Should CEOs be rewarded for the decisions they make? Delegated expertise and CEO pay *

Marco Celentani[†] Rosa Loveira[‡] Pablo Ruiz-Verdú[§]

July 15, 2011

Firms often reward CEOs for making certain strategic decisions, such as acquiring other firms. However, little is known about the optimality of making CEO pay depend on the CEO's decisions: Should CEO pay depend only on firm performance? Or should shareholders reward CEOs for the decisions they make? To address these questions, we propose a model in which shareholders delegate the firm's strategic decisions to the CEO, who must exert costly effort to obtain information about the consequences of the different decisions. The CEO's compensation contract should thus provide incentives to the CEO both to acquire information about the returns of the different decisions and to act on that information optimally. We show that: 1) It is generically optimal to make pay contingent on the decision made by the CEO; 2) The CEO is often rewarded for choosing alternatives that are ex-ante suboptimal. 3) When decisions differ in their complexity, optimal pay-performance sensitivity may be zero if the CEO chooses the complex alternative. Our model highlights novel factors that should be considered in the design of executive compensation contracts, sheds light on existing compensation practices, such as rewarding executives for acquisitions, and suggests mechanisms to promote managerial innovation.

Keywords: Executive compensation, delegated expertise.

JEL Classification Numbers: D86, M52, J33, G30.

*We gratefully acknowledge the financial support of FEDEA and the Spanish Ministry of Innovation and Science for financial support under grant SEJ2008-03516. Rosa Loveira gratefully acknowledges the financial support of Xunta de Galicia (Spain) under the Isidro Parga Pondal research grant. We thank seminar participants at Universidad Carlos III, Universidade de Vigo, Econometric Society European Meeting 2007, 2nd European Reward Management Conference 2009, EASSET 2009, EARIE 2010 and Jornadas de Economía Industrial 2010 for useful discussions and suggestions.

[†]Department of Economics, Universidad Carlos III de Madrid; marco.celentani@uc3m.es

[‡]Departamento de Fundamentos del Análisis Económico e Historia e Instituciones Económicas, Universidade de Vigo; rloveira@uvigo.es.

[§]Department of Business Administration, Universidad Carlos III de Madrid; pablo.ruiz@uc3m.es.

1 Introduction

Firms often reward their CEOs for the decisions they make. For example, CEOs who acquire other firms (Harford and Li, 2007; Grinstein and Hribar, 2004; Bliss and Rosen, 2001) or execute joint ventures, strategic alliances, seasoned equity offerings, or spin-offs (Fich et al., 2010) are paid more than the CEOs of similar firms who do not engage in such deals. Moreover, CEOs are frequently rewarded for acquisitions or other corporate strategy decisions that are associated with reductions in shareholder value. This last fact has prompted several researchers to propose that rewards for acquisitions and similar events are a form of rent extraction by CEOs. However, little is known about the optimality for shareholders of conditioning pay on observable decisions made by the CEO: Should CEO pay depend only on firm performance? Or should shareholders reward CEOs for the decisions they make? If so, how? To address these questions, in this paper we propose a model in which shareholders delegate decision-making to the CEO and in which CEO decisions are observable, so CEO pay can be conditioned on both firm outcomes and the decisions made by the CEO.

By the very nature of the CEO's role, board directors and shareholders can observe most of the key decisions that a CEO makes. Thus, directors and shareholders can observe the markets a firm enters or exits, the firm's choice of capital structure, or whether a firm acquires another firm or spins off a division. Therefore, CEO pay could be made to depend on observable decisions made by the CEO, and not only on the outcome, in terms of the firm's financial performance, of those decisions. However, the problem with conditioning managers' pay on their decisions is that, often, shareholders or board directors do not know which decisions are optimal. Indeed, as Murphy (1999) puts it, "the reason shareholders entrust their money to self-interested CEOs is based on shareholder beliefs that CEOs have superior skills or information in making investment decisions."[1]

Consider for example a board of directors at a firm faced with the opportunity to acquire another firm. Suppose that board directors know that, on average, acquisitions of the sort they are contemplating are associated with a reduction in shareholder value for the acquiring firm. However, board directors also understand that, in some cases, it may be optimal for the firm to carry out such acquisitions. If board directors want to give incentives to the CEO both to thoroughly investigate the consequences of the alternative courses of action for the firm and to implement the optimal decision given the knowledge acquired by the CEO, should they penalize

[1] Murphy (1999), page 2521.

the CEO for making the acquisition? Or should they reward the CEO for making a decision that they believe ex ante to be suboptimal? Alternatively, should they condition the CEO's pay only on firm performance and disregard the information about the decision made by the CEO?

The usual analytical framework to analyze executive compensation is the principal-agent model with hidden action. In this model, the principal knows the action that she would like the agent to take, but cannot make compensation contingent on the action taken by the agent because this action is not observable. The problem for the principal is to design a compensation contract that, by conditioning pay on observables—which typically include some measure of firm performance—provides incentives to the manager to take the right action at the minimum cost for the principal. Certain aspects of the general problem of providing incentives to executives fit this framework well. For example, executives may take actions that are difficult to observe and that reduce firm value, like diverting firm funds for their personal use, or enjoying excessive perquisites (Jensen and Meckling, 1976). In this case, the action that maximizes firm value is known in advance (no fund diversion, no superfluous perquisites), and the problem for shareholders is to design the compensation contract in such a way that the manager refrains from taking (unobservable) suboptimal actions.

The standard hidden action framework is, however, only of limited value to analyze the problem of providing incentives to CEOs to ensure that they make the right strategic decisions. These strategic decisions, such as determining whether to acquire a supplier or merge with a rival, determine firm value to a much larger extent than the manager's enjoyment of excessive private benefits. Strategic decisions are, however, typically observable, so that—unlike in the hidden action model—the problem for shareholders is not how to convince the manager to implement a decision that is known to be optimal, but, rather, how to provide incentives for the manager to learn which decision is optimal, and, then, to effectively make the optimal decision.[2] The main agency problem between shareholders and management, thus, arguably arises not because shareholders do not observe managerial actions but because shareholders do not know which actions CEOs should take in order to maximize firm value. While the standard hidden-action model is well suited to agency problems in which the principal knows what is best, but cannot observe what the agent carries out the task, it is not so well suited to study the moral hazard problem of top executives.

[2]This point has also been made by Murphy (1999): "Unobservable actions cannot be the driving force underlying executive contracts: even if shareholders (or boards of directors) could directly monitor CEO actions, they could not tell whether the actions were appropriate given the circumstances." (Murphy, 1999, page 2521.)

Therefore, in this paper we propose a *delegated expertise* (Demski and Sappington, 1987) model of the role of the CEO. In our model, shareholders delegate the firm's strategic decisions, which we represent as the choice of an investment project, to the CEO, who is uniquely skilled to acquire, process, and interpret information relevant for those decisions. Acquiring and processing information about the profitability of the alternative projects available to the firm is, however, costly for the CEO. After obtaining that information, the CEO chooses one of the projects. Shareholders can observe the CEO's project choice, but do not know which project is optimal nor whether the CEO exerted the necessary effort to acquire information about the profitability of the projects. The problem for shareholders is how to design a compensation contract to provide incentives for the CEO both to exert effort to obtain information and to choose the optimal project based on that information.

Our model yields several novel results about the design of optimal compensation contracts. Given that project choice are contractible, the first question we address is whether the CEO's pay should depend on his choice of project at all or whether pay should depend only on the firm's profits. The usual guide to determine the variables to be included in a contract is the informativeness—or sufficient statistic—principle (Holmstrom, 1979). According to this principle, a variable should be included in an executive compensation contract if it provides information as to whether the executive took the desired action. In our context, the informativeness principle would imply that the variable of interest—the project choice—should be used in the compensation contract if it provides information about the CEO's effort into information acquisition and processing. However, the project is itself chosen by the agent. Thus, even if the choice of certain projects may signal that the CEO put effort into information acquisition, rewarding the CEO for picking those projects may distort the CEO's incentives to choose the optimal project, and these incentives are key in a context in which shareholders do not know which project is optimal. Therefore, to determine whether the CEO's decision should enter (and how) the compensation contract it is necessary to understand the potential trade-off between the insurance motive underlying the informativeness principle and the objective of implementing optimal decisions ex post. We show that it is generically optimal to make the CEO's pay depend on observable decisions (project choice) even if shareholders do not know which decision is optimal. Even if the CEO could be given the right incentives by means of a contract that made his pay contingent on profits only, making pay depend on observable decisions allows the firm to provide those incentives at a lower cost.

Second, we show that the form of the optimal compensation contract crucially depends on two factors. The first factor is the precision with which the informed CEO can predict the outcome of each project. For some projects the information acquired by the CEO may allow him to form precise predictions of the project's outcome, while for other projects the information acquired by the CEO may be only a very noisy signal of the project's outcome. Thus, managerial learning may be symmetric (the informed CEO can predict the outcome of each project with similar precision) or asymmetric (the informed CEO can predict the outcome of a project with much greater precision than the outcome of the other). This learning asymmetry may arise because some projects are more complex or there is less prior knowledge about them that the CEO can draw upon to interpret the information he receives. Alternatively, the CEO may be specialized in evaluating and implementing certain kinds of projects, but may not be able to obtain a significant informational advantage with respect to shareholders regarding other projects. For example, a manager may be a takeover specialist, who is able to determine with great precision the outcome of the acquisition of a supplier firm, yet may not be better equipped than board directors to assess the outcome that would result if the firm did not carry out the acquisition. We show that the form of the optimal contract depends on the symmetry of managerial learning. In particular, when learning is sufficiently asymmetric, compensation is completely insensitive to performance if the CEO selects the project whose outcome is more uncertain. Thus, in some cases, observed pay-performance sensitivity may be low even if the CEO's contract is set optimally. In these cases, the incentives to make the right decision are provided by the performance sensitivity of pay had the alternative decision been made.

The second factor that determines the form of the compensation contract is shareholders' (or board directors) ex ante information about the optimal project. Even if board directors do not know which project is optimal, they need not think that all project are equally likely to be optimal. For example, on the basis of existing empirical research on the effect of acquisitions on the stock price of the acquiring firm, directors may expect that acquisitions (at least for large firms) are likely to be suboptimal, even if they understand that under certain circumstances it may be optimal to undertake an acquisition. We show that the ex ante information about the efficient project determines the form of the compensation contract. In particular, when the precision with which the manager learns about the two projects is similar, the manager is optimally paid more (conditional on performance) if he selects the project that is considered to be ex ante suboptimal. Therefore, our results show that it may be optimal to reward managers for the decisions they make. Moreover, it may be optimal to reward them for making decisions that are ex ante expected to be wrong. It follows from these results that it may be optimal to reward managers for acquiring other firms, even though, on average, acquisitions have been associated with reductions in the wealth of the acquiring firm's shareholders—in particular in the case of large operations (Moeller et al., 2005).

Our paper is related to the theoretical literature on delegated expertise initiated by Lambert (1986) and Demski and Sappington (1987).[3] In delegated expertise models, the principal hires an agent that has the time and skills to acquire expertise about certain decisions. As in our model, the problem for the principal is to provide incentives to the agent to acquire expertise and to use that expertise optimally. However, previous models of delegated expertise assumed that the expert's pay could be conditioned only on performance and not on the expert's actions. Therefore, they are silent about the questions we pose in this article.

Our model also speaks to the empirical literature on executive compensation,[4] especially those articles that have analyzed the relation between CEO pay and corporate strategy decisions. In particular, Bliss and Rosen (2001), Grinstein and Hribar (2004), and Harford and Li (2007) study the impact of mergers and acquisitions on the pay of the acquiring firm's CEO and find that CEOs experience significant wealth increases even if the acquisition appears to reduce shareholder value. Fich et al. (2010) obtain similar results when the corporate strategy decision considered is the execution of joint ventures, strategic alliances, seasoned equity offerings, or spin-offs. Bebchuk and Grinstein (2005) also document that managers benefit from increasing the scale of the firm.

The paper is organized as follows. Section 2 presents the basic model for the case in which the realizations of the possible decisions made by the CEO are independent. In Section 3 we prove that generically there is a gain in making pay conditional not only on performance but also on the CEO's observable decisions. In Sections 4 and 5 we characterize the way in which pay should depend on observable decisions and performance. In Section 6 we consider the case in which the the realizations of the possible decisions made by the CEO are negatively correlated. In Section 7 we consider the case of uninformative priors, i.e., the case in which the two possible decisions are ex-ante equivalent. Section 8 discusses the implications of our results for the determination

[3]See also Barron and Wadell (2003,2008), Gromb and Martimort (2007); Malcomson (2010). Several of these articles interpret the agent's decision as project choice. Therefore, their authors also refer to delegated expertise models as project-selection models (see, e.g. Barron and Wadell, 2008).

[4]See the influential article by Murphy (1999) and Frydman and Jenter (2010) for surveys of the literature on executive compensation.

of CEO pay. Finally, Section 9 summarizes the results and presents some concluding remarks. All proofs are presented in the Appendices.

2 The Model

We consider a firm owned by a risk-neutral owner who delegates to a risk averse manager the tasks of investigating the return distributions of the different projects that the firm may undertake and of deciding which of these projects the firm undertakes.

The firm can undertake one of two possible projects, which we label A and B. Each project can have one of two possible realizations, S, success, or F, failure. Each project determines a probability distribution over returns, and we denote by r_A and r_B the realizations of projects A and B, respectively. It is common knowledge that the two probability distributions are independent and that a is the ex ante probability of success of A, and b is the ex ante probability of success of B. The projects may represent the consequences of entering one of two markets, of acquiring a potential target or not, or of investing in the development of two alternative technologies. Without loss of generality we assume that $a \geq b$, but we will often refer to the generic case in which $a > b$:

Assumption 1 *Decision A is ex ante optimal: $a > b$.*

The manager can either devote sufficient effort to investigate thoroughly the consequences of the two alternative projects ($e = 1$) or not devote sufficient effort ($e = 0$), but the owner cannot observe whether the manager exerts the necessary effort. If the manager exerts no effort, he obtains no additional information about A and B. But if he expends effort, he obtains two independent signals on A and B, respectively α and β. We assume that $\alpha \in \{\overline{\alpha}, \underline{\alpha}\} = \mathcal{A}$ and $\beta \in \left\{\overline{\beta}, \underline{\beta}\right\} = \mathcal{B}$, where $\overline{\alpha}$ is a favorable signal on A, $\underline{\alpha}$ is an unfavorable signal on A, and similarly for β and B.

