A Theory of Risk Management with Applications to Executive Compensation and Earnings Management*

JEREMY BERTOMEU†

Abstract

This paper presents a theory of risk management in which the choices of managers over effort and risk are imperfectly monitored by outsiders. In a principal-agent framework, hedging can reduce extraneous noise in the variables outsiders observe or create opportunities for self-dealing behavior. The model reproduces several empirical features commonly described as anomalous, in an optimal-contract setting. In equilibrium, the hedged distribution of output is hump-shaped and asymmetric: first, for outputs close to the mode of the distribution, managers are more likely to do well and less likely to do poorly; second, managers hedge against large gains and increase the likelihood of large losses. The model accounts for the prevalence of linear compensation schemes and the relatively low performance-pay coefficients observed in managerial jobs. A simple linear contract is optimal over states with large payoffs or when the manager has access to all, or nearly all, fair hedges and gambles. In the latter case, the optimal contract may feature no observable performance-pay, yet elicit some effort. Empirical implications for the detection of earnings management, risk controls, robust contracting and observed compensation schemes across industries are also examined.

Keywords: Multi-Tasking, Agency, Manipulation, CEO Compensation, Option

---

*I thank Jon Glover and Rick Green for helpful comments and discussions. Many thanks to Vincent Glode, Burton Hollifield, Pierre Liang, Chris Sleet, Steve Spear, Sevin Yeltekin and participants to the Tepper FEAST Workshop. Several examples were borrowed from final projects in Aziz Lookman's Corporate Finance class (Tepper) and I would like to thank the students for their research which has indirectly benefited the paper.

†Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213. Corresponding author: Jeremy Bertomeu. E-mail address: bertomeu@cmu.edu. Current version: June 15th 2007.

Most companies and institutions engage in off-balance sheet risk management; yet, the amount of hedging practiced by firms, even in the same sector, varies considerably. For example, Chesapeake hedged 80 percent of its gas production. On the other hand, Exxon Mobil did not hedge: "Exxon Mobil (...) doesn't use financial hedges. Many of Chesapeake's peers take a similar position, reasoning that their skills are in finding oil and gas" (WSJ Nov. 6 2006). Similarly, Southwest Airlines hedged 85 percent of its fuel expenses in 2005 and 70 percent in 2006. During the same period, less financially solid airlines reversed their hedging policies. American Airlines terminated most of its contracts ending beyond March 2004 and hedged only 12 percent (resp. 4 percent) of oil risk for 2004 (resp. 2005). Continental Airlines had virtually no hedging in 2004 and 2005.

In the wrong hands, hedging instruments can also be abused and expose the firm to large potential losses.[1] Such problems are often linked to the lack of transparency when disclosing risk exposure and the profits from hedging instruments. The FASB and the IASB recently produced comprehensive statements with respect to hedging instruments (FAS 133, IAS 39); accounting regulations define conditions under which instruments can qualify for *hedge* accounting, broadly defined as accounting recognition of gains and losses only on the balance sheet (and not on the income statement). Yet, critiques observe that, since qualifying for hedge accounting is based on each asset separately (and not on aggregate net positions), managers often use risk management for reporting purposes and risk exposure is neither fully transparent nor well-controlled by shareholders.

Following several proposals from the SEC (regarding the application of the Sarbanes-Oxley Act), the FASB is designing a conceptual framework defining the role of risk management as well as several objectives for better regulating inappropriate hedging choices. Quite strikingly, the first of these objectives is to "Discourage transactions and transaction structures primarily motivated by accounting and reporting concerns rather than economics."[2] A fairly widespread

[1] For example, the notorious Jedi and Raptor special purpose vehicles created by Enron provided some insurance against a change in value of the company's assets but, since the insurance was guaranteed using Enron stock, only if the stock price remained high. In the Orange County scandal, fund manager Robert Citron used credit derivatives to create the illusion of above-average returns, eventually losing $2.2 billion out of $7.7 billion in 1994. Other more recent well-advertised cases include the hedge fund Amaranth and the scandals at Fannie Mae and Freddie Mac.

[2] See SEC report:"Report and Recommendations Pursuant to Section 401(c) of the Sarbanes-Oxley Act of 2002 on Arrangements with Off-Balance Sheet Implications, Special Purpose Entities, and Transparency of Filings by

idea is to formulate the basic trade-off in the following terms: hedging can improve economic risk-sharing between the firm and outside investors but can be distorted by managers for pure reporting reasons. Graduate textbooks generally mention the trade-off as one key to understanding both the need for hedging and the problems it may create.

"Why do firms use derivatives? The answer is that derivatives are tools for changing the firm's risk exposure. Someone once said that derivatives are to finance what scalpels are to surgery. By using derivatives, the firm can cut away unwanted portions of risk exposure and even transform the exposures into quite different forms. [...] Derivatives can also be used to merely change or increase the firm's risk exposure. [...] Most of the sad experiences with derivatives have occurred not from their use as instruments for hedging and offsetting risk, but, rather, from speculation."

from Corporate Finance - Ross, Westerfield and Jaffe (p.697)

Other references warn of the need to understand these instruments when monitoring the actions of managers:

"There are situations where off-balance sheet obligations make good economic sense. Unfortunately, those who have wanted to cover up their actions or who have not wanted to disclose the full amount and nature of their debt leverage have abused them. Often, the complexity of off-balance sheet vehicles makes it very difficult for an outsider to understand a company's true financial picture and sometimes for insiders as well, it appears. There should not be a blanket condemnation of the practice of off-balance-sheet financing, but directors need to insure the sound rationale of using such vehicles and that they are fully disclosed in company statements."

from Corporate Governance - McGraw-Hill Executive MBA Series (p.160)

One common denominator in current debates is that they are extremely loose with respect to the objective of risk management and do not explicitly say what is meant by "unwanted risk exposure," "good economic sense" or, as described by the FASB, "accounting and reporting

Issuers" (June 2005).

concerns." In fact, this question is also heavily debated in the corporate world as shown by the different views reported in the press. Several key questions seem to be of interest in order to better understand and regulate risk management.

I. In a model with perfect outside capital markets, what informational frictions (within the firm) can explain risk management by firms held by well-diversified investors? Then, under which conditions does risk management increase expected cash flows and/or is beneficial to agents in the economy?

II. What outcomes are "unwanted" risk exposure and will be hedged, and which outcomes will not be hedged? Should the firm always hedge against large losses or large gains?

III. How can one design executive compensation contracts that give incentives to hedge in the best interest of shareholders? How does risk management affect performance-pay coefficients? How does risk-taking by the agent respond to the concavity/convexity of the contract? What type of contracts will perform well in situations where risk management is important?

IV. What "anomalous" features of the cross-section and time series of earnings are consistent with risk management? Why do firms and regulators seem to tolerate earnings management to beat threshold or income smoothing?

Except for a few notable exceptions, most of the existing literature on risk management assumes that risk is managed directly by the owners of the firm (e.g., the shareholder, the board, etc.). Yet, current debates point out that, effectively, risk is under the control of the management, whose interest may not be fully aligned with those of the owners. This paper presents a framework in which the agent privately manages risk. The risk decision is imperfectly observed by the owners of the firm which creates an informational asymmetry between owners and managers. I discuss what problems may arise from this informational asymmetry and how incentives to manage risk can conflict with incentives to work hard (I.). Giving managers the freedom to hedge through financial derivatives can, potentially, either mitigate or exacerbate agency problems. On the one hand, hedging can reduce extraneous noise in the variables outsiders observe, and thereby allow them to more accurately evaluate the consequences of managerial actions and choices. This will reduce agency costs. On the other hand, hedging can, in itself, provide opportunities for self-dealing by managers at the expense of outsiders and of the total surplus created.

Resolving this trade-off, I investigate conditions under which a firm will hedge certain states of the world (II.). These conditions are derived from explicit informational frictions and not from exogenously specified capital market frictions. Then, I show how hedging can be elicited using well-designed managerial contracts (III.). These contracts can be interpreted to rationalize different executive contract shapes and hedging strategies chosen in different industries. I analyze how the agent will respond to the contract and recover features that are consistent with several anomalous properties of the cross-section and time series of corporate earnings. In complement to the empirical evidence, the analysis provides ways to rationalize several empirical tests of earnings management and link the cross-section of earnings to executive compensation (IV.).

In broad terms, the benchmark model is based on Holmström (1979) in which the actions of an agent are imperfectly monitored by the owners. In the standard model, the agent chooses effort and then the distribution of the output signal is given. A unidimensional costly effort choice is the only way that the agent can affect output. In my model, the agent can pick a distribution from a set of available hedges and gambles. The main theoretical contribution of the paper is to introduce a novel restriction on this set, called mean-distance-ordering (MDO), which intuitively frames the risk management problem. Under MDO, the cost of managing risk is (weakly) increasing in the distance between the original distribution and the hedged distribution. Technically, this assumption allows me to analyze a broad class of risk management actions from first-order conditions on the problem of the agent.

Initially, I take the contract as given and solve for the choices of the manager. Ross (2004) explains that offering more convex contracts may not make a manager more risk-seeking. His analysis suggests that much of the simple logic used to identify the causes of excess speculation is misplaced and one cannot always map contract shape to risk-taking. By focusing on risk-aversion, however, his approach does not make the agency problem explicit.[3] I show that, when the compensation is convex, the agent chooses a hedged distribution that has more probability mass in the tail of the distribution. In particular, a risk-averse manager paid with options will concentrate the mass of the distribution strictly above the strike price and simultaneously increase the probability of large losses. This property rationalizes a cross-section of earnings that exhibits bunching to beat target performance thresholds as well as a decrease in the probability of large gains, as documented by Degeorge, Patel and Zeckhauser (1999) (hereafter DPZ99). In this respect, the model provides a simple framework that theoretically validates several cross-sectional tests of earnings management.[4]

Next, I solve for the optimal contract when considering the two sources of moral hazard. Murphy (1999) (hereafter M99) documents that there is considerable heterogeneity among contracts offered to top executives and the forces causing this heterogeneity are not yet fully understood. He mentions one aspect of the problem, describing how contracts originally designed to provide the right incentives to increase shareholder value can also be prone to manipulation. The model accounts for both sides of the trade-off described by Murphy, since risk manipulation can lead to output signals that are more informative on the actions of the agent.[5] The analysis indicates, as a function of economic primitives, which states of the world one would expect to be hedged: I show that the compensation contract should induce the agent to reduce the probability of states in which the value of the firm is high or the output signal is relatively uninformative on the action of the agent. In the first case, offsetting incentives to hedge high payoff states would require a wage that makes the agent risk-neutral to output shocks, which is costly in terms of risk-sharing. In the second case, reducing the probability of a state increases its associated likelihood ratio (the standard measure of informativeness for a state), thereby

[3] He makes the following observation: "since compensation schedules arise as equilibria in agency models, it would be interesting to explore these results within such a setting. In particular, we should examine their implication when there is asymmetric information between agent and principal" (p. 224).

[4] In the existing literature, earnings management is often defined as misreporting current earnings to increase reported earnings in future periods. I follow here a broader definition of earnings management, as any action taken by the manager that may affect reported earnings. Further, in Section 5, I show that the model can be reinterpret as the agent manipulating a report made to the principal (but not the true economic cash flow). The framework predicts income smoothing effects that are consistent with the empirical evidence (because the agent may increase the probability of states with low payoffs in the current period and increase the probability of states with high payoffs in future periods).

[5] See Arya, Glover and Sunder (2003) for a survey of the existing literature showing when manipulation can be beneficial in solving other informational frictions.

providing a better solution to the incentive problem.

While the standard theory does not exclude non-linearities in the optimal contract, the empirical evidence suggests that most of the contracts that we observe are close to linear or, at least, are linear over certain output regions (e.g., equity compensation, piece-rate incentive schemes).[6] As explained by Holmström and Milgrom (1987), linearity can be desirable in situations where the agent can manipulate the contracting variables. I show that linearity is optimal when the control of the agent over risk is important.[7] When the agent has access to all fair gambles and hedges, a contract that prescribes a compensation that is linear in the managed signal is optimal. When risk management is costly, the optimal contract converges to a linear compensation schedule in states such that the value of the firm is large or when the cost of hedging becomes small.[8] In intuitive terms, a linear compensation performs well at preventing speculation over states that are relatively uninformative on the actions of the agent.

The theoretical literature incorporating risk management to agency theory is still relatively scarce. Two notable exceptions are Diamond (1998) and Palomino and Prat (2003) which present settings which are, while closely related in terms of the problem under investigation, quite different in terms of modeling assumptions and actual results. The first paper proposes an analysis of the optimal contract for a general risk management problem.[9] His main finding is that non-linearities in the compensation scheme distort project selection and thus, when the size of the cash flows becomes large relative to the cost of effort, a linear contract becomes optimal. In comparison to this paper, my assumptions on the problem are more demanding but they allow me to analyze the problem outside of this limiting case and obtain a richer set of predictions

[6] Holmstrom comments his results as follows: "In view of risk-sharing benefits, the convexity of the second-best solution may be surprising, but this is in no way exceptional (...). Examples for which sharing rules are concave or linear or even two-peaked can be easily generated as well" (Holmström (1979), p.80).

[7] It should be emphasized that the conceptual argument at play in my model is very different from Holmström and Milgrom (1987). In their model, the agent makes decisions over a sequence of periods but the principal only observes aggregate variables. Because the agent can adjust choices dynamically, any non-linearity in the compensation can be exploited. In my model, linearity holds in a static one-shot interaction and with only few restrictions on preferences or the production technology.

[8] The theoretical literature incorporating agency theory and risk management is still relatively scarce. An exception is Palomino and Prat (2003) who solve an agency model where the agent can chooses from a set of distributions verifying certain regularity conditions. See also Chang (1997), Biais and Casamatta (1999), Povel and Raith (2004) and Parlour, Purnanandam and Rajan (2006) for other models in which the choices of the agent may affect the risk of the output.

