\ +\nu(B+F)\left(1 + \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) \end{array} \right)
$$

We are now able to show that this degree 3 polynomial, denote it $P(\tau)$, is negative for all $\tau \in [\tau_{\min}, \tau_0]$. Indeed

$$
\frac{\partial P(\tau)}{\partial \tau} = \left(\begin{array}{c} 3\tau^2 B(1-\nu)\left[\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H} - \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] \\ +2\tau\left[-\left(\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right) B + 2\nu B + \left(\nu + (1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)(B+F)\right] \\ -\left[\nu B\left(1 + \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) + 2\nu(B+F)\right] \end{array} \right)
$$

Moreover, as

$$\varepsilon \geq \varepsilon_{ib} \Longleftrightarrow \frac{\tau\left(\beta_L p_H(\tau_0-2)-\varepsilon\beta_H\tau_0\right.}{p_H\Delta\beta} \leq 1$$

we have

$$\begin{aligned} & \tau^2 B(1-\nu)\left[\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] \\ = \ & \tau B(1-\nu)\left(\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)\left(\frac{\tau\left(\beta_L p_H(\tau_0-2)-\varepsilon\beta_H\tau_0\right.}{p_H\Delta\beta}\right) \leq \tau B(1-\nu)\left(\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right) \end{aligned}$$

and thus

$$\begin{aligned} \frac{\partial P(\tau)}{\partial\tau} \quad & \leq \quad \left(\begin{array}{c} 3\tau B(1-\nu)\left(\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right) \\ +2\tau\left[\begin{array}{c} -\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)B \\ +2\nu B+(1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\left(\tau_0-1\right)(B+F) \end{array}\right] \\ -\left[\nu B\tau_0+2\nu\left(B+F\right)\right] \end{array}\right) \leq 0 \\ \Longleftrightarrow \quad & \tau \leq \frac{\nu B\tau_0+2\nu\left(B+F\right)}{\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)B+\nu B+\left(\nu+(1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)(B+F)} \end{aligned}$$

Moreover,

$$\frac{\nu B\tau_0+2\nu\left(B+F\right)}{\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)B+\nu B+\left(\nu+(1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)(B+F)} \leq \tau_0$$

Hence, $\frac{\partial P(\tau)}{\partial\tau}$ is negative for all $\tau \in \left[\tau_{\min},\tau_0\right]$. Finally, we will show that $\left(W_{CP}-W_{CF}\right)(\tau_0) \geq$

0

$$
\begin{gathered}
\left(W_{CP}-W_{CF}\right)(\tau_0)\geq 0 \Longleftrightarrow \\
\frac{(1-\gamma)\Delta\beta(p_L-\varepsilon)}{\beta_L(\Delta p+2\varepsilon)\tau_0}\left(\begin{array}{c}
\left[(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}-(1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right]B\tau_0^3 \\
\left[\begin{array}{c}
-\left(\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right)B+2\nu B \\
+\left(\nu+(1-\nu)\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)(B+F)
\end{array}\right]\tau_0^2 \\
-\left[\nu B\tau_0+2\nu\left(B+F\right)\right]\tau_0+\nu\left(B+F\right)\tau_0
\end{array}\right)\geq 0 \\
\Longleftrightarrow\left(\begin{array}{c}
B(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\tau_0\left(\tau_0-1\right)\left[\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right] \\
+(1-\nu)\tau_0\left(\tau_0-1\right)F+\nu\left(\tau_0-1\right)\left(B+F\right)
\end{array}\right)\geq 0
\end{gathered}
$$

However, as

$$
\begin{aligned}
\varepsilon &\leq \varepsilon_{nb}\Longleftrightarrow \\
\frac{\beta_H\Delta p}{p_L\beta_H-\beta_L p_H}(p_L-\varepsilon) &\geq \frac{p_H\Delta\beta p_L+\left[p_L\beta_H-\beta_L p_H\right]p_L-p_L\beta_H\Delta p}{\left[p_L\beta_H-\beta_L p_H\right]}-(p_L-\varepsilon)
\end{aligned}
$$

we have, together with $p_H\Delta\beta\geq\beta_H\Delta p$

$$
\begin{aligned}
\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)} &\geq \frac{\beta_H\Delta p}{p_L\beta_H-\beta_L p_H}-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)} \\
&\geq \frac{p_H\Delta\beta p_L+\left[p_L\beta_H-\beta_L p_H\right]p_L-p_L\beta_H\Delta p}{\left[p_L\beta_H-\beta_L p_H\right](p_L-\varepsilon)}-1-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\geq 0
\end{aligned}
$$

As $\left[\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right]\geq 0$, $\left(W_{CP}-W_{CF}\right)(\tau_0)$ is thus positive and as $\frac{\partial P(\tau)}{\partial\tau}$ is negative for all $\tau\in\left[\tau_{\min},\tau_0\right]$, $W_{CP}-W_{CF}$ is therefore positive for all $\tau\in\left[\tau_{\min},\tau_0\right]$. The optimal contract is the collusion proof contract for $\varepsilon_{ib}\leq\varepsilon\leq\varepsilon_{nb}$.