If the manager expends effort, he obtains the signals and then chooses between A and B. If the manager does not expend effort, he chooses between A and B on the basis of the prior information only. We assume that the manager's project choice is observable and contractible and, in the rest of the paper, we will refer to this choice simply as a *decision*.

The sequence of events is as follows:

1. The owner makes a take-it-or-leave-it contract offer w to the manager. A contract w specifies a salary payment for every possible public history following acceptance of the

contract, where by a public history we refer to a combination of a decision and a realization of return. Thus, a contract has the form $w = (w_{AF}, w_{AS}, w_{BF}, w_{BS}) \in \mathbb{R}^4$.

2. The manager accepts or rejects the contract. If the manager rejects the contract, he obtains reservation utility $\overline{U}$.

3. If the manager accepts the contract, he is hired and chooses whether to exert effort, $e = 1$, or not, $e = 0$.

4. If the manager does not exert effort, he does not receive any signal and makes a publicly observed decision $d \in \{A, B\}$.

5. If the manager exerts effort, he receives signals α and β and then makes a publicly observed decision $d \in \{A, B\}$.

6. The firm's return r_d is realized and publicly observed.

7. Payoffs occur according to the accepted contract and the subsequent public history (d, r_d).

The manager's preferences are described by the Bernoulli utility function $U(w) - g(e)$, where U is twice continuously differentiable with $U'(.) > 0$, $U''(w) < 0$, and $g(1) = g > g(0) = 0$.

2.1 The CEO's information

If the CEO exerts effort, he receives signals α and β. We assume that the distributions of α and β conditional on the realizations of A and B are:

$$\Pr(\alpha = \overline{\alpha} \mid (r_A, r_B)) = \begin{cases} 1 - \varepsilon_A & \text{if } r_A = S \\ \varepsilon_A & \text{if } r_A = F \end{cases} ; \Pr\left(\beta = \overline{\beta} \mid (r_A, r_B)\right) = \begin{cases} 1 - \varepsilon_B & \text{if } r_B = S \\ \varepsilon_B & \text{if } r_B = F \end{cases} . \tag{1}$$

Thus, $\varepsilon_d \in \left[0, \frac{1}{2}\right]$, for $d \in \{A, B\}$, measures the precision of the CEO's information regarding the outcome of decision d. If $\varepsilon_d = 0$, the CEO can perfectly forecast the outcome of decision d; if $\varepsilon_d = 1/2$, the signal about decision d is completely uninformative. Notice that we assume that the signals are independently distributed conditional on the realizations of returns.

We denote by $\overline{a}$ the probability of S under A conditional on the favorable signal, $\overline{\alpha}$:

$$\overline{a} = \Pr(r_A = S|\overline{\alpha}) = \frac{\Pr(\overline{\alpha}|r_A = S)\Pr(r_A = S)}{\Pr(\overline{\alpha}|r_A = S)\Pr(r_A = S) + \Pr(\overline{\alpha}|r_A = F)\Pr(r_A = F)}.$$

We define $\underline{a}$, $\overline{b}$, and $\underline{b}$ analogously. Notice that for $\varepsilon_d \in \left[0, \frac{1}{2}\right)$, $d \in \{A, B\}$, $\overline{a} > \underline{a}$ and $\overline{b} > \underline{b}$, so that $\overline{\alpha}$ and $\overline{\beta}$ are favorable signals on A and B, respectively, and $\underline{\alpha}$ and $\underline{\beta}$ are unfavorable signals.

A decision d is optimal given the pair of signals (α, β) if:

$$\Pr\left(r_d = S \mid (\alpha, \beta)\right) \geq \Pr\left(r_{d'} = S \mid (\alpha, \beta)\right), d \neq d'.$$

We denote by $\hat{\delta}(\alpha, \beta)$ the optimal decision as a function of the pair of signals (α, β).

Depending on the prior distributions of returns and on the relative precision of the signals, the optimal decision conditional on observing a pair of signals may depend on both signals or just on one:

Definition 1 *The optimal decision is*

1. *sensitive only to signal α if $\hat{\delta}(\alpha, \beta)$ depends only on α. Analogously for β.*

2. *sensitive to both α and β if $\hat{\delta}(\alpha, \beta) \neq \hat{\delta}(\alpha', \beta')$ whenever $\alpha' \neq \alpha$ or $\beta' \neq \beta$.*

3. *insensitive to both α and β if $\hat{\delta}(\alpha, \beta)$ is constant in both α and β.*

It follows from our assumption that A is ex ante optimal that the optimal decision if the manager observes a favorale signal for A and an unfavorable signal for B is A. We state this result formally in the following lemma (all proofs are in the appendices):

Lemma 1 *Under Assumpion 1 $\hat{\delta}\left(\overline{\alpha}, \underline{\beta}\right) = A$.*

If A were sufficiently superior ex ante and the signals sufficiently imprecise, then the optimal decision would be A independently of the realization of the signals. To rule out this possibility, we define:

$$\overline{\varepsilon}_A(\varepsilon_B) \equiv \frac{a(1-b)\varepsilon_B}{a(1-b)\varepsilon_B + (1-a)b(1-\varepsilon_B)}, \tag{2}$$

and make the following assumption:

Assumption 2 $\varepsilon_A < 1 - \overline{\varepsilon}_A(\varepsilon_B)$.

This assumption, which we maintain throughout the paper, guarantees that at least one signal is informative, so an unfavorable signal about the ex ante efficient decision A and a favorable signal about the ex ante suboptimal decision B make B ex post efficient. The following lemma formally states this result:

Lemma 2 *Under Assumption 2* $\hat{\delta}\left(\underline{\alpha},\overline{\beta}\right)=B.$

To characterize the optimal decision conditional on the signals when both signals are of the same sign we also define:

$$\underline{\varepsilon}_A\left(\varepsilon_B\right)\equiv\frac{\left(1-a\right)b\varepsilon_B}{\left(1-a\right)b\varepsilon_B+a\left(1-b\right)\left(1-\varepsilon_B\right)}. \tag{3}$$

We note that $0<\underline{\varepsilon}_A\left(\varepsilon_B\right)<\varepsilon_B<\overline{\varepsilon}_A\left(\varepsilon_B\right)$, and that $\overline{\varepsilon}_A\left(\varepsilon_B\right)$ may be larger than $\frac{1}{2}$.

The following lemma characterizes how the optimal decision when both signals are of the same sign depends on the precision of the signals. For simplicity, and without loss of generality, when $\varepsilon_A=\underline{\varepsilon}_A\left(\varepsilon_B\right)$ or $\varepsilon_A=\overline{\varepsilon}_A\left(\varepsilon_B\right)$, we assume $\hat{\delta}\left(\overline{\alpha},\overline{\beta}\right)=A$ and $\hat{\delta}\left(\underline{\alpha},\underline{\beta}\right)=A$ as an indifference-breaking rule.

Lemma 3 *Under Assumptions 1 and 2:*

1. *The efficient decision is sensitive only to* α *if and only if* $\varepsilon_A<\underline{\varepsilon}_A\left(\varepsilon_B\right)$*; in this case* $\hat{\delta}\left(\overline{\alpha},\overline{\beta}\right)=A$ *and* $\hat{\delta}\left(\underline{\alpha},\underline{\beta}\right)=B$*.*

2. *The efficient decision is sensitive only to* β *if and only if* $\varepsilon_A>\overline{\varepsilon}_A\left(\varepsilon_B\right)$*; in this case* $\hat{\delta}\left(\overline{\alpha},\overline{\beta}\right)=B$ *and* $\hat{\delta}\left(\underline{\alpha},\underline{\beta}\right)=A$*.*

3. *The efficient decision is sensitive to* α *and* β *if and only if* $\varepsilon_A\in\left[\underline{\varepsilon}_A\left(\varepsilon_B\right),\overline{\varepsilon}_A\left(\varepsilon_B\right)\right]$*; in this case* $\hat{\delta}\left(\overline{\alpha},\overline{\beta}\right)=A$ *and* $\hat{\delta}\left(\underline{\alpha},\underline{\beta}\right)=A$*.*

Thus, when the precision of α (the signal on decision A) is high enough relative to the precision of β ($\varepsilon_A<\underline{\varepsilon}_A\left(\varepsilon_B\right)$) observing β is irrelevant for determining the efficient decision: If $\alpha=\overline{\alpha}$, the efficient decision is A, and if $\alpha=\underline{\alpha}$, the efficient decision is B. Similarly, if the precision of α is low enough relative to the precision of β ($\varepsilon_A>\overline{\varepsilon}_A\left(\varepsilon_B\right)$), observing β is sufficient to determine the efficient decision. In either of these two cases, the manager could optimally discard one of the two signals. However, when the two signals have a similar precision ($\underline{\varepsilon}_A\left(\varepsilon_B\right)\leq\varepsilon_A\leq\overline{\varepsilon}_A\left(\varepsilon_B\right)$), observation of both signals is necessary to determine the efficient decision. In particular, when the two signals have similar precision, observing $(\alpha,\beta)=\left(\overline{\alpha},\overline{\beta}\right)$ or $(\alpha,\beta)=\left(\underline{\alpha},\underline{\beta}\right)$ does not lead the manager to revise the a priori ordering of the decisions. Thus, knowing that $\alpha=\underline{\alpha}$ (or that $\beta=\overline{\beta}$) is not sufficient to infer that the efficient decision is B.

2.2 The optimal contracting problem

We define an optimal contract as a contract that minimizes the expected salary payment to the CEO among all contracts thar are (i) accepted by the CEO; (ii) such that the CEO weakly prefers exerting effort to shirking; and (iii) such that the CEO weakly prefers to make efficient decisions conditional on the private information he obtains.

Denote by $E(w)$ the expected payment to the manager if the manager accepts the contract, exerts effort, and chooses the optimal decision conditional on the signals. Denote by $V_0(w)$ the manager's ex ante expected utility of monetary payments conditional on playing the optimal decision conditional on the signals. Denote by $V_1(w, d|(\alpha,\beta))$ the manager's expected utility from decision d conditional on having observed signal (α,β). An optimal contract is thus the one that solves the following problem:

$$\min_{w\in\mathbb{R}^4} \quad E(w) \tag{P}$$

$$\text{s.t.} \quad V_0(w) - g \geq \overline{U} \tag{PC}$$

$$V_0(w) - g \geq aU(w_{AS}) + (1-a)U(w_{AF}) \tag{IC-A}$$

$$V_0(w) - g \geq bU(w_{BS}) + (1-b)U(w_{BS}) \tag{IC-B}$$

$$V_1\left(w, \hat{\delta}(\alpha,\beta)\,|\,(\alpha,\beta)\right) \geq V_1(w, d|(\alpha,\beta)), \ \forall d \neq \hat{\delta}(\alpha,\beta), \forall(\alpha,\beta) \tag{IC-$\alpha\beta$}$$

Constraint (PC) is the manager's participation constraint. Constraints (IC-A) and (IC-B) are two first-stage (i.e., prior to the reception of the signals) incentive compatibility constraints that ensure that the manager weakly prefers to exert effort and choose the optimal decision conditional on the signals received over exerting no effort and making either of the decisions. Each of the constraints (IC-$\alpha\beta$), (one for each signal pair $(\alpha,\beta) \in \mathcal{A}\times\mathcal{B}$), is a second-stage (i.e., posterior to the reception of the signals) incentive compatibility constraint that requires that, conditional on the received signal, the manager weakly prefers to make the efficient decision $\hat{\delta}(\alpha,\beta)$.

The assumptions about U guarantee that a solution to the previous problem exists, is unique, and is characterized by its first order conditions.[5]

[5] To see this notice that one can rewrite the previous problem in terms of CEO utilities. By $U'' < 0$, the objective function would be strictly quasi-convex and the constraints would be linear and therefore continuous and convex.

3 Conditioning pay on decisions

Before characterizing the optimal contract, we first address the question as to whether the CEO's pay should be made contingent on the CEO's decisions or only on outcomes:

PROPOSITION 1 *Let $w = (w_{AF}, w_{AS}, w_{BF}, w_{BS})$ be the optimal contract. If $a \neq b$, then, either $w_{AF} \neq w_{BF}$ or $w_{AS} \neq w_{BS}$, or both. If $a = b$ and $\varepsilon_A = \varepsilon_B$, then $w_{AS} = w_{BS} > w_{AF} = w_{BF}$.*

Proposition 1 states that a sufficient condition for making it optimal to condition the CEO's pay on the decision made by the CEO and not only on its realization is that one of the two decisions has a higher ex ante probability of success. Only in the case in which both decisions are considered ex ante to have exactly the same probability of success it may be optimal to condition the CEO's pay only on returns, a possibility that we discuss further in Section 7. As mentioned in the introduction, Murphy (1999) has argued, in his oft-cited review of the executive compensation literature, that uncertainty about the efficient decision would lead to contracts that are based on the principal's objective. Proposition 1 shows that a necessary condition for this is the existence of extreme uncertainty about the efficient decision, in the sense that the principal's prior assigns the same probability of success to both decisions.

To understand Proposition 1, it is worth noting that any contract w that conditions pay only on returns ($w_{AF} = w_{BF} = w_F < w_S = w_{AS} = w_{BS}$) will give the CEO the incentives to make the efficient decision contingent on the signal received, since the efficient decision by definition leads to a higher probability of success. It is possible, however, to maintain those incentives while reducing the risk faced by the CEO by conditioning pay on the CEO's decision. Since the CEO is risk averse, this reduces the cost of the contract to the principal. To see this, take any contract w that satisfies the participation constraint and all incentive compatibility constraints, and that does not condition pay on the CEO's decision. This contract must establish $w_S > w_F$, so that all second-stage incentive compatibility constraints hold with slack. Moreover, since $a > b$, ensuring that the CEO prefers to exert effort rather than shirking and choosing A also ensures that he prefers to exert effort rather than shirking and choosing B. Starting from this contract, it is possible to compress the CEO's pay while ensuring that the participation constraint and all incentive compatibility constraints still hold. A possible way to do this is to reduce pay in case of success if the CEO takes decision A and to increase pay in case of failure if the CEO takes decision B in such a way that the ex ante expected utility of the CEO is not changed. Since the CEO is risk averse, this reduces the expected cost of the contract. Further, for small enough changes, all but one of the incentive compatibility constraints will still hold, since they held with slack for contract w. But the remaining incentive compatibility constraint—the one ensuring that the CEO prefers to exert effort over exerting no effort and choosing A—will also hold, since the proposed deviation does not change the CEO's expected utility, yet it reduces the expected utility of exerting no effort and choosing A. It is worth noting that, even though the proposed deviation would implement the efficient action profile at a lower cost than a contract that does not condition pay on the CEO's decision, it does not follow that the optimal contract necessarily entails $w_{AS} < w_{BS}$ and $w_{AF} < w_{BF}$, as implied by the proposed deviation.