[9] Diamond's model is stated with risk-neutrality by the agent and the principal and three possible outputs; but this is done for expositional simplicity since most of the results extend to the general case considered here.

that match a large body of empirical evidence. While I also obtain linearity in limiting cases, linearity in my model is obtained by leaving the original moral hazard problem unchanged and making assumptions on the risk management problem.[10]. The second paper shares with my paper its focus on the optimal contract, even outside of limiting cases. The authors show that the optimal contract takes the form of a two-step bonus scheme under certain (single-crossing) restrictions on the set of hedges available to the agent. This is, to my knowledge, the first model to *derive* an interpretable optimal contract with risk management. The type of contract that they find, however, seems to be only rarely used in practice.[11]. My modeling framework accommodates risk-aversion and predicts a richer set of contracts. In my model, I can also characterize more precisely how the agent manages risk when given an optimal contract.[12]

1. The Model

I state the risk management problem for a firm, owned by a principal and operated for a single period by an agent (or manager). To keep the model simple, I assume that the firm is liquidated after this period ends and yields a net cash flow $y \in X$. The infimum of X is denoted $\theta \in [-\infty, +\infty)$ (it is the maximum loss that the manager may cause). The manager privately chooses an action $a \in [\underline{a}, \overline{a}]$ and then can manage risk by selecting a distribution $\hat{F}(.)$ from a non-empty set $\Gamma(a)$. The set $\Gamma(a)$ is defined as the choice set of the agent for a given effort and corresponds to the set of all hedges and gambles that are available to the agent.

[10]For example, a natural concern is that the risk management problem may change as cash flows become large (i.e., it may be harder to hedge for a very large firm) which could work against Diamond's limiting argument.

[11]It should be emphasized that they present their findings in the context of money managers (for which tournament-like payments have similarities with a two-step bonus scheme), yet their modeling assumption are based on a generic agency model and thus may be valid for CEOs (or many other occupations); even the bonus scheme component of executives is not a *two-step* bonus scheme. Murphy (1999) documents that about 70% of his sample of CEOs use $80/120$ or Modified Sum-of-Targets as their bonus scheme. In a $80/120$, the compensation is capped for performance equal to 80% or less of the target and 120% or more of the target, but it is linear and increasing between 80% and 120%. Qualitatively, the Modified Sum-of-Targets is similar. The remaining firms in their sample used more complex formulae with a non-trivial dependence of payments on performance targets.

[12]See also Chang (1997), Biais and Casamatta (1999), Povel and Raith (2004) and Parlour et al. (2006) for other models in which the choices of the agent may affect the risk of the output.

Risk Management I define $F(.|a) \in \Gamma(a)$ as the distribution of y when the agent does not manage risk and assume that it has mean a.[13] In order to reflect the idea that effort increases the value of the firm, I assume that $F(.|a)$ first-order stochastically dominates $F(.|a')$ if $a \geq a'$. I restrict the attention to cases in which managing risk cannot directly increase the value of the firm and thus, $\Gamma(a)$ must include distributions that have mean (weakly) below a. As a result, conditional on a, $F(.|a)$ maximizes the expected value of the firm.

More explicit restrictions on the sets $\Gamma(a)$ are delayed until Sections 2 and 3; yet, it is helpful at this point to think about a as value-increasing effort, $F(.|a)$ as the distribution of output without risk management and then $\Gamma(a)$ as a set of distributions that can change the risk of y but may not increase value.

Preferences For now, I assume that the principal is risk-neutral. This assumption is made to focus the attention on the most distinctive aspects of the framework since the role of risk-management in the presence of exogenous market frictions is already well-understood (Froot, Scharfstein and Stein 1993). In contrast, if capital markets are perfect, risk management should be *a-priori* irrelevant to the value of a firm owned by well-diversified investors; in this setting, hedging may be required only as a result of an incentive problem. In Section 5, I extend the framework to risk-aversion by the principal and show that most of the results are robust to other (exogenous) capital market frictions.

The principal can provide incentives to take a desired action by offering a compensation contract $w(y)$ (defined over $\mathbb{R}$). For outcome y and action a, the manager achieves a utility $u(w(y)) - \psi(a)$ satisfying standard regularity conditions ($u' > 0$, $u'' < 0$, $\psi(\underline{a}) = \psi'(\underline{a}) = 0$, $\lim_{a \to \overline{a}} \psi(a) = +\infty$, $\psi'', \psi''' > 0$ except possibly at $a = \underline{a}$). Unless explicitly stated otherwise, I assume that the agent has limited liability and must be paid $w(y) \geq \underline{w}$ for all y. The model incorporates two important aspects of an agency situation: (i) increasing the value of the firm is privately costly, (ii) the agent may be risk-averse and dislike volatile compensation. Finally, the contract must prescribe a minimum reserve utility equal to b which corresponds to the outside option of the agent.

[13] The statement is without loss of generality when the mean of $F(.|a)$ is continuous in a. For example, even if the mean of $F(.|a)$ is not a, one may always relabel effort $A = \int y dF(y|a)$.

Contracting Problem A contract $(w(.), \hat{F}(.), a)$ is incentive-compatible if for a given $w(.)$ the agent chooses $(\hat{F}, a)$. Taking into consideration the actions of the manager, the principal will choose an optimal contract which solves the following problem.

$$(P) \qquad \max_{a, w(.) \geq \underline{w}, \hat{F}(.)} \int (y - w(y)) d\hat{F}(y)$$

$$s.t. \qquad \int u(w(y)) d\hat{F}(y) - \psi(a) \qquad \geq b \tag{1}$$

$$(a, \hat{F}(.)) \in \arg\max_{\tilde{a}, F(.) \in \Gamma(\tilde{a})} \int u(w(y)) dF(y) - \psi(\tilde{a}) \tag{2}$$

While I focus in most of the analysis on incentive-compatibility as stated in Equation (2), it will also be important to interpret certain situations in which a solution to the problem of the agent does not exist but the contract is "unreasonable" and should not be offered. First, I say that a contract is incentive-free if any sequence of policies converging to the supremum in the problem of the agent prescribes an action converging to $\underline{a}$. From the perspective of the principal, an incentive-free contract does not provide any incentives to work hard and thus can be replaced by a constant wage. Second, I say that a contract is agent-unbounded if the supremum of the problem of the agent is infinite. Because the value of the firm is bounded, an agent-unbounded contract cannot be optimal.

Definitions The *first-best* is an optimum to this Problem when the incentive-compatibility condition is omitted. I introduce some additional terminology to simplify the exposition. I say that a compensation scheme is linear when $w(y) = h_0 + h_1 y$. For a given a, I say that the agent fully hedges when y is deterministic (i.e., $\hat{F}(.)$ is degenerate) and the agent hedges (against) an outcome y when its probability according to $\hat{F}(.)$ is lower than under $F(.|a)$. In contrast, an agent speculates when there is in $\Gamma(a)$ a distribution that (strictly) second-order monotonically stochastically dominates $\hat{F}(.)$.[14]

[14]See Huang and Litzenberger (1988), p.49-50 for a discussion of second-order monotonic stochastic dominance. Formally, a random variable A (with distribution F_A) second-order monotonically stochastically dominates a random variable B (with distribution F_B) if and only if $\int_{-\infty}^{y} (F_A(z) - F_B(z)) dz \leq 0$ for all y and

2. Perfect Risk Management

Before moving to the general problem, it is useful to gain some preliminary intuition in a simplified setting. I assume here that the manager has access to all fair gambles and hedges and thus $\Gamma(a) = \{\hat{F}(.)/ \int y d\hat{F}(y) \leq a\}$. Under perfect risk management, the choice of the manager over gambles or hedges is unrestricted provided it does not increase firm value. To avoid uninteresting issues caused by a discrete support, I assume here that X is an interval including $[\underline{a}, \overline{a}]$.[15] I present next two simple examples showing which situations perfect hedging can capture.

Example 1: A fund manager has access to a risk-free trading strategy (or perfect arbitrage) that, for each dollar invested, yields a return 10%. The manager can invest $\$10a$ but then incurs management fees equal to $\psi(a)$ (see Berk and Green (2004) for an interpretation of these management fees). In addition, the manager can speculate and purchase zero net-present-value gambles. The principal cannot monitor these gambles because the manager may always claim that these gambles are part of the trading strategy. In this first example, the output conditional on a is risk-free (i.e. $F(y|a)$ prescribes $y = a$ with probability one) but the principal cannot prevent the agent from taking gambles.

Example 2: Suppose a manager chooses effort a and conditional on a, the output is distributed according to a distribution $F(.|a)$ with full support on X. Then, assume that the manager may provide verifiable information on a to a financial intermediary (e.g., an expert on the industry). If the intermediary is competitive, it will offer any gamble or hedge substituting the original output y drawn from $F(y|a)$ with an output drawn from a distribution $\hat{F}(.) \in \Gamma(a)$. In this second example, the effort choice and the risk management decisions are observable but not contractible (see **?**) for a discussion of variables that are observable but not contractible).

$\int_{-\infty}^{+\infty} z dF_A(z) \geq \int_{-\infty}^{+\infty} z dF_B(z)$. Monotonic second-order dominance extends second-order dominance to cases in which the random variables do not have the same mean. Any risk-averse and non-satiated individual would prefer A to B.

[15] In other terms, $X = (-\infty, \theta'']$, $X = [\theta', +\infty)$ or $X = (\theta', \theta'')$ where θ' and θ'' are two real numbers. All results carry over provided the elicited effort is in X.

Even in settings where perfect risk management may seem too extreme (e.g., the information of financial intermediaries on a is imperfect), the assumption can help better understand how the risk management problem is affecting the contractual relationship. It should be emphasized that I take perfect risk management as a thought experiment in a limiting case, not as a model to be matched to the empirical evidence; to this effect, frictions to perfect risk management are developed in the Sections 3 and 4. To ease the exposition, I assume for now that $b \geq u(\underline{w})$ (the other case $b < u(\underline{w})$ will be studied separately since, then, the limited liability replaces the participation of the agent).

2.1. First-Best

For reference, I state the first-best solution to the problem. Since the principal is risk-neutral but the agent may be risk-averse, there is an optimal policy such that the agent is given a constant wage W. In first-best, this wage binds the participation of the agent, i.e. $u(W^*) = b + \psi(a^*)$. The principal solves the following Problem:

$$(P_{fb}) \qquad \max_a a - u^{-1}(b + \psi(a))$$

The first-best effort is given by the following first-order condition.

$$\psi'(a^*) = u'(u^{-1}(\psi(a^*) + b)) \tag{3}$$

Let (a^*, W^*) be the first-best outcome; it equates the marginal disutility of effort with the marginal cost of compensating the agent to work more.[16] Note that, in first-best, the irrelevance of the capital structure holds: risk management does not increase the utility achieved by the principal or the agent.

In first-best, the principal controls both effort and risk management. I claim that the first-best surplus can also be attained provided the principal controls only risk management. Suppose for example the principal chooses risk management but not effort: ex-ante, the principal chooses

[16] First-best effort would still be given by Equation (3) even if the principal were risk-averse since the principal would then fully hedge and set $y = a^*$ with probability one. More generally, it can be verified that risk-aversion by the principal would not affect any of the results presented in this Section.

a hedged distribution $\hat{F}^a(.) \in \Gamma(a)$ for all possible a (although the effort a remains chosen by the agent). Then, the principal could set $y = a$ with probability one for all a and offer W^* only if $y = a^*$. The agent will respond to the contract by choosing $a = a^*$, leading to first-best payoffs.[17] In contrast, I will show next that first-best can fail when the agent controls both effort and risk management.

2.2. Optimal Risk-Sharing

I develop next the optimal risk management strategy in second-best, when the decision to provide effort or hedge is under the control of the agent. With perfect risk management, any manager who does not fully hedge must be speculating (as defined earlier in Equation (3)). This is because fully hedging, i.e. setting $y = a$ with probability one, second-order monotonically stochastically dominates any other distribution with mean less than a. Taking for now the action as given, the next lemma describes what risk management strategies should be elicited by the principal.

Lemma 2.1. *An optimal contract must prescribe fully hedging.*

The first result is intuitive: it states that the agent should hedge away all the risk because he/she can do so; yet, in the context of a moral hazard problem, it may seem surprising. In the standard model, a contract should impose (unwanted) risk on the agent to preserve incentives. With perfect risk management, however, it is optimal to set a compensation that is deterministic.

At first sight, a deterministic compensation may seem incompatible with incentives to work hard. To see why a deterministic compensation provides incentives here, note that while the compensation is constant ex-post, it is not necessarily constant in the ex-ante problem if the agent deviates to work less or speculate. For example, let W and $\overline{y}$ be two scalars and suppose that the principal offers a compensation $w(y) = W$ if $y \geq \overline{y}$ and $w(y) = \underline{w} < W$ if $y < \overline{y}$. By fully hedging and choosing effort $a = \overline{y}$, the agent will receive a constant compensation W. By deviating to less effort, the agent would have to set positive weight on outcomes $y < \overline{y}$

[17] For obvious reasons, first-best can also be achieved if the principal chooses effort but the agent controls risk management (i.e., the principal will then pay a flat compensation).

and thus would receive an expected compensation strictly below W. To summarize, perfect risk management implies a constant wage on-equilibrium but not necessarily off-equilibrium.

2.3. Incentive-Compatibility

Since the agent fully hedges (by Lemma 2.1), the distribution $\hat{F}(.)$ will be degenerate and thus the output signal will be *in equilibrium* perfectly informative and equal to the chosen effort $y = a$. The presence of a perfectly informative output signal $y = a$ may suggest that the principal should be able to reach first-best. This is not (always) the case here. In Figure 1, I show by way of an example why first-best may fail. Assume that the principal offers a wage $W^* = b + \psi(a^*)$ conditional on $y = a^*$ and $\underline{w}$ for any other realization of y. This compensation scheme binds the participation of the agent and seems to be geared to elicit a^* by minimizing payoffs over states that should not occur when the agent fully hedges.