3. $\varepsilon_{ib} \leq \varepsilon_{nb} \leq \varepsilon$

$$\begin{aligned} W_{CP} - W_{CF} &= -(1-\gamma)\Delta\beta \frac{\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right]}{\beta_L(\Delta p + 2\varepsilon)} \left[\begin{array}{c} \nu(1-\xi(\tau))^2(p_L - \varepsilon) \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})(p_H+\varepsilon) \end{array}\right] \\ &\quad +(1-\gamma)\left[(p_H+\varepsilon) - \nu(\Delta p + 2\varepsilon)\right] \frac{B\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)} \geq 0 \text{ for all } \tau \in [\tau_{\min}, \tau_0] \end{aligned}$$

with $\tau_0 = \frac{p_H+\varepsilon}{p_L-\varepsilon} + 1$ if $\varepsilon \geq \varepsilon_{ib}$. Indeed,

$$W_{CP} - W_{CF} = \frac{(1-\gamma)(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)} \left(\begin{array}{c} \left[\nu + (1-\nu)(\tau_0 - 1)\right] B \\ -\left[B - \frac{\xi(\tau)}{(1-\xi(\tau))}F\right] \left[\begin{array}{c} \nu(1-\xi(\tau))^2 \\ +(1-\nu+\frac{\xi(\tau)\nu}{\tau})(\tau_0-1) \end{array}\right] \end{array} \right)$$

As, we have $\xi^{ED}(\tau) = \frac{1}{\tau}$, this gives

$$W_{CP} - W_{CF} = \frac{(1-\gamma)(p_L-\varepsilon)\Delta\beta}{\beta_L(\Delta p + 2\varepsilon)(\tau - 1)} \left(\begin{array}{c} \left[\nu(2B+F) + (1-\nu)(\tau_0-1)F\right]\tau^2 \\ -\nu\left[B\tau_0 + 2(B+F)\right]\tau \\ +\nu(B+F)\tau_0 \end{array} \right)$$

This polynomial in τ with a positive second degree term has 2 positive roots. If those roots are both lower than τ_0, then, $W_{CP} - W_{CF} \geq 0$ for all $\tau \in [\tau_{\min}, \tau_0]$. The lowest root is

$$\tau_2 = \frac{\nu\left[B\tau_0 + 2(B+F)\right] - \sqrt{\begin{array}{c} \nu^2\left[B\tau_0 + 2(B+F)\right]^2 \\ -4\nu(B+F)\tau_0\left[\nu(2B+F) + (1-\nu)(\tau_0-1)F\right] \end{array}}}{2\left[\nu(2B+F) + (1-\nu)(\tau_0-1)F\right]}$$

We have

$$\tau_2 \geq \tau_0$$

$$\Longleftrightarrow 4\tau_0^2 \begin{bmatrix} \nu\left(2B+F\right) \\ +(1-\nu)\left(\tau_0-1\right)F \end{bmatrix}^2 - 4\nu\tau_0\left[B\tau_0+(B+F)\right] \begin{bmatrix} \nu\left(2B+F\right) \\ +(1-\nu)\left(\tau_0-1\right)F \end{bmatrix} \geq 0$$

$$\Longleftrightarrow 4\tau_0\left(\tau_0-1\right)\left[\nu\left(2B+F\right)+(1-\nu)\left(\tau_0-1\right)F\right]\left[\nu\left(B+F\right)+(1-\nu)\tau_0 F\right] \geq 0$$

which is true. $W_{CP}-W_{CF}$ is therefore positive for all $\tau \in [\tau_{\min}, \tau_0]$. The optimal contract is the collusion proof contract for $\varepsilon_{ib} \leq \varepsilon_{nb} \leq \varepsilon$. ■