4 Rewarding decisions

Consider the benchmark case in which both signals are equally informative ($\varepsilon_A = \varepsilon_B = \varepsilon$). In this case, it follows from Lemma 3 that if both signals have the same sign, the ex-post optimal decision is the ex ante optimal decision A. Thus, decision B is optimal only conditional on $(\underline{\alpha}, \overline{\beta})$. In this case, the optimal contract is as follows:

Proposition 2 *The optimal contract satisfies:*

$$w_{AF} < w_{BF} < w_{AS} < w_{BS}.$$

Therefore, pay is increasing in returns, both conditionally on the decision made ($w_{dF} < w_{dS}$, for any d) and unconditionally ($w_{dF} < w_{d'S}$, for any $d \neq d'$). This monotonicity, by itself, implies that, if the CEO exerts effort, he has the incentive to make the decision that is optimal given the signal.

In line with Proposition 1, the CEO's pay depends on the decision made ($w_{Ar} \neq w_{Br}$, for $r \in \{F, S\}$). Moreover, pay is monotonic in the decision made, in the sense that, conditional on a given outcome, one decision always leads to strictly higher pay. Further, the decision that leads to greater pay is ex ante suboptimal. Thus, when A is the ex ante optimal decision, the optimal contract pays the CEO more, conditional on returns, if he takes decision B. This is noteworthy, as it implies that the CEO is paid more for making decisions that are wrong on average (and that are known to be wrong on average by the principal). The rationale for the result, however, is compelling. If pay is weakly increasing in performance, then a CEO that does not exert effort and that, as a result, does not obtain any additional information about the return distributions, will always find it optimal to make the ex ante optimal decision. Only an informed CEO could possibly choose the ex ante suboptimal decision if, upon exerting effort, he learnt that it could be expected to be superior to the alternative. In other words, the only ex ante incentive constraint that may be binding is the one that ensures that the manager prefers to exert effort and choose optimally rather than to not exert effort and choose A. The principal can relax this constraint and, thus, lower the cost of the contract by reducing pay if the ex ante optimal project is chosen (and increasing pay if the the ex ante suboptimal project is chosen to ensure that the manager's participation constraint is satisfied). In other words, it is cheaper for the principal to elicit effort if the contract penalizes the manager for choosing what an uninformed manager would choose (given the monotonicity of the contract).

To shed more light on Proposition 2, it is useful to analyze the optimal contract if the CEO always made the optimal decision contingent on the signals, so that the only problem for the principal is to induce effort provision.[6] In this case, as long as the contract is weakly monotonic in returns, a CEO who shirks will never choose project B, the ex-ante suboptimal project. Therefore, the likelihood of observing project B relative to project A is higher when the CEO exerts effort than when he shirks (in fact, it is infinitely higher). This implies that, in order to induce effort, the CEO's pay if project B is chosen should be higher than if project A is chosen. If the manager needs to be provided incentives to choose the optimal project conditional on the signals, this conclusion need not hold any more, since paying the manager for choosing B may distort project choice. However, in the baseline case with $\varepsilon_A = \varepsilon_B$, the result that the manager should be paid more if he chooses project B carries over to the case in which the principal needs to provide incentives to the manager both to exert effort and to select the optimal project conditional on the signal.

5 Decision complexity and pay-performance sensitivity

In the previous section, we assume that the signals about the likely returns of the different projects have the same precision. However, CEOs may often be able to forecast the returns of different projects with different degrees of precision. For example, consider a firm that has to decide whether to enter a market that is similar to its home market or a completely new market. Even if the manager allocates his effort optimally to the tasks of investigating the consequences

[6]This amounts to reducing the problem to a simple moral hazard problem with unobservable effort.

of each alternative, it is likely that in the end he will have a more precise profit forecast for the known market than for the completely new market. Thus, an asymmetry in the manager's ability to learn the consequences of the different alternatives may arise because some decisions are more complex or there is less prior knowledge about them that the manager can draw upon to interpret the information he receives. But the asymmetries in the precision with which the manager can forecast the outcomes of different decisions may also be idiosyncratic to the manager. A manager may be specialized in assessing the consequences of one of the available decisions, but may not be able to obtain a significant informational advantage with respect to shareholders regarding an alternative decision. Continuing with the example of a firm deciding which market to enter, the firm's manager may have a long experience with one of the two markets, while he may not be better equipped than board directors to assess the outcome of the entry into the other market.

The purpose of this section is to characterize how the optimal contract depends on the precisions of the two signals received by the CEO. We will refer to two cases. In the first case, we will analyze situations in which the optimal decision is sensitive to both signals, α and β. Because the optimal decision is sensitive to α and β when the signals received by the CEO have sufficiently similar precisions, we will refer to this case as one in which decisions have *similar complexity*. In the second case, we will analyze situations in which the optimal decision is sensitive to only one of the two signal, α or β. Because this happens when the signals received by the CEO have sufficiently different precisions, we will refer to this case as one in which decisions have *different complexity*.

5.1 Decisions of similar complexity

Consider the case in which the optimal decision is sensitive to α and β. By Lemmas 1, 2 and 3, under Assumptions 1 and 2 the optimal decision is A unless the CEO observes $(\underline{\alpha}, \overline{\beta})$, in which case the optimal decision is B.

PROPOSITION 3 *Suppose that Assumptions 1 and 2 are satisfied and the optimal decision is sensitive to α and β, i.e., $\varepsilon_A \in [\underline{\varepsilon}_A(\varepsilon_B), \overline{\varepsilon}_A(\varepsilon_B)]$. Then, there exist $\underset{\sim}{\varepsilon}_A(\varepsilon_B)$ and $\widetilde{\varepsilon}_A(\varepsilon_B)$ satisfying $\underline{\varepsilon}_A(\varepsilon_B) < \underset{\sim}{\varepsilon}_A(\varepsilon_B) < \varepsilon_B < \widetilde{\varepsilon}_A(\varepsilon_B) < \frac{1}{2}$ such that:*

1. *If $\underline{\varepsilon}_A(\varepsilon_B) \leq \varepsilon_A \leq \underset{\sim}{\varepsilon}_A(\varepsilon_B)$, the optimal contract w satisfies:*

$$w_{AS} < w_{BS};$$

2. If $\underset{\sim A}{\varepsilon}(\varepsilon_B) < \varepsilon_A \leq \widetilde{\varepsilon}_A(\varepsilon_B)$*, the optimal contract* w *satisfies:*

$$w_{AF} < w_{BF} < w_{AS} \leq w_{BS};$$

3. If $\widetilde{\varepsilon}_A(\varepsilon_B) < \varepsilon_A \leq \widehat{\varepsilon}_A(\varepsilon_B)$*, the optimal contract* w *satisfies:*

$$w_{AF} < w_{BF} < w_{BS} < w_{AS}.$$

Therefore, when signals have very similar precision (case 2) the optimal contract has the same form as the one described in Proposition 2 which describes the case in which the signals have identical precisions. However, as the precision of one of the signals increases relative to the other, the contract can cease to be monotonic in the decisions, that is, which decision commands a higher pay depends on the realization of profits.

5.2 Decisions with different levels of complexity

Consider the case in which the optimal decision is sensitive to only one of the two signals, α or β.

PROPOSITION 4 *Suppose that Assumptions 1 and 2 are satisfied. (1) If the optimal decision is sensitive only to* α*, the optimal contract satisfies*

$$w_{AF} < w_{BF} = w_{BS} < w_{AS}.$$

(2) If the optimal decision is sensitive only to β*, the optimal contract satisfies*

$$w_{BF} < w_{AF} = w_{AS} < w_{BS}.$$

As in the case of decisions of similar complexity, the optimal contract is weakly increasing in performance to ensure that the CEO makes the optimal decision contingent on the signal received. However, when the optimal decision is sensitive only to one of the signals, pay is independent of profits if the CEO makes the more complex decision. This result implies that we may observe pay that is unresponsive to performance, even when the contract is optimally set. Another feature of the optimal contract is that, contrary to the previous cases, the form of the contract depends on which decision is more complex and not on which decision is ex ante optimal. Even if in both cases contemplated in Proposition 4 the ex ante optimal decision is A (Assumption 1) the ranking of salaries depends only on the precision with which the outcomes of the two decisions can be forecasted by the CEO. It is important to note that the CEO's pay is not monotonic in the decision made, that is, it is not the case that, contingent on returns, one of the decisions always leads to a higher pay.

To understand part 1 of Proposition 4, it is worth discussing why there should be a positive pay-performance sensitivity (PPS) in our context. Since the manager is risk averse and has no preference for any project, flat pay would be optimal from the point of view of risk sharing and would implement the optimal decision conditional on the signals. However, a contract with flat pay would not elicit effort from the agent. Thus, some PPS is needed to induce effort provision, but the manager's risk aversion implies that the principal will want to limit PPS as much as possible to reduce the cost of the contract. Now, if the optimal decision is sensitive only to α, the manager should choose A when $\alpha = \overline{\alpha}$ and B when $\alpha = \underline{\alpha}$. That is, it has to be the case that for any β:

$$V_1(w, A|(\overline{\alpha}, \beta)) \geq V_1(w, B|(\overline{\alpha}, \beta)) = V_1(w, B|(\underline{\alpha}, \beta)) \tag{4}$$

$$V_1(w, A|(\underline{\alpha}, \beta)) \leq V_1(w, B|(\underline{\alpha}, \beta)) = V_1(w, B|(\overline{\alpha}, \beta)) \tag{5}$$

Therefore, to elicit effort provision the contract has be such that $V_1(w, A|(\overline{\alpha}, \beta)) > V_1(w, A|(\underline{\alpha}, \beta))$, which implies that $w_{AS} > w_{AF}$, so there must be positive PPS if the manager selects project A. Now, as long as the above inequalities hold for both $\overline{\beta}$ and $\underline{\beta}$, the manager will choose the optimal project irrespectively of whether $w_{BS} > w_{BF}$ or $w_{BS} = w_{BF}$. Since the manager is risk averse, it is thus optimal for the principal to set $w_{BS} = w_{BF}$. In other words, if the optimal decision is sensitive only to α, a positive PPS if A is chosen is sufficient both to induce effort and to elicit the optimal decision. Note, however, that this not possible in the case in which both decisions have similar complexity, since in such case it is necessary to ensure that the manager obtains more if he chooses B when $\beta = \overline{\beta}$ than when $\beta = \underline{\beta}$.

6 Correlated outcomes

Until now we have considered the case in which the returns of projects A and B are independent. In many cases, however, these returns may not be independent. For example, if the CEO has to decide whether to launch a new product in one of two countries which are not too dissimilar, it is reasonable to expect that if the product is good, returns will tend to be large in both countries and if the product is bad returns will tend to be low in both. In cases such as this, r_A and r_B may be positively correlated. In other cases, however, r_A and r_B may be negatively correlated. For example, if the CEO has to decide whether to adopt one of two technological standards, network externalities associated with standards imply that the returns from adopting the two standards are negatively correlated: if standard A succeeds, r_A will tend to be large and r_B will tend to be low, and vice versa. When the returns from the two decisions have a positive correlation, obtaining information about their likely realizations will have a smaller impact on the optimal decision and the value of a CEO that can obtain this information will also be lower. For this reason, in the following we focus on the case in which the returns of the projects are negatively correlated and, for the sake of simplicity, consider the case in which there exists negative perfect correlation between A and B. In other words, we assume that with probability a, $(r_A, r_B) = (S, F)$, and with probability $1 - a$, $(r_A, r_B) = (F, S)$. Notice that the equivalent of Assumption 1 that guarantees that A is ex-ante optimal is that $a > \frac{1}{2}$.

To make the results comparable with the previous section we assume that if the CEO exerts effort, he obtains only one signal, α, rather than two. The reason is that this implies that, as in the previous section, the CEO receives a signal with only two possible intensities *for each of the two projects*. For example, $\overline{\alpha}$ is favorable and $\underline{\alpha}$ is unfavorable concerning A; the opposite holds for B. With two independent signals, the perfect negative correlation of the returns of A and B would imply that the CEO receives a signal with four different intensities *for each of the two projects*. For example, for A, $\left(\overline{\alpha}, \underline{\beta}\right)$ is very favorable, $\left(\underline{\alpha}, \overline{\beta}\right)$ is very unfavorable and $\left(\overline{\alpha}, \overline{\beta}\right)$ and $\left(\underline{\alpha}, \underline{\beta}\right)$ are intermediate; the opposite holds for B.

The conditional probability distribution of signal α is

$$\Pr\left(\alpha = \overline{\alpha} \mid (r_A, r_B)\right) = \begin{cases} 1 - \varepsilon & \text{if } (r_A, r_B) = (S, F) \\ \varepsilon & \text{if } (r_A, r_B) = (F, S) \end{cases},$$

with $\varepsilon \in \left[0, \frac{1}{2}\right)$. As in the previous sections, if the CEO shirks, he receives no signal.

We consider only the case in which the optimal decision is sensitive to α (or otherwise the signal would have no value). Therefore, the optimal decision is A conditional on $\overline{\alpha}$ and B conditional on $\underline{\alpha}$. The following proposition provides a characterization of the optimal contract for this information structure:

PROPOSITION 5 *Consider the case of perfect negative correlation of A and B. Suppose that A*

is ex-ante optimal and that the optimal decision is sensitive to α. Then the optimal contract satisfies

$$w_{AF} < w_{BF} \leq w_{AS} < w_{BS}$$

Therefore, the optimal contract yields the same ranking of salaries as in the case in which decisions are independent and signal precisions are sufficiently similar (compare with Proposition 2 and with part 2 of Proposition 3). The rationale for the contract form is the same as in that case. On the one hand, monotonicity with respect to returns guarantees that if the CEO exerts effort, he will make the decision that is optimal contingent on the signal received. On the other hand, paying the CEO more if he makes the ex ante suboptimal decision allows the principal to elicit effort at a lower cost.