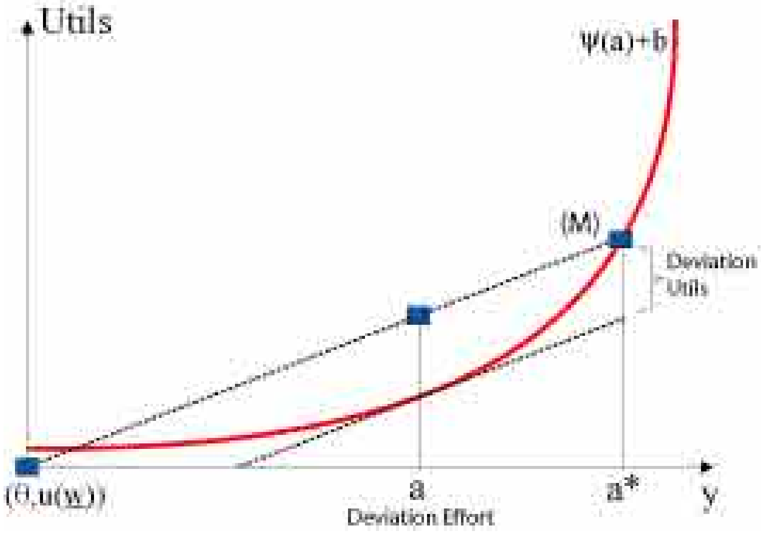

Figure 1. Optimal Effort Choices

It turns out that that the agent does not choose first-best in this example. In Figure 1, the rightmost square corresponds to the first-best choice; it implies no rents for the agent. By managing risk, the agent may also speculate and choose any distribution over $y = \theta$ (the worst-case outcome) and $y = a^*$, which leads to a set of feasible utilities (gross of effort cost) delineated by the straight line (M). Note first that the agent would not speculate while still choosing a^* since this would lead to an expected utility below b. On the other hand, the agent may also deviate in effort and choose $a < a^*$. In the plot, I find the effort a that maximizes the utility of the agent (i.e. the distance between (M) and $\psi(a) + b$). The analysis shows how in response to the contract, the agent deviates jointly in effort and risk management and will speculate over the worst possible state of the world.

Proposition 2.1. *If the maximum downside is unbounded ($\theta = -\infty$), a contract that is not agent-unbounded must be incentive-free. That is, the principal cannot elicit any effort above $\underline{a}$ for a finite expected wage.*

To begin with, I discuss the incentive problem when the worst state of the world generates an unbounded loss. In this case, I show that the principal can only elicit the minimum possible effort. This is because the agent can design a risk management strategy that exposes the firm to very large losses with a small probability and use such gambles to almost always achieve the maximum possible compensation. Here, these gambles act as substitutes for effort and make the managed signal useless for contracting purposes. The model provides a simple condition under which an indicator of performance is not robust to manipulation by the manager (and thus is not usable in the contract).[18] Note that while the agent will choose $\underline{a}$, this may be hard to detect because $y > \theta$ will occur with a probability that is almost one. To give content to the contract design problem, therefore, I assume in the rest of the paper that θ is finite and meets the condition in Lemma 2.2.

Lemma 2.2. *There exists a contract eliciting effort $a > \underline{a}$ for an expected wage below w if and only if:*

$$\frac{u(w) - u(\underline{w})}{a - \theta} \geq \psi'(a) \tag{4}$$

Even when the maximum loss is bounded, the principal must be mindful of a joint deviation in effort and risk management and, to deter unwanted risks, must offer a compensation that

[18] See also Liang and Nan (2006) for a LEN model in which an informative signal is not used in the presence of risk management. A related idea of robustness of signals is developed in Glover, Ijiri, Levine and Liang (2007) in which the agent can alter the output signal and may have private information about his ability to do so. They show that in certain cases it is optimal to ignore the measure that can be manipulated.

mitigates incentives to speculate. Equation (4) shows that the relevant statistic in the contract is the utility of the agent received per unit of output. If the agent is risk-neutral, this ratio captures the performance-pay coefficient, i.e. the share of the output that should be given to the manager. An incentive-compatible contract must ensure that the agent receives a sufficient proportion of the total output produced.

It follows that the principal has two ways to make a contract incentive-compatible. One, the principal can elicit less effort, reducing $\psi'(a)$. This will reduce the attractiveness of working less but will also reduce the firm's value. Two, the principal may pay above the reserve, increasing $(u(w) - u(\underline{w}))/(a - \theta)$. This will make speculation more costly to the agent but will also increase total compensation cost.

2.4. Second-Best Effort and Rents

Under perfect risk management, it is (weakly) optimal to maximize punishments over outcomes that should not be observed in equilibrium. Therefore, one may restrict the attention without loss of optimality to contracts (W, a) such that the principal recommends effort a, pays W when $y = a$ and $\underline{w}$ for any other realization of y. This contract clearly minimizes incentives to speculate (uniqueness is delayed until Section (2.6)).[19]

I define two other contracts which may solve the moral-hazard problem when first-best cannot be attained. First, let (W^{**}, a^{**}), as defined in the previous paragraph, correspond to the contract such that the participation of the agent binds, i.e. $W^{**} = b + \psi(a^{**})$, and the incentive-compatibility binds:

$$\psi'(a^{**}) = \frac{b + \psi(a^{**}) - u(\underline{w})}{a^{**} - \theta} \tag{5}$$

To guarantee a unique solution to Equation (5), I assume that, for all a, $\psi'(a)/\psi(a) \leq 1/(a - \theta) + (b - u(\underline{w}))/\psi(a)$. This condition implies that the elasticity of the cost of effort is not too large. In the model, it guarantees that the performance coefficient obtained in Equation

[19] This type of contract is similar to the contract proposed by Palomino and Prat (2003) for similar conceptual reasons. The authors show that for a given set $\Gamma(a)$ (different from the set studied here), risk-neutrality by the agent and a monotonic compensation, a two-step bonus scheme is optimal. In their model, the bonus contract maximizes punishments for outcomes that indicate low actions.

(4) is non-increasing when more effort is elicited.

Second, let (W^{***}, a^{***}) be the optimal contract such that only the incentive-compatibility binds. This contract will always dominate (W^{**}, a^{**}) provided it makes the agent participate. By incentive-compatibility, the wage can be written $W^{***} = u^{-1}(\psi'(a^{***})(a^{***} - \theta) + u(\underline{w}))$. Substituting in the problem of the principal, a^{***} must maximize $a - u^{-1}(\psi'(a)(a - \theta) + u(\underline{w}))$.[20] The (unique) solution to the first-order condition in this problem characterizes a^{***}:

$$\psi'(a^{***}) = u'(W^{***}) - \psi''(a^{***})(a^{***} - \theta) \tag{6}$$

In the next Proposition, I collect these three contracts and solve for the optimal contract.

Proposition 2.2. *Suppose $b \geq u(\underline{w})$. The optimal contract is given as follows:*

(i) If $a^* \leq a^{**}$, the first-best contract (W^*, a^*) is optimal.

(ii) If $a^* > a^{**}$ and $u(W^{***}) < b + \psi(a^{***})$, the contract (W^{**}, a^{**}) is optimal.

(iii) In all other cases, the contract (W^{***}, a^{***}) is optimal.

The elicited effort is always smaller than a^* (strictly in (ii)-(iii)) and greater than $\underline{a}$ (strictly when $b > u(\underline{w})$).

In contrast to unbounded downside risk, the principal always elicits an effort strictly greater than $\underline{a}$ (using a contract that is not flat). I also show that in second-best the expected value of the firm is strictly lower than in first-best. The result shows that hedging by the agent will reduce expected cash flows, except in cases where first-best is attained. Intuitively, incentives to speculate are greater when the agent contributes more to the firm. Internalizing the cost of giving incentives, the principal responds to informational frictions by eliciting less effort. When first-best cannot be attained, I find two possible second-best contracts.

The contract (W^{**}, a^{**}) selects the effort level such that an agent paid his/her outside option is indifferent between speculating and hedging. As is intuitive, more downside risk (θ) or a greater limited liability ($\underline{w}$) leads to less effort in this contract. On the other hand, a greater

[20] Uniqueness is obtained because ψ''' is positive and thus this program is concave.

outside option (b) for the agent works to increase how much the agent may lose if he/she does not select the appropriate action and thus increases the elicited effort. An interesting feature of this situation is that the resulting allocation is Pareto-dominated by first-best, i.e. the agent receives no rent and the principal is worse-off. This aspect provides some evidence that hedging can be undesirable to *all* contracting parties.

The contract (W^{***}, a^{***}), on the other hand, leads to a utility for the agent that is strictly greater than in first-best. It should be noted that agent rents occur even though the limited liability never binds *ex-post*, i.e. the agent is paid $\underline{w}$ with probability zero. This feature of the model contrasts with standard models with a limited liability (such as Innes (1990)). Further, in practice, one rarely managers being driven to levels of poverty (e.g., personal bankruptcy) that seem fully consistent with a binding ex-post limited liability, even after very low performance.

Equation (6) can be compared to first-best. The optimal effort equates the marginal cost of effort to the marginal utility from consumption (as in first-best) and, in addition, incorporates the required compensation to avoid speculation, which increases the marginal cost of increasing effort by $\psi''(a^{***})(a^{***} - \theta) > 0$. The comparative statics with respect to θ and $\underline{w}$ are similar to those with contract (W^{**}, a^{**}). Note that in this framework, these comparative statics are simple and intuitive while they would be ambiguous in a model without risk management.

I extend next the analysis to the case in which the limited liability is *greater* than the outside option of the agent, i.e. $u(\underline{w}) > b$ (versus $u(\underline{w}) \leq b$ previously). In this case, the participation of the agent cannot bind and thus the contracts (W^*, a^*) and (W^{**}, a^{**}) are no longer feasible. On the other hand, the principal may now choose to bind the limited liability everywhere, which leads to a contract $(\underline{w}, \underline{a})$. In the next Proposition, I compare this contract to (W^{***}, a^{***}).

Proposition 2.3. *Suppose $b < u(\underline{w})$. The optimal contract is given as follows:*

(i) If $u'(\underline{w}) > \psi''(\underline{a})(\underline{a} - \theta)$, (W^{***}, a^{***}) is optimal.

(ii) Else, the principal chooses to pay always $\underline{w}$ and elicits $\underline{a}$.

The elicited effort is always strictly smaller than a^*.

Corollary 2.1. *The elicited effort is strictly greater than $\underline{a}$ (and the contract is not flat) if and* *only if either* $\psi''(\underline{a}) = 0$ *or:*

$$\theta > \underline{a} - \frac{u'(\underline{w})}{\psi''(\underline{a})} \quad (7)$$

In contrast to Proposition 2.2, it may be optimal to elicit the minimum effort when the maximum downside risk is sufficiently important. The statement shows that the conclusions of Proposition 2.1 may be robust to a bounded downside risk, but only when the limited liability lies above the outside option of the agent. Intuitively, when the limited liability becomes larger than the outside option of the agent, the principal may no longer use the threat of paying $\underline{w}$ an agent contributing a very low effort.

2.5. Discussion

The analysis at this stage yields several practical implications for the analysis of managerial contracts. I interpret here several comparative statics in the context of existing debates on risk management.

Remark 1: Imperfectly-monitored Risk Management can cause Low Output

In the model, first-best could be attained if the risk management process was monitored by the principal, even if effort were still chosen by the agent (as argued in Section 2.1). However, when the agent controls the risk management process, the principal may choose a lower effort than first-best. An implication of the framework is that the lack of transparency on risk choices can, jointly with a standard moral-hazard problem, cause low output and large absolute pay. This concern alone has attracted considerable attention from regulators. While previously risk management was viewed as being the sole responsibility of management, the Sarbanes-Oxley Act and various implementation notes from the FASB and SEC require better disclosure on the risk taken by managers, in particular regarding hedging instruments. In line with these recommendations, the Committee of Sponsoring Organizations of the Treadway Commission (COSO) requires internal auditors to identify and manage all risks faced by the organization.[21]

[21] See 2004 COSO published Enterprise Risk Management - Integrated Framework. Risk Assessment is one of the eight components of the COSO framework. COSO is a US private-sector initiative created in 1985 and sponsored by the main accounting institutes.

In the financial industry, the Third Pillar of the Basel II accord increases the disclosure of financial institutions over their portfolio of investments. These regulatory changes suggest that, unlike other operational decisions, risk management should be under the direct supervision of boards and voting shareholders, and not only decided by managers.

Remark 2: More liability by managers increases efficiency

Increasing the liability of managers can be helpful to better solve the risk management problem. For example, many professions in which risk takes an important role are structured in partnerships with, among other things, varying degrees of personal liability. The choice of the right liability level may account for this organization form for managerial occupations in the law, accounting, architecture industry and explain the dependence of insurance premia on performance in medical professions. A distinctive feature of these industries is a certain control over risk: certain pleas may have greater chances of success but higher losses if they fail, audit reports may be qualified or unqualified, architectural designs must balance creativity and cost uncertainty, and an experimental medical treatment may worsen a condition.[22] As evidence that the limited liability may be complementary to better disclosure, the Sarbanes-Oxley Act increased legal liability for executives taking risks that are not well monitored; this should increase productive efficiency toward first-best, but may also reduce the rents achieved by current insiders.[23]

Remark 3: More downside risk aggravates agency cost

The resolution of the moral hazard problem presupposes some control over maximum downside risk. In the fund management and trading professions, firms generally restrict the asset classes taken by managers and bound maximum losses. Such controls may, of course, be more difficult for large financial institutions. Yet, at least in principle, the First Pillar of the Basel

[22] Since 1996 (Uniform Partnership Act), limited liability partnerships or LLPs are legal in many US states. This organization form is fairly new and still rare in non-US countries. In addition, many US states restrict the domain of application of the limited liability clause to negligence claims.

[23] I interpret here an increase in liability as $\underline{w}$, not as a wasteful punishment. The latter, however, produces the same comparative statics and is omitted to save space and notations. Further, the exact ranking of a^{**} and a^{***} and their response to a change in θ or $\underline{w}$ is ambiguous as well as the effect of $\underline{w}$ on the rent of the agent conditional on a^{***}; yet, the statement is justified if one compares a^{***} (with rents) to a^{*} (no rents).