Proof of Proposition 6. When $\varepsilon \leq \widehat{\varepsilon}$ and $\widehat{\tau} \leq \tau_0$, we have

$$W_{CB}(\tau) = \begin{cases} -\left[B-\frac{1}{\tau-1}F\right] \begin{bmatrix} \nu(\frac{\tau-1}{\tau})^2(p_L-\varepsilon) \\ +((1-\nu)+\frac{\nu}{\tau^2})(p_H+\varepsilon) \end{bmatrix} & \text{if } \tau \leq \widehat{\tau} \\ -\left[B-\frac{1}{\tau-1}F\right] \begin{bmatrix} \nu(\frac{\tau-1}{\tau})^2 \\ +\left[(1-\nu)\left(\frac{\tau-1}{\tau}\right)+\nu\frac{(\tau-1)}{\tau^3}\right]\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H} \end{bmatrix} & \text{if } \widehat{\tau} \leq \tau \leq \tau_0 \\ -\frac{1}{\tau}\left[\left(\tau-1\right)B-F\right]\left[\nu+(1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right] & \text{if } \tau \geq \tau_0 \end{cases},$$

When $\varepsilon \geq \widehat{\varepsilon}$ or $\widehat{\tau} \geq \tau_0$, we have:

$$W_{CB}(\tau) = \begin{cases} -\frac{1}{(\tau-1)\tau^2}\left[\left(\tau-1\right)B-F\right] \begin{bmatrix} \nu\left(\tau-1\right)^2\left(p_L-\varepsilon\right) \\ +((1-\nu)\,\tau^2+\nu)(p_H+\varepsilon) \end{bmatrix} & \text{if } \tau \leq \tau_0 \\ -\frac{1}{\tau}\left[\left(\tau-1\right)B-F\right]\left[\nu(\tau-1)(p_L-\varepsilon)+(1-\nu)(p_H+\varepsilon)\right] & \text{if } \tau \geq \tau_0 \end{cases},$$

Assume first $\varepsilon \leq \widehat{\varepsilon}$ and $\widehat{\tau} \leq \tau_0$. As F has to be set as high as possible, we have $\left[B-\frac{1}{(\tau-1)}F\right] =$

K, due to the CEO's limited liability constraint. We thus have, if $\tau \leq \widehat{\tau}$

$$\begin{aligned}\frac{\partial W_{CB}(\tau)}{\partial \tau} &= -\left(K\left[\nu(p_L-\varepsilon)\left(\frac{2\tau^2(\tau-1)-2\tau(\tau-1)^2}{\tau^4}\right)-\frac{2\nu}{\tau^3}(p_H+\varepsilon)\right]\right) \\ &= -\left(2K\nu\frac{[(p_L-\varepsilon)(\tau-1)-(p_H+\varepsilon)]}{\tau^3}\right)\end{aligned}$$

However,

$$(p_L-\varepsilon)(\tau-1)-(p_H+\varepsilon) \leq 0$$

as $\tau \leq \tau_0 = 1+\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}$. And then $\frac{\partial W_{CB}(\tau)}{\partial \tau} \geq 0$. If $\widehat{\tau} \leq \tau \leq \tau_0$,

$$W_{CB}(\tau) = -K\left[\begin{array}{c}\nu(\frac{\tau-1}{\tau})^2 \\ +\left[(1-\nu)(\frac{\tau-1}{\tau})+\nu\frac{(\tau-1)}{\tau^3}\right]\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\end{array}\right]$$

The first derivative of this objective function is in this case:

$$\begin{aligned}\frac{\partial W_{CB}(\tau)}{\partial \tau} &= -\left(K\left[\nu\left(\frac{2(\tau-1)}{\tau^3}\right)+\left[\frac{(1-\nu)}{\tau^2}+\nu\frac{3-2\tau}{\tau^4}\right]\frac{p_H\Delta\beta}{p_L\beta_H-\beta_L p_H}\right]\right) \\ &= -\frac{K}{\tau^4}\left[2\nu\tau(\tau-1)+(1-\nu)\tau^2\tau_0+\nu(3-2\tau)\tau_0\right] \\ &= -\frac{K}{\tau^4}\left[2\nu\tau^2-2\nu\tau+(1-\nu)\tau^2\tau_0+3\nu\tau_0-2\nu\tau\tau_0\right] \\ &= -K\left[\frac{2\nu}{\tau^2}-\frac{2\nu}{\tau^3}+\frac{(1-\nu)\tau_0}{\tau^2}+\frac{3\nu\tau_0}{\tau^4}-\frac{2\nu\tau_0}{\tau^3}\right] \\ &= \frac{K}{\tau^4}\left[-\tau^2(2\nu+(1-\nu)\tau_0)+2\nu\tau(1+\tau_0)-3\nu\tau_0\right]\end{aligned}$$

The sign of this expression is equivalent to the sign of a second degree concave polynomial in τ. This polynomial has two positive roots. We will show below that it is negative in $\widehat{\tau}$ and τ_0 and that its derivative in $\widehat{\tau}$ and τ_0 is also negative. This implies that it is negative for all