7 Uninformative prior

The results in Propositions 2-4 are derived under the assumption that one of the decisions is ex ante optimal (Assumption 1). In this section we briefly discuss the extent to which our results depend on this assumption. To do this we make the extreme assumption that $a = b$, that is, we assume that the priors are so uninformative that the two decisions are equally attractive ex-ante. For the sake of simplicity we omit the proof of the results and we simply relate them to the results of the previous sections.

Consider first the case in which decisions are independent. If $\varepsilon_A = \varepsilon_B$, then the complete symmetry between the two decisions implies that a contract that does not condition on the decision is optimal. However, if $\varepsilon_A \neq \varepsilon_B$ this symmetry is broken and it becomes optimal to condition of the decision. Since $a = b$, the optimal decision is sensitive only to α if and only if $\varepsilon_A < \varepsilon_B$ and the optimal decision is sensitive only to β if and only if $\varepsilon_A > \varepsilon_B$. Therefore, the optimal contract for this case would lead to the same rankings of salaries as the ones obtained in Proposition 4. This is not surprising, since Proposition 4 shows that the optimal decision is sensitive only to one signal, the ranking of salaries does not depend on which decision is ex ante optimal, but only on the precisions of the signals received by the CEO..

In the case in which the returns of the two decisions are perfectly negatively correlated and in which $a = b$, the ex ante symmetry carries through to the optimal contract, which would make the CEO's pay depend only on the return realization.

Therefore, uncertainty about the optimal decision should translate into contracts that condition pay only on the principal's objective (returns, in our case) exclusively in cases of extreme symmetry in which the priors on the two decisions are identical ($a = b$), and the precision of the signals received by the CEO are also identical.

8 Discussion and Applications

8.1 Implications for Contract Design

The model presented here has several implications for the design of executive compensation contracts:

1. *Pay should generally depend on observable decision and not only on returns.* When return-relevant decisions are observable and the optimal return-relevant decision is not known by shareholders, then it is generally optimal to make the manager's pay depend on the observable decision. Pay should not depend on the observable decision only in the extreme case in which shareholders (or board directors) are unable to rank decisions ex ante and the set of available decisions or the manager's idiosyncratic features are such that the manager would receive information with the same degree of precision about all feasible decisions.

2. *Expectations about the optimal decision matter.* The relation between pay and observable decisions may depend on which decision shareholders *expect* to be optimal, even if shareholders (or board directors) *do not know* with certainty which decision is optimal.

3. *Optimal contracts may pay managers to go against the (ex ante) odds.* This implication has the same statistical inference logic of the informativeness principle. If managers are not specialized, then they will go against the ex ante odds only if they have exerted effort and obtained information that favors the ex ante suboptimal decision. Therefore, to elicit effort in information acquisition, it may be optimal to explicitly reward decisions that go against shareholders' expectations.

4. *The precision of information matters.* A manager may be able to predict the consequences of some decisions better than those of others. This may be due to his own idiosyncratic specialization or may occur because the outcomes of some decisions are inherently more difficult to forecast. In these cases the ranking of salaries is determined by the precision of information and not by shareholders' ex-ante expectations on the optimal decision.

5. *Lack of performance-sensitivity may be optimal following some decisions.* When the information a manager obtains about a decision is sufficiently less precise than the information he obtains on the alternative, his pay should be respond to performance if he makes the decision on which he obtains more precise information. By contrast if the manager makes the decision on which he has less precise information he should not be rewarded for success or penalized for failure.

6. *Pay may not be monotonic in decisions.* This occurs in all cases in which the manager obtains information with sufficiently different degrees of precision on the feasible decisions. In these cases, whether the manager is paid more for making one decision or the other will depend on performance: conditional on low performance, pay will be higher when the manager makes the decision on which he has less precise information; conditional on high performance, however, pay will be higher when the manager takes the decision in which he has more precise information.

8.2 Pay-Performance Sensitivity

CEOs' pay-performance sensitivity is often argued to be too low (Jensen and Murphy, 1990), though conclusions depend on the metric used to measure it (see, e.g., Murphy, 1999; Hall and Liebman, 1998; Baker and Hall, 2004). Several authors have also documented a positive association between managerial stock ownership and firm value (at least for moderate levels of managerial ownership) and between the level of equity incentives and firm value (see Frydman and Jenter, 2010). Although the empirical results are not conclusive,[7] the early evidence and the influential article by Jensen and Murphy (1990) suggested that (suboptimally) low pay-performance sensitivity would be associated with lower firm performance.

Our results show that, in some cases, a lack of performance sensitivity is optimal. As discussed above, when managerial specialization or inherent forecasting complexity make it harder to forecast the returns of one of the projects, it is optimal to make pay unresponsive to performance when the CEO selects the hard-to-forecast project. Moreover, the incidence of a lack of performance sensitivity may be high for plausible parameter values. For example, if $a = 0.6$, $b = 0.5$, $\varepsilon_A = 0.3$, and $\varepsilon_B = 0.4$ zero performance sensitivity would occur with probability of 0.46.

[7] See Frydman and Jenter (2010) for an assessment of the empirical literature relating executive pay and firm value.

Moreover, our results have the implication that when only one signal is informative, a low performance sensitivity is associated with lower average performance.

Corollary 1 *Suppose only one signal is informative. Then, expected performance is lower conditional on choosing the decision whose signal is not informative.*

To see this, suppose that only α is informative. In this case, if the manager chooses A, his pay is sensitive to performance, and the probability of success is $\overline{a} > a$. If the manager chooses B, his pay is completely insensitive to performance and the probability of success is $b < \overline{a}$. Therefore, our results suggest caution when interpreting observed positive correlations between PPS and firm performance as an indication that firms with low realized PPS are acting suboptimally and would benefit from an increase in PPS.

8.3 Mergers and Acquisitions

Managers often receive large bonuses when they acquire other firms (Grinstein and Hribar, 2004), even if acquisitions have been shown to be associated, on average, with a reduction in the acquiring firms' value (Andrade et al., 2001; Moeller et al., 2004 and 2005). That being the case, it seems puzzling that the managers of acquiring firms are rewarded for acquisitions. Several possible explanations for this puzzle have been proposed. The first one is, simply, that the acquisition process takes extraordinary effort from managers, so they need to be paid more in the event of an acquisition to induce them to put that extra effort or, simply, to satisfy their ex post participation constraint. Although the length of the acquisition process seems to increase the manager's bonus, the explanation based on effort is, at best, partial (Grinstein and Hribar, 2004). Further, other authors have proposed that managers have intrinsic incentives to acquire other firms because they are motivated by empire building (Jensen, 1986) or that they tend to overestimate their ability to benefit from the acquisition (the hubris hypothesis of Roll, 1986). According to these authors, incentives would, in fact, be needed, to dissuade managers from acquiring other firms. Acquisition bonuses may be the result of managers' discretion over their own compensation. Thus, managers may use acquisitions as a way to justify pay increases to investors (Bebchuk and Fried, 2004, ch. 10).

Our model shows that, irrespectively of the sign of the effect of acquisitions on firm value, acquisitions should generally be taken into account when determining CEOs' pay. More importantly our model characterizes CEO's pay when acquisitions, despite lowering the stock price of the acquiring firm, are optimal for the acquiring firm's shareholders. This may be the case if the acquisition, despite being optimal, reveals negative information to the market about the acquiring firm (Jovanovic and Braguinsky, 2004),[8].

To analyze the implications of the model for the relation between acquisitions, CEO pay, and firm value, consider a firm with two alternatives: Alternative A is to not acquire any other firm; alternative B is to acquire a potential target. Assuming for simplicity that in either case, the firm can either succeed (in which case, firm value is $S = 1$) or fail (in which case, firm value is $F = 0$), the expected profitability of an alternative is fully determined by the probability of success.

The firm's directors have an assessment of the probability of success of each alternative and, thus, of their expected profitability. Suppose that, on the basis of publicly available information about the acquiring firm, potential target, and market conditions, the board of directors believes that the probability of success is greater if the firm does not carry out the acquisition, that is: $a > b$. However, the board of directors understands that further investigation by a CEO with unique skills and access to information about the firm may show that the acquisition is optimal.

Suppose that $\varepsilon_A \in \left(\underset{\sim A}{\varepsilon}(\varepsilon_B), \widetilde{\varepsilon}_A(\varepsilon_B)\right)$. In this case, as shown in Proposition 2, $a > b$ implies that the optimal compensation contract for the CEO has the form $w_{AF} < w_{BF} < w_{AS} < w_{BS}$. Therefore, if the board believes the acquisition to be (ex ante) suboptimal, then it is optimal to offer the CEO a contract that rewards him if he chooses to make the acquisition. Note that the contract is nonetheless monotonic in firm performance, so not only $w_{BS} > w_{BF}$, but $w_{AS} > w_{BF}$ as well.

If the contract is designed optimally, then the CEO will not only obtain and process all possible information relevant for the decision (i.e., exert effort), but will also make the optimal decision conditional on that information. It follows that the acquisition will take place if and only if it is ex post optimal. However, news of the acquisition may be accompanied by a reduction in firm value. To see this, consider an example with $\varepsilon_A = \varepsilon_B = 0.2$, $a = 0.7$, and $b = 0.4$. Let W_0 be the value of the firm before the decision is made but with knowledge that the optimal contract has been implemented, let W_A be the value of the firm if the CEO makes decision A with knowledge that the optimal contract has been implemented, and let W_B be defined analogously.

[8] Proponents of this view, however, have not discussed whether managerial pay should be tied to acquisitions and if so how.

It follows from $\varepsilon_A = \varepsilon_B$ and Lemma 3 that:

$$W_A = \frac{P(\overline{\alpha})\overline{a} + \pi_{\underline{\alpha\beta}}\underline{a}}{P(\overline{\alpha}) + \pi_{\underline{\alpha\beta}}} = 0.77 \tag{6}$$

$$W_B = \overline{b} = 0.73 \tag{7}$$

$$W_0 = (P(\overline{\alpha}) + \pi_{\underline{\alpha\beta}})W_A + \pi_{\underline{\alpha}\overline{\beta}}W_B = 0.76 \tag{8}$$

Therefore, for this example, $W_B < W_0$, so that news of the acquisition would reduce value even if the market correctly understands that the manager will acquire the target if and only if it is optimal to do so.

This example parallels the result obtained by Jovanovic and Braguinsky (2004) that even if an acquisition is ex post optimal for the acquiring firm, the news of the acquisition may nonetheless lead to a reduction of firm value. In our case, observing an acquisition shows that the CEO received good news about the target firm and bad news about the firm's own projects. For the chosen parameter values (and for a broad range of other plausible parameter values), the bad news conveyed by an acquisition outweigh the good news, resulting in a reduction in firm value.[9]

To fully understand the empirical implications of the example, let us assume that we have N ex ante identical firms like the acquiring firm described above and, for simplicity, that these firms are not competing for targets nor interacting in any other way. Then, if we examined the relation between observed acquisitions, firm value, and pay, we would observe that: a) controlling for firm returns, pay is higher for managers who have acquired the target, b) upon announcement of the decision to acquire the target, firm value goes down, and c) realized value is lower on average in firms that acquire the target.

9 Conclusion

In this paper, we propose a model in which shareholders delegate decision-making to the CEO. The reason why they do so is that the CEO has the time and skills, which shareholders lack, to learn which of the alternative strategies available to the firm is optimal. Although the CEO has superior skills, he needs to exert costly effort to acquire and process the information necessary to improve his estimate of the returns of the different strategies. Shareholders observe both

[9] For simplicity, we are abstracting from the determination of the price paid for the target firm. Thus, the profit in case of success if there is an acquisition should be understood as net of the acquisition price. Also for simplicity, we are disregarding the CEO's salary when computing firm value. One can easily extend the model to allow for an acquisition premium and to compute firm value net of CEO pay and obtain the same results.

the CEO's decision and the resulting profits, but not the effort exerted by the CEO to obtain information nor the information acquired by the CEO. Therefore, shareholders can condition the CEO's pay both on profits and on the observable decision. In this context, the CEO's contract has to solve two related incentive problems: it has to a) provide incentives to the CEO to exert effort to acquire information and b) ensure that the CEO has the incentives to make the optimal decision conditional on his information.

The assumptions of the model are meant to capture what we consider to be the main agency problem between shareholders and managers. Shareholders and, certainly, board directors—who actually set the manager's compensation—can observe the most relevant decisions made by managers, namely those firm-wide strategic decisions that determine the firm's scope, organizational form, technology, financing and the markets in which it operates. Therefore, they can, in principle, condition the manager's pay on those strategic decisions. The problem is, however, that the very reason why they delegate those decisions to a CEO is, most often, that they believe that the CEO is better equipped than they are to make optimal decisions for the firm.

We study the optimal contract in this context and obtain several novel predictions about the form of the executive compensation contract. First, we show that, under very general conditions, the optimal contract conditions pay not only on profits, but also on the decision made by the manager. Second, we show that shareholders' (or board directors') ex ante beliefs about the optimal decision determine the form of the optimal compensation contract, a prediction that, to our knowledge, had not been obtained before in agency models. Third, we show that degree of precision with which a CEO manager learns about different decisions also influences the form of the compensation contract. Fourth, we characterize the relation between profits, decisions, and pay implied by the optimal contract in different contexts and find, among other implications, that: i) it may optimal to reward managers for making decisions which are ex ante suboptimal; ii) it may be optimal to set a very low pay-performance sensitivity following certain decisions; and iii) optimal contracts may be nonmonotonic in decisions in the sense that the decision that is rewarded conditional on a success may be different from the decision that is rewarded conditional on failure.

We have discussed the implications of the model for the questions as to whether CEOs should be rewarded for making certain strategic decisions, such as acquiring other firms. We believe, however, that our model has normative implications which have a much broader application, since they provide guidelines concerning the variables that should be included, and in which way, in compensation contracts. Thus, our model could be applied to the determination the optimal compensation of other executives and lower-level managers, to the extent that these managers are delegated decisions that are observable by superiors. Our results can also be applied to the study of the compensation contracts of expert agents other than CEOs, such as investment analysts, portfolio managers, medical doctors, or lawyers.