II accord explicitly requires financial institutions to maintain a regulatory capital in proportion to the risk of each investment. Many financial failures, such as the Barings bank, the Orange County pension fund or, more recently, the Amaranth hedge fund are stereotype examples of a large imperfectly-monitored downside risk. A less clear-cut but interesting case is the hedge fund Long-Term Capital Management whose positions were often described as nearly-perfect arbitrage. Over its existence the fund was given increasingly large leverage ratios and, in 1997, returned nearly half of its capital to shareholders. The conjunction of these two effects increased maximum downside risk without increasing the maximum liability. In addition, as arbitrage opportunities decreased, the compensation of the management team decreased in levels. Thus, if management contracts were not adjusted to reflect such changes to the environment, the fund may have fallen into the region where it was ex-ante desirable to move away from pure arbitrage positions. In line with this idea, Lowenstein (2002) describes changes to LTCM's portfolio over the latter years of its existence (such as equity arbitrage) with different correlation structures and greater levels of risk.

2.6. Robustness of the Linear Contract

It should be noted that the model with perfect risk management predicts the equilibrium compensation and effort, but the off-equilibrium payments (if the agent deviates) are not uniquely predicted. For example, if the principal is willing to elicit a, offering a targeted bonus that pays above $\underline{w}$ only if y is equal to the recommended action is always optimal but many other contracts such that pay is sufficiently low for $y \neq a$ are also optimal. Since analyzing the optimal contract is the focus of this paper, refining the multiplicity of optimal contracts is important. I give next a simple argument showing why a simple linear contract is the unique contract that is robust to other frictions that one may reasonably expect.

Assume that the cost of effort is decomposed as: $\psi(a) = c\tilde{\psi}(a)$, where $c \in \mathbb{R}\backslash\{0\}$. I will suppose that $\tilde{\psi}(a)$ is common-knowledge and the parameter c is known to the agent but not the principal. It is drawn ex-ante from a distribution $H(.)$ which has full support on $\mathbb{R}^+$ (and no mass point at $c = 0$). Uncertainty about the cost of the effort reflects the idea that some aspects of the agency problem may be better known by the agent. To avoid self-selection, I assume here that $u(\underline{w}) > b$ and therefore agents always participate. In addition, I assume that $\psi''(\underline{a}) = 0$ so that, if c was common knowledge, an effort strictly above $\underline{a}$ would be elicited.[24]

The contracting environment is similar to the previous problem:

1. The principal proposes a contract $w(.)$.

2. The agent observes c. Then, he/she decides on an effort $a(c)$ and a distribution $\hat{F}_c(.)$, which depends on c, with the constraint that the expected output under $\hat{F}_c(.)$ is less than $a(c)$.

3. The output is drawn from $\hat{F}_c(.)$. The principal receives $y - w(y)$ and the agent achieves $u(w(y)) - c\psi(a(c))$.

Let $u^j(.)$ $(j > 1)$ be a strictly concave utility function. I state next the extended contract design problem:

$$(P') \qquad \max_{w(.)\geq \underline{w}, \hat{F}_c(.), a(c)} \int \int (y - w(y)) d\hat{F}_c(y) dH(c) \tag{8}$$

s.t., for all c,

$$(a(c), \hat{F}_c(.)) \in \arg\max_{\tilde{a}, F(.)\in\Gamma(\tilde{a})} \int u^j(w(y)) dF(y) - c\tilde{\psi}(\tilde{a}) \tag{9}$$

The model is equivalent to the previous problem when the distribution $H(.)$ has mass on only one value of c. A small amount of risk-aversion is important in order to remove solutions optimal only when agents are perfectly risk-neutral; to remove risk-aversion in the limit, I assume that the sequence of functions $u^j(.)$ converges to a linear function when j increases. This technique is used to obtain a notion of near risk-neutrality and I use it to select a particular optimal contract.[25] In the next Proposition, I assume that $1/(a - \theta) + \psi''(a)/\psi'(a)$ is concave. This restriction implies that the maximum downside risk is sufficiently important and the elasticity of the cost of effort does not grow too fast. For reference, I say (P') is solved in partial-information if the principal observed c (and $w(.)$ may depend on c). The partial-information problem was solved in Section (2.4).

[24]These assumptions ensure that the principal does not shut down any cost type; if this were the case, the contract would not be unique or linear over regions that are never attained. This is because certain realizations of y would never be observed on the equilibrium path.

[25]The same argument would have been true if the principal had a small amount of risk-aversion but the agents are risk-neutral. If both parties are perfectly risk-neutral, however, speculation becomes costless which no longer makes convex contracts costly.

Proposition 2.4. *As the utility of the agents becomes linear (i.e., j increases), any solution to (P') converges to a linear contract.*

Corollary 2.2. *When j becomes large, there exists a threshold $\overline{c}$ such that if $c < \overline{c}$ (resp. $c > \overline{c}$) a manager with cost of effort $\psi(a) = c\tilde{\psi}(a)$ is strictly better-off (resp. worse-off) when c is unobservable by the principal than in the partial-information problem.*

Corollary 2.3. *If, in addition, $H(.)$ converges to a mass point at c_0 (i.e. $c = c_0$ deterministic), any solution to (P') must converge to a linear contract: $w(y) = \underline{w} + \psi'(a^{***})(a^{***} - \theta)$, where a^{***} is the elicited solution in Proposition 2.2 with a cost $\psi(a) = c_0\tilde{\psi}(a)$.*

A linear contract is obtained as the unique solution in (P) with a small uncertainty about the cost of effort and almost risk-neutral agents.[26] When c is unobservable, the high-cost types (i.e., the "bad" types) are better-off than if cost was observable whether those with low cost of effort are better-off. This is because the low-cost types must be compensated above their value to avoid speculation; some of this additional compensation cost is transferred to the low-cost types.[27]

I shall try to explain why linearity is optimal here. Uncertainty about the cost of effort implies that the output (y) will no longer be deterministic. Agents with higher cost will produce less firm value while those with lower cost will produce more firm value. Faced with a compensation such that $u(w(y))$ is not concave, a positive mass of agents will always take gambles to reach up to the concavification of $u(w(y))$. As argued in Lemma 2.1, this will cause greater compensation cost than if the principal offers a compensation that does not elicit speculation. This argument should convince the reader that, given risk management and uncertainty about cost, the principal is constrained to a compensation scheme that is weakly concave over the whole domain X.

Concavity works against what a principal would have chosen if the cost of effort was observable. Because the cost of effort is convex, the wage should have been convex in the elicited

[26]Note that by a minor notational change in the proof of Proposition 2.4, this contract yields the same surplus for the principal as a revelation mechanism in which agents declare their type and then are given a surplus.

[27]Note that this property of the model implies that, if managers could truthfully reveal their cost types ex-ante, all low-cost types would reveal their types. This would trigger, in equilibrium, all types truthfully revealing their cost types.

effort if c was publicly known. The principal chooses the compensation that is closest to this scheme but constrained to a weakly concave scheme. Solving for the best concave compensation scheme that comes close to this convex schedule yields an optimal compensation that is linear. In intuitive terms, a convex compensation scheme is required to compensate the high-skilled workers while a concave compensation is required to avoid speculation by the low-skilled workers.

3. Imperfect Risk Management

3.1. Mean-Distance-Ordering

I develop now the general framework by modeling more explicitly economic frictions that may prevent the agent from fully hedging. Let $X = \{x_k\}_{k=1}^N$ be an increasing set of possible outcomes where $N \geq 3$ is possibly infinite and which includes a^* as defined earlier.[28] I move away from $\Gamma(a)$ being the set of all distributions with mean below a and assume now that any distribution $\hat{F}(.) \in \Gamma(a)$ different from $F(.|a)$ must have mean strictly less than a. I define next a simple separability restriction on the cost required to hedge.

Definition 3.1. $(\Gamma(a))_{a\in[\underline{a},\overline{a}]}$ *is mean-distance-ordered (MDO) if for all* a*, there exists a distribution* $(p_k(a))_{k=1}^N$ *with* $p_k(a) > 0$ *such that:* $\hat{F}(.)$ *given by* $(\hat{p}_k)_{k=1}^N$ *is an element of* $\Gamma(a)$ *if and only if* $\sum_{k=1}^N \hat{p}_k x_k \leq \sum_{k=1}^N p_k(a) x_k - \sum_{k=1}^N C(a, x_k, \hat{p}_k - p_k(a))$ *where* C *is a positive twice continuously-differentiable function verifying for any* a, x_k, δ,

(i) $C(a, x_k, \delta)$ *is twice differentiable in all its arguments, and strictly convex in* δ *with* $C(a, x_k, 0) = C_\delta(a, x_k, 0) = 0$, $C_\delta(a, x_k, \delta) < 0$ *(resp.* $C_\delta(a, x_k, \delta) > 0$*) for* $\delta < 0$ *(resp.* $\delta > 0$*),*

(ii) $\lim_{\delta\to p_k(a)} C(a, x_k, -\delta) \geq \lim x_n - x_1$ *(possibly* $+\infty$*).*

Under MDO, distributions are ordered in terms of their distance from a reference $(p_k(a))_{k=1}^N$ with distribution denoted $F(.|a)$. Here, I interpret hedging as moving the distribution from $F(.|a)$ to $\hat{F}(.)$. As the agent hedges more, in that the distribution becomes more distinct from

[28] Discreteness in the outcome space is useful to avoid measure-theoretic considerations, but otherwise not required for most of the results.

the original distribution, the cost of hedging increases.[29] In the model that follows, outsiders will observe an output generated from $\hat{F}(.)$. The specification implies that there is a deadweight cost associated to hedging. Intuitively, this cost can be interpreted as the search cost of finding hedging partners or the managerial attention required to change the risk of the cash flow. Assumptions (i)-(ii) are imposed to avoid boundary solutions but are otherwise not essential. Note that when C is (nearly) zero, MDO will correspond to the case of perfect risk management considered earlier.

An important aspect of the definition should be emphasized. The concept of MDO does not impose that mean-preserving spreads should be cheaper to induce than distributions with greater precision. I do not attempt to model here how the agent may increase the risk of the project by investing in risky publicly-traded securities at very little cost. Such cases would likely be observable and controllable by the principal and thus do not fit in a theory in which the decision to hedge is private. In my setting, engaging in mean-preserving spreads would require the agent to use derivative arrangements that are not easily observable and could involve substantial cost.[30]

The concept of MDO is consistent with the idea of sophisticated financial markets with a large set of available hedges priced by intermediaries. To illustrate this point, let y' denote the original unhedged signal (drawn from $F(.|a)$) and assume that an agent wants to purchase a security which pays $y - y'$ where y is some arbitrary random variable. If markets are complete, there should always exist a price $P \geq 0$ for this security, so that the net value of the firm should be $y' - P$. Under MDO, this price can be computed for any possible random variable y': the model places almost no restrictions on which securities the agent can purchase. This is in contrast with nearly all other agency models in which the agent is restricted to a class of distributions.

For tractability, assume that $p'_k(a)$ is well-defined and differentiable. The choice of the manager is denoted $\hat{p}_k = \delta_k + p_k(a)$. One may also reinterpret the current model as a pure reporting

[29]This specification is similar to the Integrated Squared Error used in non-parametric estimation to measure the fit of a density estimation (see for example Pagan and Ullah (1999), p. 24).

[30]See for example Morellec and Smith (2007) for a model in which the principal decides to hedge by buying securities. In their model, the decision to hedge is public and the moral hazard problem occurs after the hedging decision. One may tie their analysis to the portfolio of hedges held by a firm and fully disclosed in financial filings but not the unobserved risk exposure due to project selection and certain off-balance sheet instruments.

manipulation such that the true output is drawn from $F(.|a)$ instead of y (although wages remain based on y) without any consequence on the results. In this alternative interpretation of the results, the manager may be compensated using a short-term indicator of performance y although the firm, in the long-run, will achieve the true performance minus the cost of hedging.

3.2. Agent's Problem

It is helpful to rewrite the problem of the agent as a choice of δ_k rather than $\hat{p}_k$. The agent must choose a probability distribution whose mean is given by the effort provided minus the cost of risk management.

$$\begin{aligned} \sum_{k=1}^{N} \hat{p}_k(a) &= 1 && (10) \\ \sum_{k=1}^{N} \hat{p}_k x_k &\leq \sum_{k=1}^{N} p_k(a) x_k - \sum_{k=1}^{N} C(a, x_k, \hat{p}_k - p_k(a)) && (11) \end{aligned}$$

Replacing $\hat{p}_k$ in this expression, the agent's problem is stated as follows:

$$(A) \qquad \max_{(\delta_k)_{k=1}^{N}, a} \sum_{k=1}^{N} (p_k(a) + \delta_k) u(w(x_k)) - \psi(a)$$

s.t.

$$\begin{aligned} \sum_{k=1}^{N} \delta_k &= 0 && (\lambda) && (12) \\ \sum_{k=1}^{N} \delta_k x_k &= -\sum_{k=1}^{N} C(a, x_k, \delta_k) && (\mu) && (13) \end{aligned}$$

Denote $\mathcal{C}^{x_k}(a, .)$ the inverse of $C_\delta(a, x_k, \delta)$ in δ. The next Proposition characterizes the optimal choice of the agent.