τ in $[\widehat{\tau}, \tau_0]$ and that consequently $W_{CB}(\tau)$ is non increasing on this interval. Indeed, we have

$$\begin{aligned}
\frac{\partial W_{CB}(\tau)}{\partial \tau}|\tau_0 &= -\frac{K}{\tau_0^4}\left[(1-\nu)\,\tau_0^3 + \nu\tau_0\right] \leq 0 \\
&= -K\left[\frac{(1-\nu)}{\tau_0} + \frac{\nu}{\tau_0^3}\right] \leq 0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial W_{CB}(\tau)}{\partial \tau}|\widehat{\tau} &= -\frac{K}{\widehat{\tau}^2}\left[2\nu\left(1-\frac{1}{\frac{\tau_0}{\tau_0-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}}}\right) + (1-\nu)\,\tau_0 + \frac{3\nu\tau_0}{\left(\frac{\tau_0}{\tau_0-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}}\right)^2} - \frac{2\nu\tau_0}{\left(\frac{\tau_0}{\tau_0-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}}\right)}\right] \\
&= -\frac{K}{\widehat{\tau}^2}\left[\begin{array}{c} 2\nu\left(\frac{\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}}{\tau_0}\right) + (1-\nu)\,\tau_0 \\ +\frac{3\nu\left(\tau_0-\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)^2}{\tau_0} - 2\nu\left(\tau_0 - \frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) \end{array}\right] \\
&= \frac{K}{\widehat{\tau}^2}\left[-\tau_0^2 + 4\nu\tau_0\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)} - 2\nu\left(\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) - 3\nu\left(\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)^2\right]
\end{aligned}$$

This is always negative as $\Delta = 16\nu^2\left(\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right)^2 - 4\nu\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\left(2+3\frac{(p_H+\varepsilon)}{(p_L-\varepsilon)}\right) \leq 0$ because $\nu \leq \frac{1}{2}$. Moreover, it is easy to check that the derivative of the second degree concave polynomial in τ (having the same sign as $\frac{\partial W_{CB}(\tau)}{\partial \tau}$) is negative in τ_0 and in $\widehat{\tau}$ (because $\nu \leq \frac{1}{2}$). This implies that

$$\frac{\partial W_{CB}(\tau)}{\partial \tau} \leq 0 \text{ for all } \tau \in [\widehat{\tau}, \tau_0]\,.$$

If $\tau \geq \tau_0$,

$$\begin{aligned}
\frac{\partial W_{CB}(\tau)}{\partial \tau} &= -\frac{(\tau-1)}{\tau}K\left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right] \\
&= -\frac{1}{\tau^2}K\left[\nu + (1-\nu)\frac{p_H\Delta\beta}{p_L\beta_H - \beta_L p_H}\right] \leq 0
\end{aligned}$$

When $\varepsilon \geq \widehat{\varepsilon}$, or $\widehat{\tau} \geq \tau_0$ we have, if $\tau \leq \tau_0$

$$\frac{\partial W_{CB}(\tau)}{\partial \tau} = -\left(K\frac{2\nu}{\tau^3}\left[(\tau-1)(p_L-\varepsilon)-(p_H+\varepsilon)\right]\right) \geq 0$$

If $\tau \geq \tau_0$

$$W_{CB}(\tau) = -K\left[\nu\frac{(\tau-1)^2}{\tau}(p_L-\varepsilon)+(1-\nu)\frac{\tau-1}{\tau}(p_H+\varepsilon)\right]$$

$$\begin{aligned}\frac{\partial W_{CB}(\tau)}{\partial \tau} &= -K\left[\nu\frac{2\tau(\tau-1)-(\tau-1)^2}{\tau^2}(p_L-\varepsilon)+(1-\nu)\frac{1}{\tau^2}(p_H+\varepsilon)\right] \\ &= -K\left[\nu(\tau-1)\frac{\tau+1}{\tau^2}(p_L-\varepsilon)+(1-\nu)\frac{1}{\tau^2}(p_H+\varepsilon)\right] \leq 0\end{aligned}$$

This allows us to conclude that when $\varepsilon \leq \widehat{\varepsilon}$ and $\tau_0 \geq \widehat{\tau}$, it is optimal for the shareholders to select a board of directors with a low intensity of monitoring, i.e. $\tau^* = \widehat{\tau}$. In all other cases, it is optimal for the shareholders to select a board of directors with a high intensity of monitoring, i.e. $\tau^* = \tau_0$. ■

References

[1] Adams, R. B., and D. Ferreira, 2007, "A Theory of Friendly Boards," Journal of Finance 62, 217-250.