References

[1] Andrade G., M. Mitchell, and E. Stafford (2001) "New Evidence and Perspectives on Mergers," *The Journal of Economic Perspectives*, 15, 103–120.

[2] Baker, G. P. and J. Hall (2004). CEO Incentives and Firm Size. *Journal of Labor Economics* 22, 767–798.

[3] Barron, J.M., Waddell, G.R. (2003). Executive rank, pay and plan selection. *Journal of Financial Economics* 67, 305–349.

[4] Barron, J.M., Waddell, G.R. (2008). Work hard, not smart: Stock options in executive compensation. *Journal of Economic Behavior & Organization* Vol. 66 (2008) 767–790.

[5] Bebchuk, L. and J. Fried (2004). *Pay without Performance: The Unfulfilled Promise of Executive Compensation.* Harvard University Press.

[6] Bebchuk, L. and Y. Grinstein (2005). Firm Expansion and CEO Pay. *NBER Working Paper* 11886.

[7] Bliss, R.T., Rosen, R.J. (2001). CEO compensation and bank mergers. *Journal of Financial Economics* 61, 107–138.

[8] Dai, C., Lewis, T. R. and Lopomo, G. (2006). Delegating management to experts. *RAND Journal of Economics* 37(3), 503–520.

[9] Demski, J. S. and Sappington, D. E. M. (1987). Delegated expertise. *Journal of Accounting Research* 25(1), 68–89.

[10] Fich, E. M., L. T. Starks, and A. S. Yore (2010). CEO deal-making activity, CEO compensation and firm value. *Working Paper, Drexel University.*

[11] Frydman, C. and Jenter, D. (2010). CEO Compensation. *Annual Review of Financial Economics* 2(1).

[12] Grinstein Y., and Hribar P. (2004). CEO Compensation and Incentives: Evidence from M&A Bonuses. *Journal of Financial Economics* 71, 119–143.

[13] Gromb, D. and Martimort, D. (2007). Collusion and the organization of delegated expertise. *Journal of Economic Theory* 137(1), 271–299.

[14] Harford, J., Li, K. (2007). Decoupling CEO wealth and firm performance: The case of acquiring CEOs. *Journal of Finance* 62, 917–949.

[15] Hall, B.J. and Liebman, J.B. (1998). Are CEOs Really Paid Like Bureaucrats?. *The Quarterly Journal of Economics*, 113, 653–691.

[16] Holmström, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics* 10, 231–259.

[17] Holmström, B. (1982). Moral Hazard in Teams. *Bell Journal of Economics* 13, 324-340.

[18] Jensen, M.C. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review* 76, 323—329.

[19] Jensen, M. and W. Meckling (1976. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* 3, 305–360.

[20] Jovanovic, B. and S. Braguinsky (2004). Bidder Discounts and Target Premia in Takeovers. *American Economic Review* 94, 46–56.

[21] Lambert, R.A., 1986. Executive effort and the selection of risky projects. *RAND Journal of Economics* 17, 77–88.

[22] Malcomson, J. M. (2009). Principal and Expert Agent. *The B.E. Journal of Theoretical Economics (Contributions)* 9 (1), Article 17.

[23] Moeller S., F. Schlingemann, and R. Stultz (2005). Wealth Destruction on a Massive Scale? A Study of Acquiring-Firm Returns in the Recent Merger Wave. *Journal of Finance* 60, 757–782.

[24] Moeller S., F. Schlingemann, and R. Stultz (2004). Firm Size and the Gains from Acquisitions. *Journal of Financial Economics* 73, 201-228.

[25] Murphy, K.J. (1999). Executive Compensation" in O. Ashenfelter and D. Card (1999) (eds.) *Handbook of Labor Economics*, Vol. 3, North Holland, Amsterdam.

[26] Roll, R. (1986). The Hubris Hypothesis of Corporate Takeovers. *Journal of Business* 59, 197–216.

A Proofs

A.1 Proof of Lemma 1

Immediate from observing that Assumption1 implies that $\overline{a} > \underline{b}$.

A.2 Proof of Lemma 2

Notice that $\hat{\delta}\left(\underline{\alpha}, \overline{\beta}\right) = B$ is equivalent to $\underline{a} < \overline{b}$. From the definitions of $\underline{a}$ and $\overline{b}$ we obtain

$$a\varepsilon_A(b(1-\varepsilon_B) + (1-b)\varepsilon_B) \leq b(1-\varepsilon_B)(a\varepsilon_A + (1-a)(1-\varepsilon_A))$$

which is equivalent to

$$\varepsilon_A < \frac{(1-a)\, b\, (1-\varepsilon_B)}{a\,(1-b)\,\varepsilon_B + (1-a)\, b\,(1-\varepsilon_B)} = 1 - \overline{\varepsilon}_A\left(\varepsilon_B\right).$$

A.3 Notation

We denote by $\pi_{\alpha\beta}$ the unconditional probability of the CEO receving signals (α, β) if he exerts effort. By the assumption of independence of A and B and of α and β, we have

$$\pi_{\alpha\beta} = \Pr\left(\alpha, \beta\right) = \Pr\left(\alpha\right)\Pr\left(\beta\right) \tag{9}$$

We denote by $V_{\overline{A}}$ the expected utility of the salary payment to the CEO when he chooses A after having observed a favorable signal on A, i.e. after having observed signal $\overline{\alpha}$. In a similar way we denote by $V_{\underline{A}}$ the expected utility of the salary payment to the CEO when he chooses A after having observed an unfavorable signal on A, i.e. after having observed signal $\underline{\alpha}$. Notice that the assumption of independence of A and B and of α and β implies that the expected utility of the salary payment to the CEO when he chooses A depends on α, but is independent of β. We define $V_{\overline{B}}$ and $V_{\underline{B}}$ in a similar way.

We denote by V_A (respectively, V_B) the unconditional expected utility of the salary payment to the CEO when he chooses A (respectively, B). We therefore have

$$V_A = \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\underline{A}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}}$$

$$V_B = \pi_{\overline{\alpha}\underline{\beta}} V_{\underline{B}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{B}}$$

Finally we denote by V the expected utility of the CEO of a contract, under the assumption that the CEO exerts effort and makes the optimal decision. In the case in which the optimal decision is sensitive to α and β (which implies that the optimal decision is B if and only if $(\alpha, \beta) = \left(\underline{\alpha}, \overline{\beta}\right)$), for example we have

$$V \equiv \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}} - g \tag{10}$$

A.4 Proof of Propositions **1**, 2, and 3

We start by characterizing the solution to the cost minimization problem when the optimal decision is sensitive to α and β. Recall that when the optimal decision is sensitive to α and β, the optimal decision is B if and only if $(\alpha, \beta) = \left(\underline{\alpha}, \overline{\beta}\right)$. The contract that minimizes the salary cost of inducing effort and implementing the optimal decision is therefore a solution to the following problem

$$\begin{aligned}
\min_{w \in \mathbb{R}^4} \quad & \pi_{\overline{\alpha}\underline{\beta}} \left(\overline{a} w_{AS} + (1 - \overline{a}) w_{AF}\right) + \pi_{\underline{\alpha}\overline{\beta}} \left(\overline{b} w_{BS} + (1 - \overline{b}) w_{BF}\right) + \\
& + \pi_{\overline{\alpha}\overline{\beta}} \left(\overline{a} w_{AS} + (1 - \overline{a}) w_{AF}\right) + \pi_{\underline{\alpha}\underline{\beta}} \left(\underline{a} w_{AS} + (1 - \underline{a}) w_{AF}\right) \\
\text{s.t.} \quad & \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}} - g \geq \overline{U} && (11) \\
& \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}} - g \geq \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\underline{A}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}} && (12) \\
& \pi_{\overline{\alpha}\underline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{A}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{A}} - g \geq \pi_{\overline{\alpha}\underline{\beta}} V_{\underline{B}} + \pi_{\underline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\overline{\alpha}\overline{\beta}} V_{\overline{B}} + \pi_{\underline{\alpha}\underline{\beta}} V_{\underline{B}} && (13) \\
& V_{\overline{A}} \geq V_{\underline{B}} && (14) \\
& V_{\overline{B}} \geq V_{\underline{A}} && (15) \\
& V_{\overline{A}} \geq V_{\overline{B}} && (16) \\
& V_{\underline{A}} \geq V_{\underline{B}} && (17)
\end{aligned}$$

Lemma 4 (15) *is not binding.*

Proof. We can rewrite (12):

$$\pi_{\underline{\alpha}\overline{\beta}} (V_{\overline{B}} - V_{\underline{A}}) \geq g,$$

which implies that $V_{\overline{B}} > V_{\underline{A}}$. ■

Lemma 5 $w_{AS} > w_{AF}$.

Proof. Lemma 4 and (16) imply $V_{\overline{A}} \geq V_{\overline{B}} > V_{\underline{A}}$. Thus,

$$V_{\overline{A}} > V_{\underline{A}}. \tag{18}$$

Since $\overline{a} > \underline{a}$, this requires $w_{AS} > w_{AF}$. ■

LEMMA 6 (14) *is not binding.*

Proof. (18) and (17) imply $V_{\overline{A}} > V_{\underline{B}}$. ■

Given Lemmas 4 and 6, the first order conditions are:

$$\begin{aligned}
\frac{1}{U'(w_{AS})} &= \lambda_P + \lambda_B + \frac{\left(\lambda_{\underline{\alpha}\underline{\beta}} - \lambda_A \pi_{\underline{\alpha}\overline{\beta}}\right)\underline{a} + \lambda_{\overline{\alpha}\overline{\beta}}\overline{a}}{\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}}\right)\overline{a} + \pi_{\underline{\alpha}\underline{\beta}}\underline{a}}; &(19)\\
\frac{1}{U'(w_{AF})} &= \lambda_P + \lambda_B + \frac{\left(\lambda_{\underline{\alpha}\underline{\beta}} - \lambda_A \pi_{\underline{\alpha}\overline{\beta}}\right)(1-\underline{a}) + \lambda_{\overline{\alpha}\overline{\beta}}(1-\overline{a})}{\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}}\right)(1-\overline{a}) + \pi_{\underline{\alpha}\underline{\beta}}(1-\underline{a})}; &(20)\\
\frac{1}{U'(w_{BS})} &= \lambda_P + \lambda_A - \frac{\lambda_B\left[\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\underline{\alpha}\underline{\beta}}\right)\underline{b} + \pi_{\overline{\alpha}\overline{\beta}}\overline{b}\right] + \lambda_{\overline{\alpha}\overline{\beta}}\overline{b} + \lambda_{\underline{\alpha}\underline{\beta}}\underline{b}}{\pi_{\underline{\alpha}\overline{\beta}}\overline{b}}; &(21)\\
\frac{1}{U'(w_{BF})} &= \lambda_P + \lambda_A - \frac{\lambda_B\left[\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\underline{\alpha}\underline{\beta}}\right)(1-\underline{b}) + \pi_{\overline{\alpha}\overline{\beta}}\left(1-\overline{b}\right)\right] + \lambda_{\overline{\alpha}\overline{\beta}}\left(1-\overline{b}\right) + \lambda_{\underline{\alpha}\underline{\beta}}(1-\underline{b})}{\pi_{\underline{\alpha}\overline{\beta}}\left(1-\overline{b}\right)} &(22)
\end{aligned}$$

LEMMA 7 $w_{BS} > w_{BF}$.

Proof. Lemma 4 and (17) imply $V_{\overline{B}} > V_{\underline{A}} \geq V_{\underline{B}}$. Thus,

$$V_{\overline{B}} > V_{\underline{B}}.$$

Since $\overline{b} > \underline{b}$, this requires $w_{BS} > w_{BF}$. ■

LEMMA 8 $w_{AS} > w_{BF}$ *and* $w_{BS} > w_{AF}$.

Proof. Suppose that $w_{AS} \leq w_{BF}$. This implies that $w_{AF} < w_{AS} \leq w_{BF} < w_{BS}$ and contradicts (14), (16), and (17). Suppose that $w_{BS} \leq w_{AF}$. This implies that $w_{BF} < w_{BS} \leq w_{AF} < w_{AS}$ and contradicts (15). ■

LEMMA 9 (12) *is binding.*

Proof. Suppose (12) is not binding at an optimum. This implies that

$$\pi_{\underline{\alpha}\overline{\beta}}(V_{\overline{B}} - V_{\underline{A}}) > g.$$

Consider the following local deviation: $dw = (0, \Delta_{AF}, \Delta_{BS}, 0)$, such that $\Delta_{AF} > 0$, $\Delta_{BS} < 0$ and $\mathbf{D}V dw = 0$ (i.e., the change keeps V unchanged):

$$\pi_{\overline{\alpha}\underline{\beta}} U'(w_{AF})(1-\overline{a})\Delta_{AF} + \pi_{\underline{\alpha}\overline{\beta}} U'(w_{BS}) \overline{b}\Delta_{BS} + \pi_{\overline{\alpha}\overline{\beta}} U'(w_{AF})(1-\overline{a})\Delta_{AF} + \pi_{\underline{\alpha}\underline{\beta}} U'(w_{AF})(1-\underline{a})\Delta_{AF} = 0$$

i.e.,

$$\Delta_{BS} = -\Delta_{AF} \left(\frac{U'(w_{AF})}{\pi_{\underline{\alpha}\overline{\beta}} \overline{b} U'(w_{BS})} \right) [(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}})(1-\overline{a}) + \pi_{\underline{\alpha}\underline{\beta}}(1-\underline{a})].$$

By construction, (11) and (13) hold, since the change does not affect V and reduces V_B. If the change is small enough (12), (14) and (15) hold (recall that (14) and (15) have been shown to be nonbinding and (12) is nonbinding by hypothesis). (16) and (17) hold since the change increases $V_{\overline{A}}$ and $V_{\underline{A}}$ and decreases $V_{\overline{B}}$ and $V_{\underline{B}}$. The change in expected cost is:

$$\begin{aligned} &\pi_{\overline{\alpha}\underline{\beta}}(1-\overline{a})\Delta_{AF} + \pi_{\underline{\alpha}\overline{\beta}}\overline{b}\Delta_{BS} + \pi_{\overline{\alpha}\overline{\beta}}(1-\overline{a})\Delta_{AF} + \pi_{\underline{\alpha}\underline{\beta}}(1-\underline{a})\Delta_{AF} = \\ &\quad = \Delta_{AF}[(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}})(1-\overline{a}) + \pi_{\underline{\alpha}\underline{\beta}}(1-\underline{a})] \left(1 - \frac{U'(w_{AF})}{U'(w_{BS})}\right) < 0, \end{aligned} \tag{23}$$

with the inequality following from $w_{BS} > w_{AF}$ and the concavity of U. This contradicts the hypothesis that (12) is nonbinding. ■

Lemma 10 *If* (13) *is not binding,* (17) *is binding.*

Proof. Assume that (17) is not binding. From (21) and (22) we obtain

$$\begin{aligned} \frac{1}{U'(w_{BS})} &= \lambda_P + \lambda_A - \frac{\lambda_{\overline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\overline{\beta}}}; \\ \frac{1}{U'(w_{BF})} &= \lambda_P + \lambda_A - \frac{\lambda_{\overline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\overline{\beta}}}; \end{aligned}$$

which implies $w_{BS} = w_{BF}$, and contradicts Lemma 7. ■

We now consider three cases:

- Case 1: $\pi_{\overline{\alpha}\underline{\beta}} > \pi_{\underline{\alpha}\overline{\beta}}$;
- Case 2: $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$;

- Case 3: $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$.