Proposition 3.1. *Suppose that* $\sum_{k=1}^{N} |u(w(x_k))/(x_k - \theta)| < +\infty$. *Then, there exists a solution* *to* (A)*. If, in addition,* $a > \underline{a}$ *is elicited,* $\mu > 0$ *and:*

$$
\begin{aligned}
0 &= \sum_{k=1}^{N} \mathcal{C}^{x_k}(a, \frac{u(w(x_k)) - \lambda}{\mu} - x_k) && (14)\\
\sum_{k=1}^{N} x_k \mathcal{C}^{x_k}(a, \frac{u(w(x_k)) - \lambda}{\mu} - x_k) &= -\sum_{k=1}^{N} C(a, x_k, \mathcal{C}^{x_k}(a, \frac{u(w(x_k)) - \lambda}{\mu} - x_k)) && (15)\\
\psi'(a) &= \sum_{k=1}^{N} p_k'(a) u(w(x_k)) - \mu \sum_{k=1}^{N} C_a(a, x_k, \delta_k) && (16)\\
\delta_k &= \mathcal{C}^{x_k}(a, \frac{u(w(x_k)) - \lambda}{\mu} - x_k) && (17)
\end{aligned}
$$

Corollary 3.1. *The agent does not hedge if and only if* $u(w(y)) = h_0 + h_1 y$*. In particular, a risk-neutral agent does not hedge if and only if the contract is linear.*

Proof: First, $\hat{p}_k = p_k(a)$ for all x_k implies that $u(w(x_k)) = \lambda + \mu x_k$ (necessity). Second, let $u(w(x_k)) = h_0 + h_1 x_k$. Then, the agent must achieve $h_0 + h_1 a$, which can be achieved with no hedging (sufficiency).□

In the model, non-linearities in the contract induce the agent to strategically hedge to align the contract payments with his/her self interest. The manager separates all outcomes in X into outcomes that yield a relatively favorable compensation versus less favorable ones. The optimal hedging choice can be characterized as a simple linear utility threshold: the manager increases (resp. decreases) the likelihood of outcomes above (resp. below) this threshold. Intuitively, in this framework, the manager is given considerable control over the complete probability distribution and thus can exploit any deviation from linearity. As shown in Equations (14)-(17), the hedging choices δ_k depend only on the compensation of the manager and are invariant to the original distribution $F(.|a)$ available to the agent. All other things being equal, the agent reduces the likelihood of high-payoff states versus low-payoff states and increases the likelihood of states with more compensation.

Corollary 3.2. *Assume that* N *is finite,* $C(a, y, \delta) = c(a)\delta^2/2$ *with* $c(a) > 0$ *sufficiently large* *and*[31] $u(x) = x$. *Then:*

$$\lambda = \mathbb{E}(w(y)/p(y)|a) - \mathbb{E}(y/p(y)|a)\frac{\sigma(w(y)/p(y)|a)}{\sigma(y/p(y)|a)} \quad (18)$$

$$\mu = \frac{\sigma(w(y)/p(y)|a)}{\sigma(y/p(y)|a)} \quad (19)$$

$$\delta_k = \frac{w(x_k) - \lambda - \mu x_k}{\mu c(a)} \quad (20)$$

where $p(y)$ *is the probability of outcome* y *and* $\sigma(.)$ *is the standard deviation of the random variable.*

To interpret the model further, I recover the hedging choices explicitly when C is quadratic. The slope of the hedging threshold is captured by the ratio of the volatility of the wage to the volatility of the output (scaled by $p(y)$). When this ratio is high, that is the utility of the agent is very volatile, the agent requires a higher compensation in order not to hedge away high-payoff states (Equation (19)). This aspect is intuitive because an agent with greater wage variability is more sensitive to output shocks. There is a second force which may go against this intuition. In Equation (20), for a given threshold $w(x_k)-\lambda-\mu x_k$, δ_k is decreasing in μ, that is the magnitude of the hedging choice can be decreasing in wage variability. This is because a variable wage makes the agent more sensitive to the cost of hedging and thus incentivizes the agent to hedge less. This second part of the trade-off gives some preliminary intuition why giving incentives to work hard (with high performance-pay coefficients) may not be well-aligned with giving incentives to manage risk (with high values of δ_k).

3.3. Common Compensation Schemes

I analyze next how concavity or convexity in the contract can affect risk choices. For example, concavity of $u(w(y))$ will occur whenever a linear or concave compensation is offered to a risk-averse manager. I show in Corollary 3.3 that the manager concentrates the mass of the distribution on intermediate outcomes, away from the tails of the distribution. Intuitively, the manager seeks insurance against extreme events but, in doing so, may unravel some of the

[31] The results presented here do not require strict risk-aversion by the agent. Further, one may also take $u(x) = x$ as a limiting case when risk-aversion becomes small.

performance-pay sensitivity required in the standard moral hazard problem.

Corollary 3.3. *Suppose that $u(w(y))$ is increasing, continuous and strictly concave (resp. strictly convex) then there exists (y_l, y_h) such that $\inf X < y_l < y_h < \sup X$ and:*

1. *If $x_k \in [\theta, y_l) \cup (y_h, \sup X)$, $\hat{p}_k < p_k(a)$ (resp. $\hat{p}_k > p_k(a)$).*
2. *If $x_k \in (y_l, y_h)$, $\hat{p}_k > p_k(a)$ (resp. $\hat{p}_k < p_k(a)$).*

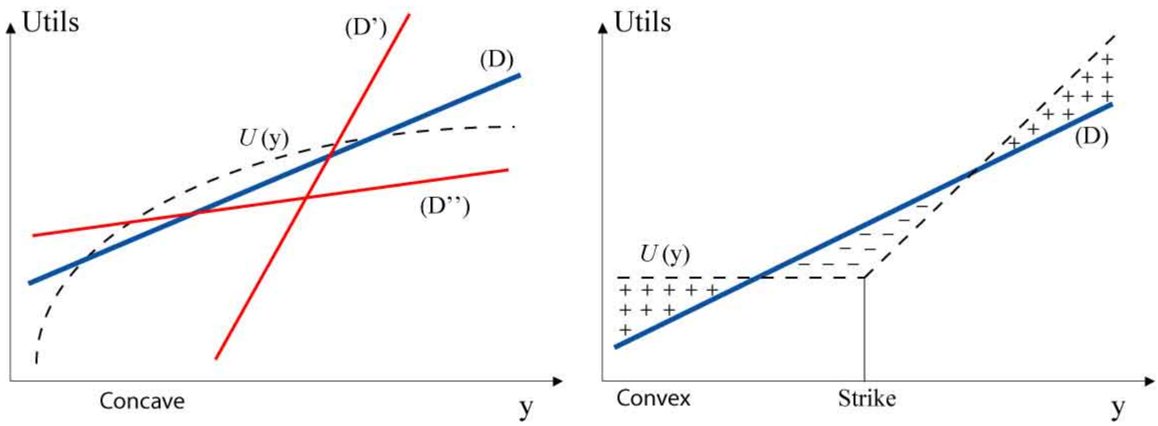

Figure 2. Examples with Concave (left) or Convex (right) wage

The statement can be proved using simple graphical arguments. In the left-hand side of Figure 2, the utility $u(w(y))$ is plotted.[32] One needs to place the hedging threshold $\lambda + \mu y$ which is identified as a straight line (D). It must intercept $u(w(y))$ at least once (by Equation (14)) and no more than twice by concavity. If the two curves intersect once, then among $\hat{F}(.)$ and $F(.|a)$, one distribution must first-order stochastically dominate the other, which is either a contradiction to costly hedging (with (D')) or $\hat{F}$ being preferred to $F(y|a)$ (with (D'')). Then, the hedging strategy must be of the form represented in Figure 2 (in bold) such that $u(w(y))$ is above the hedging threshold only for intermediate outcomes.

Corollary 3.4. *Suppose the agent is risk-neutral and given a Call option, i.e. $w(y) = \max(y - K, 0)$ with $x_1 < K < x_N$. Then, for x_k sufficiently close to K (resp. far from K), $\hat{p}_k < p_k(a)$ (resp. $p_k(a) > \hat{p}_k$). That is, the agent hedges states such that the option matures nearly at-the-money.*

[32]The grid of values for X is set arbitrarily tight in order not to burden the plot.

I analyze next the predictions of the model in the presence of option compensation. In the right-hand side of Figure 2, a Call is given to a manager who is almost risk-neutral (the risk-averse case is plotted in Figure 3). The Call features a kink located at the strike price which must be located in the intermediate region. The manager hedges the kink away so that it becomes unlikely that the option ends at-the-money. A bonus contract will have similar properties. With a standard two-step bonus contract, the agent is paid $\underline{W}$ if y lies below a performance threshold and $\overline{W}$ if the agent beats the threshold. In response to this contract, the manager will concentrate the mass of the distribution strictly above the bonus threshold at the expense of outcomes strictly below it. As observed empirically in DPZ99, the distribution of the output in the model will exhibit bunching to beat certain performance thresholds. A novelty of the approach developed here is that bunching occurs as a result of risk management rather than earnings manipulation between periods and may not be interpreted as actually misreporting current performance.

In addition, in the tails, the manager will shift some of the mass of the distribution from states with high outputs to states with low outputs. In this respect, bonus/option payments may induce the agent to increase the probability of large losses as observed in recent hedging scandals. In M99, Murphy comments: "When expected performance is moderately below the incentive zone, the discontinuity in bonus payments at threshold yields strong incentives to achieve the performance threshold (through counterproductive earnings manipulation as well as through hard work), because the pay-performance slope at the threshold is effectively infinite" (p. 15). He does not explain why corporations tolerate these seemingly adverse incentives. The model presented here presents both sides of the trade-off: the manager modifies risk exposure but in doing so may choose a distribution that is more informative on effort. The hedged distribution may have more probability mass concentrated around the effort level (for example, when the threshold is set slightly below the elicited effort).

Murphy also reports evidence that the shape of the contract offered to CEOs of large corporations varies across industries. He documents that in the industrial sector, 27% of contracts are convex-shaped whereas only 14% in the insurance and finance sector.[33] He does not explain what features of the contracting problem could cause these differences. I provide here a simple

[33]Recall that a contract with pay linear in output will imply that $u(w(y))$ is concave.

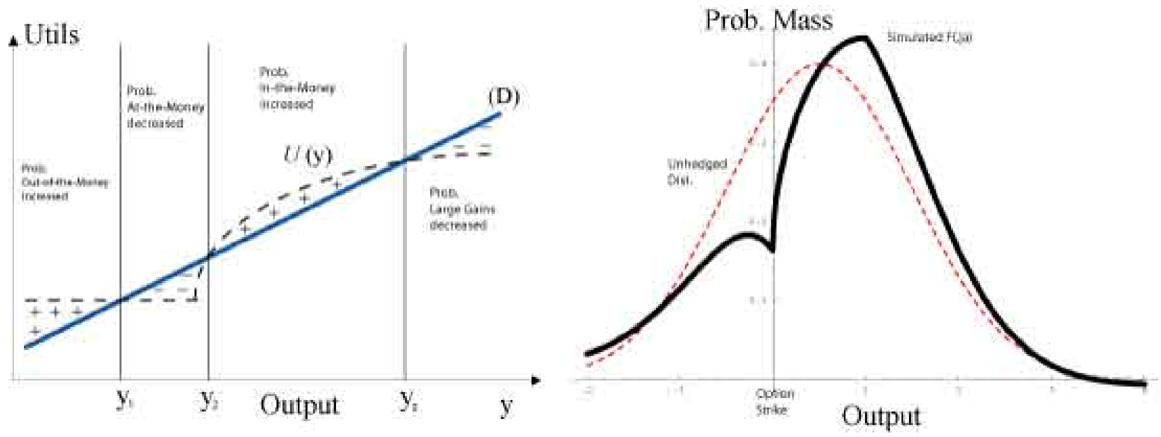

Figure 3. Call with Risk-Aversion

plausible explanation for this evidence. Reasonably, output in the finance industry may be affected by many factors that make inference on the actions of the manager difficult. For example, returns in the financial industry are more driven by systematic risk factors (as evidenced by the low Price to Earnings ratios for banks) which are unrelated to the manager's contribution. By offering a compensation that is linear in output, the firm will induce management to hedge away some of these risk factors.[34] In the industrial sector, on the other hand, the output signal may already be sufficiently informative. To remove excess incentives to hedge, it may be optimal to offer a pay convex in output which induces a compensation offsetting the concavity of the agent's utility function.

3.4. An Example

I develop this argument further with an example which shows if qualitatively the resulting distribution of y may feature the extreme bunching or asymmetry found empirically. I assume that the agent has a CRRA utility function $u(x) = 2\sqrt{x}$. Since I take the contract as given here (this will be relaxed in the next Section), I assume that the agent receives a single Call option with strike normalized at zero and, conditional on this option, chooses an effort $a = 0.5$. The distribution of output is assumed to be Normally distributed with mean a and standard deviation

[34]One may ask why the principal does not directly hedge these risk factors in the compensation, for example by writing a contract based only on idiosyncratic risk. Such an approach would require the principal to have prior knowledge of the systematic risk of the firm (e.g., the stock's beta) whereas, more realistically, this will be private information to the manager or, even, a choice variable.

one. Finally, I assume that the cost of hedging is $C(.5, y, \delta) = e^{r|y-1|}\delta^2$ for any y.[35]

The resulting distribution of y is plotted against that of x, on the right-hand side of Figure 3. Without risk aversion, the model can predict the fact that very few firms announce below the strike price.[36] Note that, unhedged, the distribution of the output should be a smooth bell-shaped distribution. On the contrary, when the manager hedges, the distribution features a kink at zero output and then a large number of firms reporting earnings to beat zero output. This feature is consistent with empirically-observed patterns for corporate earnings (as documented in DPZ99).[37]

4. Problem of the Principal

4.1. Optimal Contract

I consider now the full contract design problem taking into consideration the choice of effort as well as the hedging strategy. To simplify the exposition, I assume in this section that there is no limited liability but the agent is strictly risk-averse. The results are essentially unchanged with a limited liability but a binding limited liability burdens the exposition of the first-order conditions. I use here the first-order approach. It should be noted, however, that it is not necessary for many of the results (including convergence to linearity) and, further, it is valid for a class of problems (see Appendix A).[38]

Proposition 4.1. *The agent is given a contract eliciting $a > \underline{a}$.*

The intuition for Proposition 4.1 is the same as in the standard model. Note that by incentive-compatibility, the compensation of the agent $w(y)$ may not be constant and thus there must be

[35]Since the model is stated discretely, I use here a version of the model with an arbitrarily fine grid to approximate the Normal distribution. The theoretical details are stated formally in Section 5. I use here $\theta = -\infty$ and an exponential penalty. In general, when $\theta = -\infty$, an optimum in the problem of the agent may not exist but given that the exponential penalty effectively prevents the agent from speculating, the solution to the problem of the agent exists (and is actually in closed-form for a given effort).

[36]I use here $\theta = -\infty$ and an exponential penalty. In general, when $\theta = -\infty$, an optimum in the problem of the agent may not exist but given that the exponential penalty effectively prevents the agent from speculating, the solution to the problem of the agent exists (and is actually in closed-form for a given effort).