[2] Adams, R.B., B.E. Hermalin and M.S. Weisbach, 2010, "The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey," Journal of Economic Literature 48, 58-107.

[3] Bebchuk, L. and M.S. Weisbach, 2010, "The State of Corporate Governance Research," Review of Financial Studies 23, 939-961.

[4] Bizjak, J., M. Lemmon and L. Naveen, 2008, “Has the Use of Peer Groups Contributed to Higher Levels of Executive Compensation?,” Journal of Financial Economics 90, 152-168.

[5] Boone, A. L, L.C. Field, J.M. Karpoff and C.G. Raheja, 2007, “The Determinants of Corporate Board Size and Independence: An Empirical Analysis”, Journal of Financial Economics 85, 65-101.

[6] Cespa G. and G. Cestone, 2007, “Corporate Social Responsibility and Managerial Entrenchment,” Journal of Economics and Management Strategy 16, 741-771.

[7] Chhaochharia, V. and Y. Grinstein, 2009, “CEO Compensation and Board Structure,” The Journal of Finance 64, 231-261.

[8] Cornelli F., Z. Kominek and A. Ljungqvist, 2012, "Monitoring Managers: Does it Matter?," forthcoming, Journal of Finance.

[9] Core, J., R. Holthausen and D. Larcker, 1999, “Corporate Governance, Chief Executive Officer Compensation, and Firm Performance,” Journal of Financial Economics 51, 371-406.

[10] Dahya, J. and J. McConnell, 2007, “Board Composition, Corporate Performance, and the Cadbury Committee Recommendation,” Journal of Financial and Quantitative Analysis 42, 535-564.

[11] Demsetz, H. and K. Lehn, 1985, “The Structure of Corporate Ownership: Causes and consequences,” Journal of Political Economy 33, 3-53.

[12] Faleye, O., R. Hoitash and U. Hoitash, 2011, “The Costs of Intense Board Monitoring,” Journal of Financial Economics, 101, 160-181.

[13] Faure-Grimaud, A., Laffont, J.-J. and D. Martimort, 2003, "Collusion, Delegation, and Supervision with Soft Information," Review of Economic Studies 70, 253-279.

[14] Ferreira, D., M. Ferreira and C. Raposo, 2011, "Board Structure and Price Informativeness," Journal of Financial Economics 99, 523-545.

[15] Harris M., and A. Raviv, 2006, "A Theory of Board Control and Size," Review of Financial Studies 21, 1797-832.

[16] Hermalin, B.E., 2005, "Trends in Corporate Governance," The Journal of Finance 60, 2351–2384.

[17] Hermalin, B.E, and M.S. Weisbach, 1998, "Endogenously Chosen Boards of Directors and Their Monitoring of the CEO," American Economic Review 88, 96-118.

[18] Ittner, C. D., R. A. Lambert and D. F. Larcker, 2003, "The Structure and Performance Consequences of Equity Grants to Employees of New Economy Firms," Journal of Accounting and Economics 34, p. 89-127.

[19] Khalil, F., and J. Lawarrée, 1995, "Collusive Auditors," American Economic Review 85, 442-446.

[20] Kofman, F., and J. Lawarrée, 1993, "Collusion in Hierarchical Agency," Econometrica 61, 629-656.

[21] Kumar, P. and K. Sivaramakrishnan, "Who Monitors the Monitor? The Effect of Board Independence on Executive Compensation and Firm Value," Review of Financial Studies, 21 (3), 1371-1401.

[22] Laffont, J.J., and D. Martimort, 2002, "The Theory of Incentives 1: The Principal-Agent Model," Princeton University Press.

[23] Linck, J.S., J.M. Netter and T. Yang, 2008, "The Determinants of Board Structure," Journal of Financial Economics 87, 308-328..

[24] Murphy, K.J., 1999, "Executive Compensation," Orley Ashenfelter and David Card (eds.), Handbook of Labor Economics 3, North Holland.

[25] Murphy, K.J., 2003, "Stock-based Pay in New-Economy Firms," Journal of Accounting and Economics 34, 129-147.

[26] Raheja, C. G, 2005, "Determinants of Board Size and Composition: A theory of corporate boards," Journal of Financial and Quantitative Analysis 40, 283-306.

[27] Schleifer, A. and R. Vishny, 1997, "A Survey of Corporate Governance," The Journal of Finance 2, 737-783.

[28] Tirole, J., 1986, "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," Journal of Law, Economics, & Organization 2, 181-214.

[29] Tirole, J., 2001, "Corporate Governance," Econometrica 69, 1-35.