Lemmas 11 to 15 refer to Case 1:

Lemma 11 (13) *is not binding at an optimum.*

Proof. Suppose it is binding. Since by Lemma 9 (12) is binding, it follows that $V_A = V_B$. Thus:

$$\pi_{\overline{\alpha}\underline{\beta}}(V_{\overline{A}} - V_{\underline{B}}) + \pi_{\overline{\alpha}\overline{\beta}}(V_{\overline{A}} - V_{\overline{B}}) + \pi_{\underline{\alpha}\underline{\beta}}(V_{\underline{A}} - V_{\underline{B}}) = \pi_{\underline{\alpha}\overline{\beta}}(V_{\overline{B}} - V_{\underline{A}}) \tag{24}$$

From (16) and (17)

$$\pi_{\overline{\alpha}\overline{\beta}}(V_{\overline{A}} - V_{\overline{B}}) + \pi_{\underline{\alpha}\underline{\beta}}(V_{\underline{A}} - V_{\underline{B}}) \geq 0. \tag{25}$$

From (24) and (25)

$$\pi_{\underline{\alpha}\overline{\beta}}(V_{\overline{B}} - V_{\underline{A}}) \geq \pi_{\overline{\alpha}\underline{\beta}}(V_{\overline{A}} - V_{\underline{B}}). \tag{26}$$

Given that $\pi_{\overline{\alpha}\underline{\beta}} > \pi_{\underline{\alpha}\overline{\beta}}$, from (26), $V_{\overline{B}} - V_{\underline{A}} > V_{\overline{A}} - V_{\underline{B}}$. But then, (16) implies $V_{\overline{B}} - V_{\underline{A}} > V_{\overline{B}} - V_{\underline{B}}$, which implies $V_{\underline{A}} < V_{\underline{B}}$ and a contradiction to (17) is obtained. ■

Lemma 12 (17) *is binding at an optimum.*

Proof. Immediate from Lemmas 10 and 11. ■

Lemma 13 *Suppose* (16) *is binding at an optimum. The optimal contract satisfies:*

$$\begin{aligned}
&\textit{For} \quad \underline{\varepsilon}_A(\varepsilon_B) \leq \varepsilon_A \leq \underset{\sim}{\varepsilon}_A(\varepsilon_B), w_{BF} \leq w_{AF} < w_{AS} < w_{BS} \\
&\textit{For} \quad \underset{\sim}{\varepsilon}_A(\varepsilon_B) < \varepsilon_A \leq \widetilde{\varepsilon}_A(\varepsilon_B), w_{AF} < w_{BF} < w_{AS} \leq w_{BS}; \\
&\textit{For} \quad \widetilde{\varepsilon}_A < \varepsilon_A \leq \bar{\varepsilon}_A(\varepsilon_B), w_{AF} < w_{BF} < w_{BS} < w_{AS}.
\end{aligned}$$

Proof. Define

$$\begin{aligned}
\underset{\sim}{L} &= \overline{a}\underline{b} - \underline{a}\overline{b} \\
\widetilde{L} &= \left(\overline{a}\underline{b} - \underline{a}\overline{b}\right) + \left(\overline{b} - \underline{b}\right) - \left(\overline{a} - \underline{a}\right).
\end{aligned}$$

Notice that $\underset{\sim}{L}$ and $\widetilde{L}$ are functions of ε_A and ε_B because the conditional probabilities of success are functions of ε_A and ε_B.

Recall from Lemma 12 that (17) is binding. From (16) and (17) with equality

$$U(w_{BS}) = \frac{\left[\underline{a}\left(1-\overline{b}\right)-\overline{a}(1-\underline{b})\right]U(w_{BF})}{\overline{a}\underline{b}-\underline{a}\overline{b}} - \frac{[\underline{a}(1-\overline{a})-\overline{a}(1-\underline{a})]U(w_{AF})}{\overline{a}\underline{b}-\underline{a}\overline{b}}; \tag{27}$$

$$U(w_{AS}) = \frac{\left[\underline{b}\left(1-\overline{b}\right)-\overline{b}(1-\underline{b})\right]U(w_{BF})}{\overline{a}\underline{b}-\underline{a}\overline{b}} - \frac{\left[(1-\overline{a})\underline{b}-(1-\underline{a})\overline{b}\right]U(w_{AF})}{\overline{a}\underline{b}-\underline{a}\overline{b}}. \tag{28}$$

Substituting (27) and (28) into (15) we obtain

$$\frac{U(w_{BF}) - U(w_{AF})}{\underset{\sim}{L}} < 0. \tag{29}$$

Given that $\underset{\sim}{L}$ is decreasing in ε_A and $\underset{\sim}{L} = 0$ when $\varepsilon_A = \underset{\sim}{\varepsilon}_A(\varepsilon_B)$, risk aversion implies that $\operatorname{sign}(w_{AF} - w_{BF}) = \operatorname{sign}\underset{\sim}{L}$.

From (27) and (28), $w_{BS} \geq w_{AS}$ if and only if

$$\widetilde{L} \times \frac{U(w_{BF}) - U(w_{AF})}{\underset{\sim}{L}} \geq 0. \tag{30}$$

By (29), (30) implies that $w_{BS} \geq w_{AS}$ if and only if $\widetilde{L} \leq 0$. Given that $\widetilde{L}$ is increasing in ε_A and $\widetilde{L} = 0$ when $\varepsilon_A = \widetilde{\varepsilon}_A(\varepsilon_B)$, risk aversion implies that $\operatorname{sign}(w_{AS} - w_{BS}) = \operatorname{sign}\widetilde{L}$. ■

Lemma 14 *If* (16) *is not binding, the optimal contract satisfies:*

$$\begin{aligned} &\textit{For} \quad \underline{\varepsilon}_A(\varepsilon_B) \leq \varepsilon_A \leq \underset{\sim}{\varepsilon}_A(\varepsilon_B) : w_{AS} < w_{BS}; \\ &\textit{For} \quad \underset{\sim}{\varepsilon}_A(\varepsilon_B) < \varepsilon_A < \widetilde{\varepsilon}_A(\varepsilon_B) : w_{AF} < w_{BF} < w_{AS} < w_{BS}. \end{aligned}$$

Proof. By Lemma 5, $w_{AS} > w_{AF}$. From (19), (20) and (21) this implies that $w_{BS} > w_{AS}$. From (17) with equality

$$U(w_{BS}) = \frac{\underline{a}U(w_{AS}) + (1-\underline{a})U(w_{AF}) - (1-\underline{b})U(w_{BF})}{\underline{b}}. \tag{31}$$

Substituting (31) into (16), (16) is not binding if and only if

$$U(w_{AS})\underset{\sim}{L} > \left(\underline{b} - \overline{b}\right)U(w_{BF}) + \left[\left(\overline{b} - \underline{b}\right) - \left(\underline{a}\overline{b} - \overline{a}\underline{b}\right)\right]U(w_{AF}). \tag{32}$$

If $\underset{\sim}{L} < 0$, condition (32) is equivalent to

$$-\frac{\bar{b}-\underline{b}}{\underset{\sim}{L}}U\left(w_{BF}\right) - \frac{\left(\underline{a}\bar{b}-\bar{a}\underline{b}\right)-\left(\bar{b}-\underline{b}\right)}{\underset{\sim}{L}}U\left(w_{AF}\right) > U\left(w_{AS}\right). \tag{33}$$

Substitute (31) into (15) and recalling from Lemma (4) that (15) is not binding we obtain

$$U\left(w_{AS}\right) > \frac{1}{\underline{a}}U\left(w_{BF}\right) - \frac{\left(1-\underline{a}\right)}{\underline{a}}U\left(w_{AF}\right). \tag{34}$$

A necessary condition for (33) and (34) to both hold, is that the left hand side of (33) is greater than the right hand side of (34). Straightforward algebra shows that this inequality holds if and only if $w_{BF} > w_{AF}$. ■

Lemma 15 *If* $\varepsilon_A \in \left[\widetilde{\varepsilon}_A\left(\varepsilon_B\right), \overline{\varepsilon}_A\left(\varepsilon_B\right)\right]$, (16) *is binding*

Proof. Notice first that for $\varepsilon_A \in \left[\widetilde{\varepsilon}_A\left(\varepsilon_B\right), \overline{\varepsilon}_A\left(\varepsilon_B\right)\right]$, $\underset{\sim}{L} < 0$. Suppose now that (16) is not binding. From Lemma 14, $w_{BS} > w_{AS}$. From (31),

$$U\left(w_{BS}\right) > U\left(w_{AS}\right)$$

is equivalent to

$$U\left(w_{AS}\right) > \frac{\left(1-\underline{b}\right)}{\left(\underline{a}-\underline{b}\right)}U\left(w_{BF}\right) - \frac{\left(1-\underline{a}\right)}{\left(\underline{a}-\underline{b}\right)}U\left(w_{AF}\right). \tag{35}$$

Notice that since $\underset{\sim}{L} < 0$ (33) holds and a necessary condition for (35) to also hold is that the left hand side of (33) is greater than the right hand side of (35) and this is equivalent to

$$\widetilde{L}\left(U\left(w_{BF}\right) - U\left(w_{AF}\right)\right) < 0. \tag{36}$$

Notice now that in Lemma 14 it was established that $\underset{\sim}{L} < 0$ and (16) not binding imply $w_{BF} > w_{AF}$. This implies that a necessary condition for (36) to hold is $\widetilde{L} < 0$. But, because this requires $\varepsilon_A < \widetilde{\varepsilon}_A\left(\varepsilon_B\right)$, a contradiction is obtained. ■

Lemmas 16 and 17 refer to Case 2:

Lemma 16 *Suppose* $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$. *If* (13) *is binding, then wages are as in Lemma 13.*

Proof. From (12) and (13) with equality and from $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$:

$$\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}}\right)\left(V_{\overline{A}} - V_{\overline{B}}\right) + \left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\underline{\alpha}\underline{\beta}}\right)\left(V_{\underline{A}} - V_{\underline{B}}\right) = 0. \tag{37}$$

Given that (16) and (17) imply $\left(V_{\overline{A}} - V_{\overline{B}}\right) \geq 0$ and $\left(V_{\underline{A}} - V_{\underline{B}}\right) \geq 0$, (37) implies that (16) and (17) hold with equality. The rest of the proof follows along the lines of Lemma (13). ■

LEMMA 17 *Suppose* $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$. *If* (13) *is not binding, then wages are as in Lemma 14.*

Proof. If (13) is not binding, (17) is binding by Lemma 10. From $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$, (12) with equality and (13) with strict inequality we obtain

$$\left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\overline{\alpha}\overline{\beta}}\right)\left(V_{\overline{A}} - V_{\overline{B}}\right) + \left(\pi_{\overline{\alpha}\underline{\beta}} + \pi_{\underline{\alpha}\underline{\beta}}\right)\left(V_{\underline{A}} - V_{\underline{B}}\right) > 0. \tag{38}$$

Because (17) is binding, $V_{\underline{A}} - V_{\underline{B}} = 0$ and from (38) we obtain $V_{\overline{A}} - V_{\overline{B}} > 0$, i.e., that (16) is not binding. The rest of the proof follows along the lines of Lemma 14. ■

Lemmas 18 to 21 refer to Case 3.

LEMMA 18 *Suppose* $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$. (16) *and* (17) *cannot be binding simultaneously.*

Proof. Suppose (16) and (17) are binding. Substituting them in (12) and (13), implies that (12) and (13) hold if and only if $\pi_{\overline{\alpha}\underline{\beta}} > \pi_{\underline{\alpha}\overline{\beta}}$ and a contradiction arises. ■

LEMMA 19 *Suppose* $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$. *(1) If* $a > 1/2$, $\varepsilon_A > \widetilde{\varepsilon}_A(\varepsilon_B)$. *(2) If* $a \leq 1/2$, $\varepsilon_A < \underset{\sim}{\varepsilon}_A(\varepsilon_B)$.