[37]The model itself cannot predict a discontinuity at zero; however, the empirical evidence does not strongly indicate a discontinuity at the threshold.

[38]The first-order approach is also required in other settings with general risk management problems (Palomino and Prat 2003). By continuity, it will be valid under the standard regularity conditions in Rogerson (1985) as long as the cost of hedging is sufficiently large.

imperfect risk-sharing between the agent and the (risk-neutral) principal. This is in contrast to perfect risk management in which risk-sharing is perfect (i.e., the risk-averse agent does not bear risk). In Section 5, I extend this result to risk-aversion by the principal, showing that imperfect risk-sharing is robust to risk-aversion by both contracting parties.

Using Equation (17), one may rewrite the wage received by the agent $w(x_k)$ as a function of the hedging choice δ_k:

$$u(w(x_k)) = \mu C(a, x_k, \delta_k) + \lambda + \mu x_k \tag{21}$$

I use this Equation to substitute the wage in the contract design problem (P). The following Lemma recovers the participation of the agent as a function of the hedging choice.

Lemma 4.1. *In an optimal contract, the participation of the agent is binding and can be written:*

$$\lambda + \mu a - \psi(a) + \sum_{k=1}^{N} (\hat{p}_k C_\delta(a, x_k, \delta_k) - C(a, x_k, \delta_k)) = b \tag{22}$$

The contract design problem can be simplified into the choice of a, (λ, μ) (i.e., which outcomes should be hedged) and δ_k (i.e., how much hedging should be elicited).

$$(P_a) \qquad \max_{(\delta_k)_{k=1}^N, a, \lambda, \mu} \sum_{k=1}^{N} (\delta_k + p_k(a))(x_k - u^{-1}[C_\delta(a, x_k, \delta_k)\mu + \lambda + \mu x_k])$$

s.t.

$$\sum_{k=1}^{N} \delta_k = 0 \qquad (\alpha) \tag{23}$$

$$\sum_{k=1}^{N} \delta_k x_k = -\sum_{k=1}^{N} C(a, x_k, \delta_k) \qquad (\beta) \tag{24}$$

$$\mu = \psi'(a) - \mu \sum_{k=1}^{N} p'_k(a) C_\delta(a, x_k, \delta_k) \qquad (\gamma) \tag{25}$$

$$b - \lambda - \mu a = -\psi(a) + \mu \sum_{k=1}^{N} [(\delta_k + p_k(a)) C_\delta(a, x_k, \delta_k) - C(a, x_k, \delta_k)] \qquad (\tau) \tag{26}$$

In the next Proposition, I state the first-order conditions associated to this Problem.

Proposition 4.2. *Let a be the elicited effort. The optimal contract satisfies:*

$$\tau = \sum_{k=1}^{N}(\delta_k + p_k(a))\frac{1}{u'(w(x_k))} \tag{27}$$

$$\gamma\psi'(a) = \sum_{k=1}^{N}(p_k(a) + \delta_k)\frac{1}{u'(w(x_k))} (\lambda + \mu x_k - \mu C_\delta(a, x_k, \delta_k) - b - \psi(a)) \tag{28}$$

$$w(x_k) = u^{-1}(\lambda + \mu x_k + \mu C_\delta(a, x_k, \delta_k)) \tag{29}$$

$$w(x_k) = -\alpha + (1-\beta)x_k - \beta C_\delta(a, x_k, \delta_k) - \mu\gamma C_{\delta,a}(a, x_k, \delta_k) - \mu C_{\delta,\delta}(a, x_k, \delta_k) (\delta_k + p_k(a))(\frac{1}{u'(w(x_k))} - \tau - \gamma\frac{p'_k(a)}{\delta_k + p_k(a)}) \tag{30}$$

One aspect of the framework is that earnings management can be examined given that the wage offered to the manager is endogenously recovered from within the model.[39] I discuss next several local characteristics of the contract. In Equation (30) one may recognize the incentive Equation of the standard model, i.e. $S = 1/u'(w(x_k)) - \tau - \gamma p'_k(a)/(p_k(a) + \delta_k)$ (see Equation (7) in Holmström (1979)). Without hedging, this term should be set at zero. Here, it may not be zero but the model predicts that qualitatively, when the condition is not met and strictly negative (i.e., without risk management it would have been optimal to increase the wage), this also leads to a greater wage *all other things being equal*. Note also that if for a set of outcomes, C is quadratic and does not depend on a, the standard Equation S will be zero only if the compensation is linear. This is because a linear compensation offsets incentives to hedge more by the principal (and thus the Equation of the standard model may apply).

I give next a sufficient condition that guarantees that it is optimal to elicit some risk management.

Corollary 4.1. *Suppose that $C_{\delta,\delta}(a, x_k, 0) = C_{\delta,a}(a, x_k, 0) = 0$ for all a, x_k, then for all k, $\delta_k = 0$ cannot be optimal.*

[39]For example, Crocker and Huddart (2006) is one of the only papers in which the contract offered to the executive is endogenous and solves an earnings management problem. However, some aspects of the intertemporal preferences of the manager are caused by a moral-hazard problem which is not explicitly modeled.

Proof: If not hedging is optimal, $u(w_k) = \lambda + \mu x_k$. But then, by Equation (30), $w(x_k) = \alpha + \beta x_k$, but since u is strictly concave, this is a contradiction.□

Hedging will generally be desirable if the second-derivative of the cost of hedging is zero and the principal is risk-neutral. The intuition for this result is that some hedging can always raise the usefulness of the output signal at very little cost. An important interpretation of this result is that, while the principal would not hedge in first-best, the informational friction creates an incentive to use some risk management.

4.2. Convergence to Linearity

As with perfect risk management, I will argue that a simple linear contract can be optimal in certain situations. To begin with, I analyze the shape of the contract over only those realizations of y that are sufficiently large.

Proposition 4.3. *Suppose that for all a and δ, (i) $|C_\delta(a, x_k, \delta)|$, $|C_{\delta,\delta}(a, x_k, \delta)|$ and $|p'_k(a)|$ are bounded by a number that does not depend on k, (ii) either $C_{\delta,\delta}(a, x_k, 0)$ or $p'_k(a)$ converges to zero, (iii) $C_{\delta,a}(a, x_k, 0) = 0$. Then, $w(x_k)$ converges to $-\alpha + (1-\beta)x_k$ as x_k grows large.*

In the model, the linear part in the compensation performs well at providing an efficient risk allocation from the perspective of the principal. For states in the tail of the distribution, this concern dominates any improvements in the likelihood ratio. Consistently with this result, M99 documents many non-linearities in the compensation of executives but explains that for events that are "outside of the incentive zone" (i.e., in the tail) compensation is essentially linear. Note that if the principal is risk-neutral, pay will tend to become linear in output for large values of x_k. Interestingly, this result presents an apparent similarity with Diamond (1998). Diamond explains that, as the size of (all of) the firm's cash flow becomes large relative to the cost of effort, the optimal compensation scheme converges to a linear function. In Proposition 4.3, I show that the same is true here considering only those states of the world in which cash flows are sufficiently large. Conceptually, however, this similarity is misleading because the cause of the result in Diamond is different from mine. In his setting, a non-linear wage causes distortions in the choices of the agent which grow large. Here, on the other hand, the cost of these distortions remain small because these events are unlikely. Further, when applied jointly with risk-aversion by the agent, Diamond's approach yields a contract such that the $u(w(y))$ (and not $w(.)$) is linear.

An additional implication of this result is that $C_{\delta,\delta}(a, x_k, \delta) = 0$ is not needed for no hedging to be suboptimal (under the conditions of Proposition 4.3). That is, one should observe hedging in the tails. In addition, one may easily verify that a risk-averse agent indeed reduces the likelihood of these large positive gains in order to shift probability mass toward intermediate outcomes. This rationalizes one aspect of the empirical evidence on earnings management: managers tend to reduce the likelihood of very large gains.

A different limiting contract can be obtained when the cost of hedging becomes small. To consider this case, I state a sequence of problems with a cost function $C^j(a, x_k, \delta) = C(a, x_k, \delta)/j$ ($j > 1$) and C verifies the conditions of MDO. As j becomes large, the cost function becomes small. In addition, I make the assumption that $C(a, x_k, \delta_k)$ becomes large (resp. $\frac{C_\delta(a,x_k,\delta_k)}{C(a,x_k,\delta_k)}$ is bounded away from zero) when δ_k converges to $-p_k(a)$ or $1 - p_k(a)$ for all (x_k, a). In intuitive terms, this assumption means that the elasticity of the cost of hedging to an increase in hedging does not become too small.

Proposition 4.4. *As k becomes large, the contract converges to a unique contract:*

$$u(w(x_k)) = \psi'(a^*)(x_k - a^*) + b \tag{31}$$

The effort a^* is the first-best effort when the agent can hedge at no cost. I argue here that the contract should become linear as the cost of hedging becomes small, but provided that the cost of hedging increases sufficiently fast for hedging policies close to infeasibility. Intuitively, when the cost of hedging becomes small, only small non-linearities are required to give incentives to hedge; important non-linearities, on the other hand, may lead to hedging cost that are large.

To recover linearity, assume that $u(.)$ becomes arbitrarily close to linear. It will then be true that, without risk management, the contracting problem can be solved in first-best (transferring all the risk to the agent). This will also occur in **?**), up to the utility normalization. It should be noted that the problem with a risk-neutral agent does not pin down a unique optimal contract, while the current framework delivers a unique linear contract as a (limiting) solution to the moral-hazard problem: $w(x_k) = \psi'(a^*)(x_k - a^*) + b$. This result gives further support to the use of linear contracts when the agent is nearly risk-neutral and risk management is nearly free.

4.3. Numerical Analysis and Earnings Management

I develop a simple numerical example to quantitatively assess the shape of the contract as well as the predicted hedged distribution. The model is discretized in two effort choices $a \in \{0, \overline{a}\}$. Conditional on $a \in \{0, \overline{a}\}$, the distribution of outputs is normally distributed with mean a and variance σ^2. The set of outputs is discretized over $N = 200$ points $(x_k)_{k=1}^N$ and such that each point has equal probability conditional on $F(x|\overline{a})$. The manager has a CARA utility function $u(x) = -e^{-rx}$. Finally, the cost function is set equal to $C(a, x, \delta) = e^{-|x-a|} c_h \delta^2/2$. The model is solved using the multi-start non-linear solver MSNLP.

The benchmark parameters are set as follows: $\sigma = 1$, $\overline{a} = .5$, $r = 1$, $\psi(\overline{a}) = .1$, $c_h = 1$ and $b = -1$. The results are shown in Figure 4.3 under four treatments: 1. First-Best Contract, 2. Second-Best without Hedging (i.e., standard agency problem), 3. Second-Best with Hedging, 4. Second-Best with Hedging but constrained to a Linear contract. The optimal contract without hedging (2.) is concave because risk-aversion by the agent becomes the most important factor (versus the increasing likelihood ratio). The optimal contract such that the agent does not hedge (4.), on the other hand, is convex as the agent must be compensated to offset risk-aversion. In this example, the optimal contract with hedging lies in-between these two extremes but turns out to be close to linear (suggesting that convergence to linearity occurs fast). On the right-hand side, the distribution of output is reported with and without hedging. I show that the optimal contract induces the agent to hedge away extreme risks which makes the distribution more precise on the actions of the agent.[40]

[40] I also calculated the expected profit of the principal. In this example, the principal is worse-off under 3. than 2. (although the difference is fairly small).

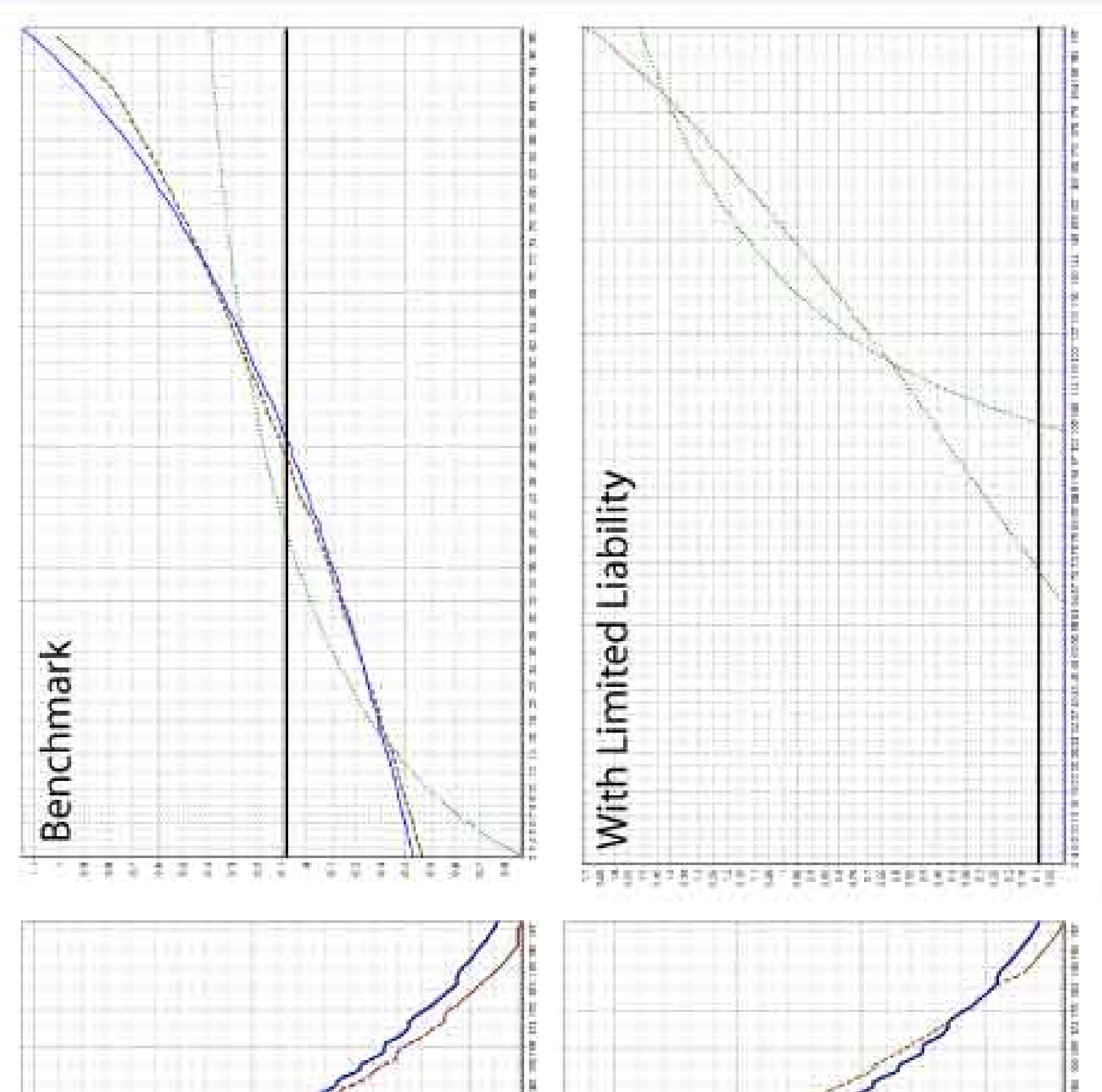

Figure 4. Simulations: Hedging (left), Wage (right)

Next, I set a limited liability constraint at -1 to test its impact on the optimal contract and on the hedging choices of the agent. Case 4. was no longer possible and therefore is omitted. In the bottom right-hand side of Figure 3, I represent the wage in the three other cases. The optimal contract without hedging takes the form of a binding limited liability for low outcomes and a concave part for higher outcomes. On the other hand, with hedging, the limited liability binds over a smaller region and the contract is close to linear over most of the outcomes. Here, the agent has excess incentives to increase the likelihood of bad outcomes which forces the principal to pay above the limited liability more frequently. In response to the contract, the agent hedges away events close to the point where the limited liability is not binding and increases the likelihood of low outcomes.

The model with limited liability produces a shape that is similar to Figure 3 and has the same interpretation. Two features are distinctive of the framework. First, the model produces a divot, i.e. a hump shape close to the median. This feature is representative of what we observe in the cross-section of earnings: few firms report earnings slightly below the mean (or previous earnings) while many firms seem to bunch and slightly beat the mean. Since there are no particular reasons for the production technology to produce these features, this is often deemed to be puzzling. This paper is among the first to recover it as a solution of a fully optimal contract. Second, the manager reduces the likelihood of very good outcomes which have been empirically observed to be relatively rare in corporate earnings disclosure and increase the likelihood of large losses. It should be noted that the model can have different implications from the existing literature in terms of how one should respond to earnings management. As in the current literature, the efficiency of contracts can be improved with better disclosures; however, earnings management is not necessarily undesirable, in the sense that the principal may prefer (if he/she could) not to constrain the agent to $F(.|a)$. In these simulations, the hedged distribution $\hat{F}(.)$ appears to be more concentrated around a than under $F(.|a)$ which leads to greater incentive gains: the divot may be interpreted as one way the principal creates distributions that are more informative on the actions of the agent.

4.4. Further Comparative Statics

I attempt next to discuss what the first-order conditions of the problem imply in terms of how hedging should respond to the payoff of the state x, the likelihood ratio $LR = p'_k(a)/p_k(a)$ and the probability $p_k(a)$ of a state of the world. For a particular state of the world, it is helpful first to rewrite Equations (29)-(30) in terms of their continuous analogue (omitting the indices):

$$
\begin{aligned}
w &= u^{-1}(\lambda + \mu x + \mu C_\delta(a,x,\delta)) && (32)\\
w &= -\alpha + (1-\beta)x - \beta C_\delta(a,x,\delta) - \mu\gamma C_{\delta,a}(a,x,\delta) \\
&\quad -\mu C_{\delta,\delta}(a,x,\delta)((\delta+p)(\frac{1}{u'(w)} - \tau) - \gamma p LR) && (33)
\end{aligned}
$$

In the model, there are only two endogenous variables that depend on the state of the world, w and δ. Provided the number of states is large and the distribution of $F(.|a)$ is sufficiently spread-out across states, the multipliers of the problem should not vary much for a comparative static applied to only one state (there would be, for example, no effect on the multipliers if the model was specified with $F(.|a)$ being a continuous distribution). As an approximation to a continuous problem, I take here the multipliers as constants. Then, I view Equations (32) and (33) as two Equations in two unknowns (w, δ) where all the other terms are taken as exogenous constants.

To obtain unambiguous comparative statics, I shut down some of the cross-effects in the cost of hedging. I make the following assumptions: For all a, x, δ, $C_{\delta,\delta,a}(a,x,\delta) = C_{\delta,x}(a,x,\delta) = C_{\delta,a,x}(a,x,\delta) = C_{\delta,a,x}(a,x,\delta) = 0$, $C_{\delta,\delta,\delta}(a,x,\delta) < 0$ and bounded away from zero. This condition restricts the cross-effects in the model and are not all necessary for each comparative static taken separately (and may be verified only on the state considered in the comparative static). The following comparative statics are obtained from the Implicit Function theorem applied on Equations (32) and (33).

Corollary 4.2. *The following comparative statics hold for a given state of the world:*

(i) For states such that the wage is sufficiently large, $\partial\delta/\partial x < 0$, i.e. the manager hedges against states with large payoffs.

(ii) If $\gamma LR \geq 0$ is sufficiently large, $\partial\delta/\partial LR > 0$, i.e. the manager hedges against states with lower likelihood ratio.

(iii) If $\gamma > 0$ and $|LR|$ is sufficiently large with $Sign(LR) > 0$ (resp. $Sign(LR) < 0$), $\partial\delta/\partial p > 0$ (resp. $\partial\delta/\partial p < 0$), i.e. if $p'_k(a)/p_k(a)$ is monotonic, the manager will produce a hump-like shape on the distribution $\hat{F}(.)$.

Many of these comparative statics are intuitive.[41] First, in the problem of the agent (studied in Section 3), the agent always has greater incentives to hedge against states with large payoffs. I show here that it would be too expensive to offset these incentives in the contract design problem since this would require a compensation that makes the agent risk-neutral to output shocks. Second, the manager is induced to increase the likelihood of states with a high likelihood ratio. These states are informative on the actions of the manager and thus the principal gains from eliciting a hedging strategy that increases their likelihood. Third, the model predicts a hump-shape in the distribution of $\hat{F}(.)$. In other terms, suppose that the distribution $F(.|a)$ has a unique mode (e.g., a bell-shaped distribution) and the likelihood ratio is zero close to the mode of the distribution and changes fast close to the mode. Then, as p increases as one moves toward the mode, the model will be in a region where $\partial\delta/\partial p < 0$. This will work to flatten the distribution. As the likelihood ratio increases close to the mode, the model will be in a region where $\partial\delta/\partial p > 0$, producing a peak in the distribution. To summarize the model can alter a bell-shaped histogram for $F(.|a)$ into an histogram for $\hat{F}(.)$ that is first flat and then peaks higher than the original distribution.[42]

5. Extensions

I propose next several applications of the framework, exploring the implications of the model for exogenous capital market frictions and income smoothing. To focus the attention on these practical issues, I use here on a simplified version of MDO. I assume now that the support of y

[41] One may guarantee that $\gamma > 0$ if the principal sets the maximum effort that the manager can choose. Further, by continuity, γ should be positive provided the cost of hedging is sufficiently large (since the property is true in Holmström (1979) and the current model becomes equivalent to his when the cost of hedging is large).

[42] In the continuous limit, the histogram will correspond to the density.

is $X = \mathbb{R}^+$ and $F(.|a)$, the unhedged distribution has a density $f(.|a)$. The concept of MDO is adapted to reflect a continuous set of outcomes. I assume that there exists a positive convex cost function $C(x)$ such that $C(0) = C'(0) = 0$, C is positive, $C'(x) < 0$ (resp. $C'(x) > 0$) for $x < 0$ (resp. $x > 0$) and $\lim_{x \to -\max_a f(x|a)} C(x) = +\infty$. The set of distributions available to the agent for an effort a, $\Gamma(a)$, is assumed to be the set of distributions $\hat{F}(.)$ with a density $\hat{f}(.)$ and such that $\int \hat{f}(y) y dy = a - \int \hat{f}(y) C(\hat{f}(y) - f(y|a)) dy$. The hedging choice over outcome y is denoted $\delta(y) = \hat{f}(y) - f(y|a)$. As in most of the literature, I assume that an optimal contract exists and the contract $w(y)$ is a smooth function of output y.

5.1. Additional Capital Market Frictions

I showed earlier that risk management is desirable when the principal is risk-neutral. I extend now the analysis to situations such that the firm is owned by investors that are not well-diversified or are facing additional exogenous capital market frictions. To model these frictions, I assume now that the principal is risk-averse.

Let $v(.)$ denote the utility function of the principal. It is assumed to be twice-differentiable, strictly increasing and strictly concave. I state first the first-best problem (effort and hedging are chosen by the principal).

$$(P_{fb}) \qquad \max_{(\delta_k)_{k=1}^N, a, \lambda, \mu} \int (f(y|a) + \delta(y)) v(y - w(y)) dy$$

s.t.

$$\int \delta(y) dy = 0 \qquad (\tilde{\alpha}) \qquad (34)$$

$$\int \delta(y) y dy = -\int C(\delta(y)) dy \qquad (\tilde{\beta}) \qquad (35)$$

$$\psi(a) + b \leq \int u(w(y))(\delta(y) + f(y|a)) dy \qquad (\tilde{\tau}) \qquad (36)$$

Proposition 5.1. *In first-best, no hedging cannot be optimal.*

In comparison to the previous Section, some hedging is always desirable as part of the first-best solution to the model when the principal is risk-averse. This is because some risk can be hedged at very little cost and it is very intuitive in the context of MDO. I focus now on the second-best problem. In contrast to Proposition 4.1 (when the principal is risk-neutral), it may now be optimal to offer a flat contract and elicit $\underline{a}$. First, the distribution $f(y|\underline{a})$ may be easier to hedge than other distributions. Second, by setting $w(y)$ constant, the principal can elicit any risk management strategy (the agent being indifferent) and attain a first-best solution to the risk management problem (although the chosen effort will be $\underline{a}$).

Suppose first that a flat contract is optimal. Then, one needs to substitute in Problem (P_{fb}), $w(y)$ by $u^{-1}(\psi(\underline{a}) + b)$ (so that the participation binds) and set $a = \underline{a}$. The risk management strategy chosen by the principal solves the resulting problem. Clearly, this problem will be similar to the problem faced by a risk-averse manager compensated with a linear wage studied in Corollary 3.3. Note that, in this situation, a risk-averse principal may be counter-intuitively fully insuring the agent whereas a risk-neutral principal would always transfer some risk to the manager. This is because solving the risk-management problem (which is easier when wage is flat) is more important for a risk-averse principal.

In the rest of this Section, I assume that eliciting $a = \underline{a}$ is not optimal and characterize the optimal hedging choices in this case.[43] As in the previous Section, I state the contract design problem, by substituting $w(y)$ in the problem of the principal from the incentive-compatibility condition. Problem (P_a) is the same as before, except that is it now states with continuous outcomes and a risk-averse principal. The contract design is restated by introducing the utility $v(.)$.

$$\max_{(\delta(.),a,\lambda,\mu} \int v(y - u^{-1}[C'(\delta(y))\mu + \lambda + \mu y])dy$$

[43]I omit the case in which $\underline{a}$ is optimal but not with a flat contract since this case is not particularly interesting.

s.t.

$$\int \delta(y)dy = \quad 0 \qquad (\alpha) \qquad (37)$$

$$\int \delta(y)ydy = \quad -\int C(\delta(y))dy \qquad (\beta) \qquad (38)$$

$$\mu = \quad \psi'(a) - \mu \int f_a(y|a)C_\delta(\delta(y))dy \qquad (\gamma) \qquad (39)$$

$$b - \lambda - \mu a =$$

$$-\psi(a) + \mu \int \quad [(\delta(y) + f(y|a))C'(\delta(y)) - C(\delta(y))]dy \quad (\tau) \qquad (40)$$

This problem is the same as Problem (P_a) except that the principal is now risk-averse and the model is stated over a continuum of outcomes. The first-order condition are similar to those derived earlier and have the same interpretation. I provide here an additional result.

Proposition 5.2. *The ratio of marginal utility* $v'(y - w(y))/u'(w(y))$ *cannot be constant, i.e. risk-sharing is imperfect.*

A well-known result in the standard agency model is the violation of perfect risk-sharing (Holmström 1979). As stated in most textbooks, should risk management fully resolve risk-sharing frictions? I show here that this is not the case. The risk taken by the agent is still required for incentive purposes and risk-sharing must remain imperfect.

I discuss next whether hedging is desirable when the principal is risk-averse. As suggested in first-best, since hedging is now used without informational frictions, a preliminary intuition would suggest that risk management should remain optimal here. It turns out that this is not the case here, i.e. no hedging is not necessarily suboptimal when the principal is risk-averse. I analyze next the shape of the optimal compensation scheme when no hedging is optimal.

Proposition 5.3. *Suppose that for any parameter values* h_0, h_1 *and* $w(y)$ *is given by* $w(y) = u^{-1}(h_0 + h_1 y)$, $v'(y - w(y))$ *cannot be linear in* $v'(y - w(y))/u'(w(y))$. *Then, no hedging cannot be optimal. Else, if no hedging is optimal and* $\lim_{z \to +\infty} u'(z) = 0$, $v'(y - w(y))/u'(w(y))$ *is strictly increasing for* y *sufficiently large.*

I give a simple condition on utility functions such that no hedging is suboptimal. Unlike with a risk-neutral principal, no hedging may occur here when this condition is not satisfied.

This aspect goes against the intuition that more financially constrained firms should always be observed to hedge more. This intuition would be valid if the principal had control over hedging. Here, given that hedging must be elicited through an appropriate compensation contract, risk management to increase precision may require the principal to hold some residual risk. This can be costly if the principal is in financial distress.