Proof. If $a > \frac{1}{2}$, $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$ is equivalent to

$$\varepsilon_A > \frac{\Pr\left(\overline{\beta}\right) - a}{1 - 2a} > \widetilde{\varepsilon}_A(\varepsilon_B).$$

If $a < \frac{1}{2}$, $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$ is equivalent to

$$\varepsilon_A < \frac{\Pr\left(\overline{\beta}\right) - a}{1 - 2a} < \underset{\sim}{\varepsilon}_A(\varepsilon_B).$$

■

LEMMA 20 *Suppose* $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$. *If* (13) *is not binding, then wages are as in Lemma 14.*

Proof. Same as in Lemma 17. ∎

Lemma 21 *Suppose $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$. If (13) is binding the optimal contract satisfies*

$$\begin{aligned} For\ \underline{\varepsilon}_A(\varepsilon_B) &\leq \varepsilon_A < \underset{\sim}{\varepsilon}_A(\varepsilon_B): w_{BF} < w_{AF} < w_{AS} < w_{BS}; \\ For\ \widetilde{\varepsilon}_A(\varepsilon_B) &< \varepsilon_A \leq \bar{\varepsilon}_A(\varepsilon_B): w_{AF} < w_{BF} < w_{BS} < w_{AS}. \end{aligned}$$

Proof. From (12) and (13) binding we obtain

$$aU(w_{AS}) + (1-a)U(w_{AF}) = bU(w_{BS}) + (1-b)U(w_{BF}). \tag{39}$$

Assume first $a < \frac{1}{2}$. From (39) we obtain that

$$aU(w_{AS}) = bU(w_{BS}) + (1-b)U(w_{BF}) - (1-a)U(w_{AF}). \tag{40}$$

Substituting (40) into (15) and (17)

$$bU(w_{BS})G > \left[\varepsilon_A \Pr\left(\overline{\beta}\right) - \varepsilon_B \Pr(\underline{\alpha})\right](1-b)U(w_{BF}) + \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)(1-a)U(w_{AF});$$

$$bU(w_{BS})H \geq \left[(1-\varepsilon_B)\Pr(\underline{\alpha}) - \varepsilon_A \Pr\left(\underline{\beta}\right)\right](1-b)U(w_{BF}) - \Pr\left(\underline{\beta}\right)(1-2\varepsilon_A)(1-a)U(w_{AF});$$

where

$$\begin{aligned} G &= (1-\varepsilon_B)\Pr(\underline{\alpha}) - \varepsilon_A \Pr\left(\overline{\beta}\right); \\ H &= \varepsilon_A \Pr\left(\underline{\beta}\right) - \varepsilon_B \Pr(\underline{\alpha}). \end{aligned}$$

Simple algebra shows that G is strictly positive and that $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$ and $a \leq \frac{1}{2}$ imply $H < 0$. Therefore

$$bU(w_{BS}) > \frac{\left[\varepsilon_A \Pr\left(\overline{\beta}\right) - \varepsilon_B \Pr(\underline{\alpha})\right](1-b)U(w_{BF}) + \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)(1-a)U(w_{AF})}{G}; \tag{41}$$

$$bU(w_{BS}) < \frac{\left[(1-\varepsilon_B)\Pr(\underline{\alpha}) - \varepsilon_A \Pr\left(\underline{\beta}\right)\right](1-b)U(w_{BF}) - \Pr\left(\underline{\beta}\right)(1-2\varepsilon_A)(1-a)U(w_{AF})}{H} \tag{42}$$

A necessary condition for (41) and (42) is that the right hand side of (42) is greater than the right hand side of (41) and this implies $w_{BF} < w_{AF}$. From (39) we obtain

$$w_{BF} < w_{AF} < w_{AS} < w_{BS}.$$

Assume now $a > \frac{1}{2}$. From (39) we obtain that

$$(1-a)\,U(w_{AF}) = bU(w_{BS}) + (1-b)\,U(w_{BF}) - aU(w_{AS})\,. \tag{43}$$

Substituting (43) into (15) and (16)

$$\begin{aligned}&(1-b)\,U(w_{BF})\,J < \\ &\left[(1-\varepsilon_B)\Pr(\underline{\alpha}) - (1-\varepsilon_A)\Pr\left(\overline{\beta}\right)\right] bU(w_{BS}) + \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)\,aU(w_{AS})\,;\end{aligned}$$

$$\begin{aligned}&(1-b)\,U(w_{BF})\,K > \\ &\left[(1-\varepsilon_B)\Pr(\overline{\alpha}) - \varepsilon_A\Pr\left(\overline{\beta}\right)\right] bU(w_{BS}) - \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)\,aU(w_{AS})\,;\end{aligned}$$

where

$$\begin{aligned} J &= (1-\varepsilon_A)\Pr\left(\overline{\beta}\right) - \varepsilon_B\Pr(\underline{\alpha})\,; \\ K &= \varepsilon_A\Pr\left(\overline{\beta}\right) - \varepsilon_B\Pr(\overline{\alpha})\,. \end{aligned}$$

Simple algebra shows that J is strictly positive and that $\pi_{\overline{\alpha}\underline{\beta}} < \pi_{\underline{\alpha}\overline{\beta}}$ and $a > \frac{1}{2}$ imply $K > 0$. Therefore

$$(1-b)\,U(w_{BF}) < \frac{\left[(1-\varepsilon_B)\Pr(\underline{\alpha}) - (1-\varepsilon_A)\Pr\left(\overline{\beta}\right)\right] bU(w_{BS}) + \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)\,aU(w_{AS})}{J} \tag{44}$$

$$(1-b)\,U(w_{BF}) > \frac{\left[(1-\varepsilon_B)\Pr(\overline{\alpha}) - \varepsilon_A\Pr\left(\overline{\beta}\right)\right] bU(w_{BS}) - \Pr\left(\overline{\beta}\right)(1-2\varepsilon_A)\,aU(w_{AS})}{K}. \tag{45}$$

A necessary condition for (44) and (45) is that the right hand side of (44) is greater than the right hand side of (45) and this implies $w_{BS} < w_{AS}$. From (39) we obtain

$$w_{AF} < w_{BF} < w_{BS} < w_{AS}.$$

■

Proof of Proposition 1 Suppose that $w = (w_{AF}, w_{AS}, w_{BF}, w_{BS})$ solves (P) and $w_{AF} = w_{BF} = w_F$ and $w_S = w_{AS} = w_{BS}$. For (IC-$\alpha\beta$) to hold, it must be the case that $w_S \geq w_F$. Now, if $w_S = w_F$, (IC-A) and (IC-B) would not hold. Therefore, if w implements the optimal action profile $\hat{s}$, $w_S > w_F$ and (IC-$\alpha\beta$) hold with slack.

Assume first that $a > b$. Then, $aU(w_S)+(1-a)U(w_F) > bU(w_S)+(1-b)U(w_F)$, so if (IC-$A$) holds, then (IC-$B$) holds with slack. Consider now contract w', with $w'_{AF} = w_F$, $w'_{AS} = w_S - \Delta_A$, $w'_{BF} = w_F + \Delta_B$ and $w'_{BS} = w_S$, where Δ_A and Δ_B are such that $V(w) = V(w')$, so that (PC) holds for w'. Further, for Δ_A and Δ_B small enough, (IC-B) and (IC-$\alpha\beta$) will also hold, since they held with slack for w. Finally, (IC-A) will hold as well, since $V(w') = V(w) \geq aU(w_S) + (1-a)U(w_F) > aU(w'_{AS}) + (1-a)U(w_F)$. Therefore, w' implements the optimal action profile $\hat{s}$. It only remains to be checked that it lowers the expected cost. Letting $h(x|\hat{s})$ denote the probability of public history x given the optimal action profile $\hat{s}$, it follows from $V(w) = V(w')$ that for Δ_A and Δ_B small enough, $h(BS|\hat{s})U'(w_F)\Delta_B = h(AS|\hat{s})U'(w_S)\Delta_A$, so

$$\Delta_B = \Delta_A \frac{h(AS|\hat{s})U'(w_S)}{h(BS|\hat{s})U'(w_F)} \tag{46}$$

Therefore, the change in the expected cost is

$$h(BS|\hat{s})\Delta_B - h(AS|\hat{s})\Delta_A = \Delta_A h(AS|\hat{s}) \left(\frac{U'(w_S)}{U'(w_F)} - 1 \right) < 0, \tag{47}$$

since risk aversion implies that $U'(w_S) < U'(w_F)$.

Assume now that $a = b$ and that $\varepsilon_A < \varepsilon_B$. This implies that the manager will base his decision only on signal α, the most precise signal. Consider now a change $dw = (0, 0, \Delta_F, \Delta_S)$, such that $\Delta_F > 0$, $\Delta_S < 0$ and $\mathbf{D}V dw = 0$ (i.e., the change keeps $V(w)$ unchanged). By definition of dw, (PC) and (IC-A) are unchanged. Further, since (IC-$\alpha\beta$) hold with slack, they will continue to hold for dw small enough. Now we check if (IC-B) still holds. By construction of dw

$$\pi_{\underline{\alpha}\overline{\beta}} \mathbf{D}V_{\overline{B}} dw + \pi_{\underline{\alpha}\underline{\beta}} \mathbf{D}V_{\underline{B}} dw = 0,$$

i.e,

$$\mathbf{D}V_{\underline{B}} dw = -\left(\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}} \right) \mathbf{D}V_{\overline{B}} dw.$$

Consider now

$$\mathbf{D}V_B dw = \Pr\left(\underline{\beta}\right)\mathbf{D}V_{\underline{B}}dw + \Pr\left(\overline{\beta}\right)\mathbf{D}V_{\overline{B}}dw = -\Pr\left(\underline{\beta}\right)\left(\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\right)\mathbf{D}V_{\overline{B}}dw + \Pr\left(\overline{\beta}\right)\mathbf{D}V_{\overline{B}}dw$$

$$= \mathbf{D}V_{\overline{B}}dw\left(\Pr\left(\overline{\beta}\right) - \Pr\left(\underline{\beta}\right)\left(\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\right)\right) = 0,$$

since

$$\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}} = \frac{\Pr(\underline{\alpha})\Pr(\overline{\beta})}{\Pr(\underline{\alpha})\Pr(\underline{\beta})} = \frac{\Pr(\overline{\beta})}{\Pr(\underline{\beta})}.$$

Therefore, (IC-B) will hold as well. Finally, it only remains to check that dw decreases the expected cost of the contract. Letting $h(x|\hat{s})$ denote the probability of public history x given the optimal action profile $\hat{s}$, it follows from $V(w) = V(w')$ that for Δ_F and Δ_S small enough, $h(BS|\hat{s})U'(w_S)\Delta_S + h(BF|\hat{s})U'(w_F)\Delta_F = 0$, so

$$\Delta_S = -\Delta_F \frac{h(BF|\hat{s})U'(w_F)}{h(BS|\hat{s})U'(w_S)}.$$

Therefore, the change in the expected cost is

$$h(BS|\hat{s})\Delta_S + h(BF|\hat{s})\Delta_F = \Delta_F h(BF|\hat{s})\left(1 - \frac{U'(w_F)}{U'(w_S)}\right) < 0,$$

since risk aversion implies that $U'(w_S) < U'(w_F)$.

Finally, we show that $a = b$ and $\varepsilon_A = \varepsilon_B$ imply $w_{AS} = w_{BS} > w_{AF} = w_{BF}$. Notice first that $a = b$ and $\varepsilon_A = \varepsilon_B$ imply $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$. By Lemma 9 we know that (12) is binding. Assume first that (13) is binding. In Lemma 16 we have shown that (12) and (13) binding and $\pi_{\overline{\alpha}\underline{\beta}} = \pi_{\underline{\alpha}\overline{\beta}}$, imply that (16) and (17) are also binding. From (12) and (13) with equality we obtain

$$aU\left(w_{AS}\right) + (1-a)\,U\left(w_{AF}\right) = aU\left(w_{BS}\right) + (1-a)\,U\left(w_{BF}\right), \tag{48}$$

and from (17) with equality we obtain

$$\underline{a}U\left(w_{AS}\right) + (1-\underline{a})\,U\left(w_{AF}\right) = \underline{a}U\left(w_{BS}\right) + (1-\underline{a})\,U\left(w_{BF}\right). \tag{49}$$

Given that $a > \underline{a}$, (48) and (49) hold if and only if $U\left(w_{AS}\right) = U\left(w_{BS}\right)$ and $U\left(w_{AF}\right) = U\left(w_{BF}\right)$ which imply $w_{AS} = w_{BS}$ and $w_{AF} = w_{BF}$.

Assume now that (13) is not binding. From (13) not binding and (12) binding we obtain

$$aU\left(w_{AS}\right)+(1-a)\,U\left(w_{AF}\right)>aU\left(w_{BS}\right)+(1-a)\,U\left(w_{BF}\right). \tag{50}$$

By Lemma 10 we know that (17) is binding and this implies (49). Notice now that a necessary condition for (49) and (50) to both hold is that $U\left(w_{AS}\right)>U\left(w_{BS}\right)$ which implies $w_{AS}>w_{BS}$. Recall now that in Lemma 17 we showed that if (13) and (17) are both binding, (16) cannot be binding. But Lemma 14 shows that this implies $U\left(w_{BS}\right)>U\left(w_{AS}\right)$ and a contradiction arises. ■

Proof of Proposition 2

Notice that $\varepsilon_A=\varepsilon_B$ implies that $\pi_{\overline{\alpha}\underline{\beta}}>\pi_{\underline{\alpha}\overline{\beta}}$ and that $\underset{\sim}{\varepsilon}_A\left(\varepsilon_B\right)<\varepsilon_A<\widetilde{\varepsilon}_A\left(\varepsilon_B\right)$. The proof follows from Lemmas 13 and 14 ■

Proof of Proposition 3

The proof follows from Lemmas 13, 14, 15, 16, 17, 20, and 21.■

A.5 Proof of Proposition 4

We characterize the salary for part (1). Since in the following we never use the ex-ante efficiency of A $(a>b)$, the proof for part (2) is symmetric.

If the optimal decision is sensitive only to α, the optimal contract is a solution to:

$$\min_{w\in\mathbb{R}^4}\quad \pi_{\overline{\alpha}\underline{\beta}}[\overline{a}w_{AS}+(1-\overline{a})w_{AF}]+\pi_{\underline{\alpha}\overline{\beta}}[\overline{b}w_{BS}+(1-\overline{b})w_{BF}]+$$

$$+\pi_{\overline{\alpha}\overline{\beta}}[\underline{a}w_{AS}+(1-\underline{a})w_{AF}]+\pi_{\underline{\alpha}\underline{\beta}}[\underline{b}w_{BS}+(1-\underline{b})w_{BF}]$$

$$s.t.\quad \pi_{\overline{\alpha}\underline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\overline{\beta}}V_{\overline{B}}+\pi_{\overline{\alpha}\overline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\underline{\beta}}V_{\underline{B}}-g\geq\overline{U} \tag{51}$$

$$\pi_{\overline{\alpha}\underline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\overline{\beta}}V_{\overline{B}}+\pi_{\overline{\alpha}\overline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\underline{\beta}}V_{\underline{B}}-g\geq\pi_{\overline{\alpha}\underline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\overline{\beta}}V_{\underline{A}}+\pi_{\overline{\alpha}\overline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\underline{\beta}}V_{\underline{A}} \tag{52}$$

$$\pi_{\overline{\alpha}\underline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\overline{\beta}}V_{\overline{B}}+\pi_{\overline{\alpha}\overline{\beta}}V_{\overline{A}}+\pi_{\underline{\alpha}\underline{\beta}}V_{\underline{B}}-g\geq\pi_{\overline{\alpha}\underline{\beta}}V_{\underline{B}}+\pi_{\underline{\alpha}\overline{\beta}}V_{\overline{B}}+\pi_{\overline{\alpha}\overline{\beta}}V_{\overline{B}}+\pi_{\underline{\alpha}\underline{\beta}}V_{\underline{B}} \tag{53}$$

$$V_{\overline{A}}\geq V_{\underline{B}} \tag{54}$$

$$V_{\overline{B}}\geq V_{\underline{A}} \tag{55}$$

$$V_{\overline{A}}\geq V_{\overline{B}} \tag{56}$$

$$V_{\underline{B}}\geq V_{\underline{A}} \tag{57}$$

We characterize the optimal contract through a series of Lemmas.

Lemma 22 *(1)* (54) *and* (56) *cannot be both binding. (2)* (55) *and* (57) *cannot be both binding.*

Proof. (1) Rewrite (53) as :

$$\pi_{\overline{\alpha}\underline{\beta}}(V_{\overline{A}} - V_{\underline{B}}) + \pi_{\overline{\alpha}\overline{\beta}}(V_{\overline{A}} - V_{\overline{B}}) \geq g > 0. \tag{58}$$

If (54) and (56) are both binding, the left hand side of (58) is equal to 0 and a contradiction is obtained.

(2) Rewrite (52) as

$$\pi_{\underline{\alpha}\overline{\beta}}(V_{\overline{B}} - V_{\underline{A}}) + \pi_{\underline{\alpha}\underline{\beta}}(V_{\underline{B}} - V_{\underline{A}}) \geq g > 0. \tag{59}$$

If (55) and (57) are both binding, the left hand side of (59) is equal to 0 and a contradiction is obtained. ■

Lemma 23 $w_{AS} \geq w_{AF}$.

Proof. From (56) and (55), $V_{\overline{A}} \geq V_{\underline{A}}$. Since $\overline{a} > \underline{a}$, this requires $w_{AS} \geq w_{AF}$. ■

Lemma 24 $w_{BS} \geq w_{BF}$.

Proof. Suppose that $w_{BS} < w_{BF}$. This implies $V_{\underline{B}} > V_{\overline{B}}$. Therefore, (54) implies that (56) is not binding and (55) implies that (57) is not binding.

Consider a change $dw = (0, 0, \Delta_{BF}, \Delta_{BS})$ such that $\Delta_{BF} < 0$, $\Delta_{BS} > 0$, and $\mathbf{D}V dw = 0$, i.e., the change keeps V unchanged[10]. Given that $\Delta_{BF} < 0$, $\Delta_{BS} > 0$, this requires

$$\mathbf{D}V_{\overline{B}} dw > 0 \tag{60}$$

$$\mathbf{D}V_{\underline{B}} dw < 0. \tag{61}$$

In the following, we will show that there exist Δ_{BF} and Δ_{BS} sufficiently small in absolute value that (i) satisfy constraints (51)-(57); (ii) the expected cost of the contract is lower.

(i) By definition of dw, (51) and (52) are unchanged. From (60) and (61), (54) and (55) also hold. Moreover, since (56) and (57) are not binding, they also hold for Δ_{BF} and Δ_{BS} sufficiently small in absolute value.

[10]$\mathbf{D}$ denotes the Jacobian operator.

Notice now that by construction of dw:

$$\pi_{\underline{\alpha}\overline{\beta}}\mathbf{D}V_{\overline{B}}dw+\pi_{\underline{\alpha}\underline{\beta}}\mathbf{D}V_{\underline{B}}dw=0, \tag{62}$$

and a sufficient condition for (53) to hold is that

$$\pi_{\overline{\alpha}\underline{\beta}}\mathbf{D}V_{\underline{B}}dw+\pi_{\overline{\alpha}\overline{\beta}}\mathbf{D}V_{\overline{B}}dw=0 \tag{63}$$

because this implies that the right hand side of (53) is unchanged (the left hand side is unchanged by construction of dw). From (62)

$$\mathbf{D}V_{\underline{B}}dw=-\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\mathbf{D}V_{\overline{B}}dw. \tag{64}$$

Substituting (64) into (63) we obtain

$$\begin{aligned}-\pi_{\overline{\alpha}\underline{\beta}}\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\mathbf{D}V_{\overline{B}}dw+\pi_{\overline{\alpha}\overline{\beta}}\mathbf{D}V_{\overline{B}}dw &= \mathbf{D}V_{\overline{B}}dw\left(\pi_{\overline{\alpha}\overline{\beta}}-\pi_{\overline{\alpha}\underline{\beta}}\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\right)\\ &= \pi_{\overline{\alpha}\underline{\beta}}\mathbf{D}V_{\overline{B}}dw\left(\frac{\pi_{\overline{\alpha}\overline{\beta}}}{\pi_{\overline{\alpha}\underline{\beta}}}-\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}\right)=0\end{aligned} \tag{65}$$

with (65) following from the independence of A and B and of α and β that imply that

$$\frac{\pi_{\overline{\alpha}\overline{\beta}}}{\pi_{\overline{\alpha}\underline{\beta}}}=\frac{\Pr\left(\overline{\alpha},\overline{\beta}\right)}{\Pr\left(\overline{\alpha},\underline{\beta}\right)}=\frac{\Pr\left(\overline{\beta}\right)}{\Pr\left(\underline{\beta}\right)}=\frac{\Pr\left(\underline{\alpha},\overline{\beta}\right)}{\Pr\left(\underline{\alpha},\underline{\beta}\right)}=\frac{\pi_{\underline{\alpha}\overline{\beta}}}{\pi_{\underline{\alpha}\underline{\beta}}}.$$

(ii) Given that dw is such that $\mathbf{D}V dw=0$

$$\left(\pi_{\underline{\alpha}\overline{\beta}}\overline{b}+\pi_{\underline{\alpha}\underline{\beta}}\underline{b}\right)U'\left(w_{BS}\right)\Delta_{BS}+\left(\pi_{\underline{\alpha}\overline{\beta}}\left(1-\overline{b}\right)+\pi_{\underline{\alpha}\underline{\beta}}\left(1-\underline{b}\right)\right)U'\left(w_{BF}\right)\Delta_{BF}=0$$

which implies that

$$\Delta_{BS}=-\frac{\left(\pi_{\underline{\alpha}\overline{\beta}}\left(1-\overline{b}\right)+\pi_{\underline{\alpha}\underline{\beta}}\left(1-\underline{b}\right)\right)U'\left(w_{BF}\right)\Delta_{BF}}{\left(\pi_{\underline{\alpha}\overline{\beta}}\overline{b}+\pi_{\underline{\alpha}\underline{\beta}}\underline{b}\right)U'\left(w_{BS}\right)} \tag{66}$$

The change in the expected salary is

$$\left(\pi_{\underline{\alpha}\overline{\beta}}\overline{b}+\pi_{\underline{\alpha}\underline{\beta}}\underline{b}\right)\Delta_{BS}+\left(\pi_{\underline{\alpha}\overline{\beta}}\left(1-\overline{b}\right)+\pi_{\underline{\alpha}\underline{\beta}}\left(1-\underline{b}\right)\right)\Delta_{BF}. \tag{67}$$

Substituting (66) into (67) shows that the change in the expected salary is negative. ■

LEMMA 25 $w_{BF} \geq w_{BS}$.

Proof. The proof is symmetric to the proof of Lemma 24 with the only difference arising from the fact that in this case (56) implies that (54) is not binding and (57) implies that (55) is not binding. ■

LEMMA 26 $w_{AS} > w_{BS} = w_{BF} > w_{AF}$.

Proof. $w_{BS} = w_{BF}$ follows from Lemmas 24 and 25. From Lemma 22, $V_{\overline{A}} > V_{\underline{B}} = V_{\overline{B}}$, since (54) and (56) cannot be both binding. Therefore, (54) and (57) imply $V_{\overline{A}} > V_{\underline{B}} \geq V_{\underline{A}}$, which implies $V_{\overline{A}} > V_{\underline{A}}$, which implies $w_{AS} > w_{AF}$. Finally, (54) to (57) imply that $w_{AS} > w_{BS} = w_{BF} > w_{AF}$. ■

Proof of Proposition 4

Immediate from Lemma 26.

■

A.6 Proof of Proposition 5

The optimal contract is a solution to

$$\min_{w \in \mathbb{R}^4} \quad \pi_{\overline{\alpha}}[\overline{a} w_{AS} + (1-\overline{a}) w_{AF}] + \pi_{\underline{\alpha}}[\overline{b} w_{BS} + (1-\overline{b}) w_{BF}] \tag{68}$$

$$\text{s.t.} \quad \pi_{\overline{\alpha}} V_{\overline{A}} + \pi_{\underline{\alpha}} V_{\overline{B}} - g \geq \overline{U} \tag{69}$$

$$\pi_{\overline{\alpha}} V_{\overline{A}} + \pi_{\underline{\alpha}} V_{\overline{B}} - g \geq \pi_{\overline{\alpha}} V_{\overline{A}} + \pi_{\underline{\alpha}} V_{\underline{A}} \tag{70}$$

$$\pi_{\overline{\alpha}} V_{\overline{A}} + \pi_{\underline{\alpha}} V_{\overline{B}} - g \geq \pi_{\overline{\alpha}} V_{\underline{B}} + \pi_{\underline{\alpha}} V_{\overline{B}} \tag{71}$$

$$V_{\overline{A}} \geq V_{\underline{B}} \tag{72}$$

$$V_{\overline{B}} \geq V_{\underline{A}} \tag{73}$$

Because of the assumptions on the utility function, the solution to this problem is interior and satisfies the following first order conditions:

$$\frac{1}{U'(w_{AS})} = \frac{(\lambda_P + \lambda_B + \lambda_1)(1-\varepsilon) - (\lambda_A + \lambda_2)\varepsilon}{(1-\varepsilon)} \tag{74}$$

$$\frac{1}{U'(w_{AF})} = \frac{(\lambda_P + \lambda_B + \lambda_1)\varepsilon - (\lambda_A + \lambda_2)(1-\varepsilon)}{\varepsilon} \tag{75}$$

$$\frac{1}{U'(w_{BS})} = \frac{(\lambda_P + \lambda_A + \lambda_2)(1-\varepsilon) - (\lambda_B + \lambda_1)\varepsilon}{(1-\varepsilon)} \tag{76}$$

$$\frac{1}{U'(w_{BF})} = \frac{(\lambda_P + \lambda_A + \lambda_2)\varepsilon - (\lambda_B + \lambda_1)(1-\varepsilon)}{\varepsilon} \tag{77}$$

We prove the Proposition through a sequence of Lemmas.

Lemma 27 (70) *and* (73) *cannot hold with slack simultaneously.*

Proof. Suppose that (70) and (73) both hold with slack. From (74)-(77) we obtain

$$w_{AF} = w_{AS} = w_{BF} = w_{BS}$$

and a contradiction to (70) and (71) arises. ■

Lemma 28 (72) *and* (73) *hold with slack.*

Proof. If (72) is binding, (71) is violated. If (73) is binding, (70) is violated. ■

Lemma 29 (70) *is binding.*

Proof. Immediate from Lemmas 27 and 28 ■

Lemma 30 (71) *is binding.*

Proof. Suppose that (71) holds with slack. From (74)-(77) we obtain

$$w_{AF} \leq w_{AS} \leq w_{BF} = w_{BS}$$

which violates (72). ■

Lemma 31 $U(w_{AS}) - U(w_{BF}) < U(w_{BS}) - U(w_{AF})$.

Proof. From Lemmas 29 and 30 we have

$$aU(w_{AS}) + (1-a)U(w_{AF}) = (1-a)U(w_{BS}) + aU(w_{BF}),$$

which is equivalent to

$$a[U(w_{AS}) - U(w_{BF})] = (1-a)[U(w_{BS}) - U(w_{AF})]. \tag{78}$$

A necessary condition for (78) to hold is

$$U(w_{AS}) - U(w_{BF}) < U(w_{BS}) - U(w_{AF}).$$

■

LEMMA 32 *$w_{AS} > w_{AF}$ and $w_{BS} > w_{BF}$.*

Proof. Immediate from Lemmas 28, 29 and 30 and from (74)-(77). ■

LEMMA 33 *$w_{AF} < w_{BF} \leq w_{AS} < w_{BS}$.*

Proof. From Lemmas 28, 29 and 30 and from (74)-(77) we obtain that

$$w_{BS} \geq w_{AS} \iff w_{BF} \geq w_{AF} \tag{79}$$

$$w_{BS} \leq w_{AS} \iff w_{BF} \leq w_{AF} \tag{80}$$

Suppose that $w_{BS} = w_{AS}$. From (79) and (80) this holds if and only if $w_{BF} = w_{AF}$. This in turn implies that (70) and (71) cannot be both binding, a contradiction to either of Lemmas 29 and 30.

Suppose now that $w_{BS} < w_{AS}$. From (79) and (80) this holds if and only if $w_{BF} < w_{AF}$. By Lemma 32, this implies one of the following

$$w_{BF} < w_{BS} \leq w_{AF} < w_{AS} \tag{81}$$

$$w_{BF} < w_{AF} < w_{BS} < w_{AS} \tag{82}$$

Notice that (81) contradicts Lemma 31 and (82) implies that (70) and (71) cannot both be binding, therefore contradicting at least one of Lemmas 29 and 30.

Suppose therefore that $w_{BS} > w_{AS}$. From (79) and (80) this holds if and only if $w_{BF} > w_{AF}$. By Lemma 32, this implies one of the following

$$w_{AF} < w_{AS} \quad < \quad w_{BF} < w_{BS} \tag{83}$$

$$w_{AF} < w_{BF} \quad \leq \quad w_{AS} < w_{BS} \tag{84}$$

Because (83) violates (72) the only possible ranking is the one in (84) and the result follows. ■

Proof of Proposition 5

Immediate from Lemma 33.