The model rationalizes why firms such as American Airlines stopped hedging during periods of financial distress while a more financially solid company such as JetBlue still hedges most of its oil expenses. While one may not conclude that, even in this model, risk-averse principals should always elicit less hedging, the ambiguous interaction between risk-sharing and incentives to hedge is worth pointing out. More generally, this finding is consistent with the ambiguous empirical relationship between financial distress and hedging choices (Mian 1996), which goes against some existing models of risk management (Froot et al. (1993), Smith and Stulz (1985)) in which there is no agency friction. Remarkably, when hedging is not desirable, the ratio of marginal utilities will be increasing for outcomes sufficiently large. This monotonicity property is similar to the standard agency problem (Holmström 1979) but occurs without the monotone likelihood ratio property.

5.2. Income Smoothing

Recently, the analysis of the time series of earnings has provided several novel tests of earnings management, focusing on the predictability of future accounting variables using current information (such as, among others, Dechow, Sloan and Sweeney (1995)). To address this issue theoretically, I recast the model as a multi-period problem and analyze its predictions in terms of the dynamics of reported earnings. I follow the standard methodology developped in Debreu (1972) to map the static model into a multi-period one with a simple change of notations. This Section will show that the framework accommodates many features linked to the time series of earnings and stock returns and does not presuppose (quite on the contrary) that managers cannot shift output across period.

Let $\mathbb{y} = (y_1, ..., y_T) \in \mathbb{X}$ where $y_i > 0$ for all i, be a sequence of realizations of *returns* from period $t = 1, ..., T$; $F(.|a)$ denotes the (multivariate) distribution of $\mathbb{y}$ when the agent does not hedge. I define next $Y = \prod_{t=1}^{T} y_t$ as the total return (with mean a) over the period and assume that the agent may choose any distribution $\hat{f}(y_1, ..., y_T)$. Intuitively, the manager will now choose a multivariate distribution but the aggregate profit remains a scalar given by Y. I redefine next MDO over the multivariate distribution (where C satisfies earlier assumptions):

$$\int_{\mathbb{X}} \hat{f}(\mathrm{y}) \prod_{t=1}^{T} y_t d\mathrm{y} \leq a - \int C(f(\mathrm{y}|a) - \hat{f}(\mathrm{y}))d\mathrm{y} \tag{41}$$

The feasibility condition for $\hat{f}(.)$ to be a density is written:

$$\int_{\mathbb{X}} \hat{f}(\mathrm{y}) = 1 \tag{42}$$

It should be clear that, by expressing the model as multiple period in this manner, the previous framework is essentially unchanged. In this Section, I suppose that a is chosen ex-ante (the previous assumptions remain valid) and the manager is compensated at the end of the period $w(1'\mathrm{y})$ as a function only of the aggregate performance. As before, let $\delta(\mathrm{y}) = f(\mathrm{y}|a) - \hat{f}(\mathrm{y})$.

The optimal hedging threshold is given by the first-order stated in the multivariate case:

$$u(w(\prod_{t=1}^{T} y_t)) - \lambda - \mu\mathrm{y} - \mu \prod_{t=1}^{T} y_t \mathcal{C}(\delta(\mathrm{y})) = 0 \tag{43}$$

The graphical analysis of the dynamic case is the same as in the static case. More generally, the properties of the optimal contract developed previously are preserved in the multi-period model (because this is only a reformulation of the previous model). More interestingly, the model yields several predictions on the dynamics of returns. Let $1 < t' < T$ be a time period, and let $\underline{y}^{t'} = \prod_{t=1}^{t'} y_t$ (resp. $\underline{x}^{t'} = \prod_{t=1}^{t'} x_t$) be the hedged (resp. unhedged) return prior to date t' and $\overline{y}^{t'} = \prod_{t=1}^{t'} y_t$ (resp. $\overline{x}^{t'} = \prod_{t=1}^{t'} x_t$) the hedged (resp. unhedged) return after date t'.

Proposition 5.4. *The following holds:*

$$cov(\underline{y}^{t'}, \overline{y}^{t'}) \leq cov(\underline{x}^{t'}, \overline{x}^{t'}) - (\mathbb{E}(\underline{x}^{t'})\mathbb{E}(\overline{x}^{t'}) - \mathbb{E}(\underline{y}^{t'})\mathbb{E}(\overline{y}^{t'})) \tag{44}$$

The inequality is strict when the manager hedges.

Comparing the covariance between periods under hedging to the original covariance, one of the following facts must be true when hedging occurs: the manager reduces the covariance of returns across periods (i.e. $cov(\underline{y}^{t'}, \overline{y}^{t'}) < cov(\underline{x}^{t'}, \overline{x}^{t'})$) or changes expected returns per period (the term $\mathbb{E}(\underline{x}^{t'})\mathbb{E}(\overline{x}^{t'}) - \mathbb{E}(\underline{y}^{t'})\mathbb{E}(\overline{y}^{t'})$). Both forms of hedging correspond to income smoothing, although their nature is different. In the first case, the manager collects positive shocks from high periods to raise returns in the next periods, and vice-versa. Whenever the covariance between returns is positive and hedging is not too intense (e.g., the cost of hedging is sufficiently large), this will lead to a sequence of hedged cash flows that is smoother than the unhedged cash flows would have been. In the second case, the manager alters the mean return per period. It is easy to verify for example that this case will generally occur when the compensation of the manager only depends on the initial periods.[44]

The divot in the cross-section of earnings described earlier is one aspect of the empirical evidence. However, the time series of periodic income also exhibits an S-Shaped response of the market price to current earnings (captured here as "current" return). An S-Shaped response means that stock prices react much to small differences in reported earnings but not much to large differences. Most observers take the divot in earnings and S-Shaped response of the market price as two aspects originating from the same underlying factors (see also Crocker and Huddart (2006) for a summary of the empirical evidence as well as recent work). In this model, the market value of the firm at t', after $\underline{y}^{t'}$ has been announced, is captured by $MV(\underline{y}^{t'})$:

$$MV(\underline{y}^{t'}) = \frac{\int \overline{y}^{t'} \hat{f}(\underline{y}^{t'}, y_{t+1}, \ldots, y_T) dy_{t+1} \ldots dy_T}{\int \hat{f}(\underline{y}^{t'}, y_{t+1}, \ldots, y_T) dy_{t+1} \ldots dy_T} \tag{45}$$

This conditional expectation is however difficult to analyze very generally. It may be concave or convex and even decreasing. However, my main purpose here is to test whether a simple parametrization model can generate a pattern that is consistent with the S-Shaped curve observed empirically. Assume that $T = 2$ and the support of x_1 and x_2 are i.i.d. and distributed uniformly on $[0, 2]$ (the location of the support does not affect the results). Finally, the cost of hedging is set as in Section 3.4. Suppose that the manager is risk-neutral and compensated

[44] The argument invoked for this result is fairly weak as it is purely driven by the constraints and the fact that no hedging cannot be optimal. Note that, as a result, if an additional constraint dictates that hedging cannot affect expected unconditional return in each period (e.g., due to regulatory supervision), the manager will always choose to reduce the covariance of returns across periods (and vice-versa).

Figure 5. Earnings Response

with a simple Call option with maturity date 2 and strike 1. In Figure 5, the hedged density $\delta(\mathrm{y})$ is plotted. Given the convexity of the option contract, the manager raises the likelihood of extreme events. When y_1 is low, for example, there are almost no gains to having y_2 high (since the option will be out-of-the money) and similarly when y_1 is high, there are great gains to having y_2 high. Plotting the corresponding MV calculated in Equation (45) as a function of y_1, the following response to y_1 which exhibits the familiar S-Shaped profile where the response is steeper near the median return.[45]

6. Conclusion

This paper presents a simple framework in which hedging decisions are part of a traditional agency-theoretic model. A basic trade-off is explored: on the one hand, hedging will induce the agent to exploit the compensation schedule to his/her advantage; on the other hand, the principal may be able to design contracts eliciting a more informative output signal. Conventional wisdom suggests that hedging should resolve some exogenous capital market imperfections (such as differences between internal and external cost of capital). Here, I endogenize the imperfec-

[45] I do however recognize that, with risk-neutrality and a uniform distribution, the response that is predicted is less steep than in the data, and does not generate anything near a discontinuity in the response (although the data is ambiguous on whether or not there is a discontinuity).

tions as a result of a moral-hazard problem between owners and managers and show that often criticized financial reporting concerns can be the essence of a sound hedging policy. That is, the executive contracts that we observe empirically may be geared to improve the informativeness of the output signal.

Appendix A: Complements

Additional Signal under Perfect Risk Management

I discuss next how the presence of an additional output signal x that cannot be manipulated may affect the contractual arrangement under perfect risk management (as described in Section 2). There are two main reasons for including this second signal. First, a model with a second signal extends the classic moral hazard problem since the principal may choose to ignore the managed signal y and use only x. Further, the theory allows me to compare how the managed signal is used in the contract versus another signal whose risk is not controlled by the agent. Second, in most realistic settings, there is some information available to the principal which is not fully controlled by the manager. On theoretical grounds, the presence of only managed information may be giving too much importance to the managed signal in two respects: (i) the principal cannot ignore it and elicit effort, (ii) any risk in the contract will be removed by risk management.

I maintain here the assumption that the agent is strictly risk-averse. I assume that the additional signal x has a density $g(.|a)$ with mean a and cannot be hedged; the wage offered by the principal is denoted $w(x, y)$. Further, I assume that risk management occurs after x is revealed (or can be conditional on the realization of x). This seems fairly reasonable as a model of how executives shift operational risk to financial intermediaries. It follows from Lemma 2.1 that, for a given x, an optimal contract must elicit $y = \rho(x)$ constant. Let $w(x, y) = \phi(x)$ be the compensation given to the manager conditional on x and $y = \rho(x)$ (and, if $y \neq \rho(x)$, $w(x, y) = \underline{w}$). I make several additional assumptions. First, I assume that the monotone likelihood ratio property holds and $x - w(x, y)$ is the net transfers received by the principal. Second, I assume that the agent must choose an action $a \in [\underline{a}, \overline{a}']$ where $\overline{a}'$ is chosen by the principal. This assumption corresponds to the idea that the principal can monitor the tasks done by the agent when these tasks are done diligently but cannot observe shirking; in my problem, it excludes situations in which the agent deviates to more effort.[46]

Proposition A. 1. *There exists $x_0 \in [\theta, +\infty]$ such that:*

(i) If $x \leq x_0$, $\rho(x) = \theta$ and $\frac{v'(x-\phi(x))}{u'(\phi(x))} = s_2 \frac{f_a(x|a)}{f(x|a)} + s_3$ where s_2 and s_3 are two positive constants.

[46] In the standard model, Holmström (1979) shows that this situation does not occur, i.e. the Lagrange multiplier on the incentive-compatibility condition is strictly positive; unfortunately, the same argument, to the best of my knowledge, does not apply in my setting. Clearly, if θ is sufficiently small, since the managed signal becomes nearly useless, the solution of the model will be close to that of Holmstrom and thus this assumption will no longer be necessary.

(ii) If $x > x_0$, $u(\phi(x)) - u(\underline{w}) = \tilde{\lambda}(\rho(x) - \theta)$ where $\tilde{\lambda}$ is a positive constant.

Further, if $x_0 = \theta$, $\tilde{\lambda} = \psi'(a)$

Proof: Suppose that the principal offers a contract that pays $\phi(x)$ when $y = \rho(x)$ is realized and $\underline{w}$ for all other outcomes. Note that this contract minimizes payment for off-equilibrium outcomes and thus is an optimal way to provide incentives. In response to this contract, the agent may deviate to a lottery with support θ and $\rho(x)$. Let $p(x)$ (resp. $1 - p(x)$) denote the probability that $y = \theta$ (resp. $y = \rho(x)$). The program of the agent is then written as follows:

$$\max_{a,p(x)} \int g(x|a) \left(p(x)u(\underline{w}) + (1 - p(x))u(\phi(x))\right) dx - \psi(a)$$

s.t.

$$\int g(x|a) \left(p(x)\theta + (1 - p(x))\rho(x)\right) dx \leq a \qquad (\tilde{\lambda} \tag{A-1}$$

In the above program, the agent maximizes utility subject to the distribution of the signal y having a mean below a. Let L_1 denote the Lagrangian associated to this problem; taking the first-order condition with respect to $p(x)$ yields:

$$\frac{\partial L_1}{\partial p(x)} = g(x|a) \left(u(\underline{w}) - u(\phi(x)) - \tilde{\lambda}(\theta - \rho(x))\right) \tag{A-2}$$

From Lemma 2.1, it is optimal to elicit $p(x) = 0$. Therefore, the Kuhn-Tucker conditions for the problem yield that: $\tilde{\lambda} \leq (u(\phi(x)) - u(\underline{w}))/(\rho(x) - \theta)$ for all x. Next, differentiating L_1 with respect to a yields the following condition:

$$-\psi'(a) + \int g_a(x|a)u(\phi(x))dx + \tilde{\lambda} \left(1 - \int g_a(x|a)\rho(x)dx\right) = 0 \tag{A-3}$$

I analyze now the problem of the principal. First, the principal can offer a contract such that Equation (A-1) does not bind. In this case the optimal contracting problem will be similar to the standard moral hazard problem in Holmström (1979). Second, the principal can offer a contract such that Equation (A-1) binds. The problem of the principal can be written as follows:

$$\max_{\rho(.),\phi(.)\geq\underline{w},a,\tilde{\lambda}} \int g(x|a)(x - \phi(x))dx$$

s.t.

$$a = \int \rho(x)g(x|a)dx \qquad (s_1) \tag{A-4}$$

$$\psi'(a) = \int g_a(x|a)u(\phi(x))dx + \tilde{\lambda} \left(1 - \int g_a(x|a)\rho(x)dx\right) \qquad (s_2) \tag{A-5}$$

$$b \geq \int g(x|a)u(\phi(x))dx - \psi(a) \qquad (s_3) \tag{A-6}$$

$$\tilde{\lambda} \leq \inf(u(\phi(x)) - u(\underline{w}))/(\rho(x) - \theta)$$

In this problem, the principal maximizes total revenue subject to the participation of the agent, feasibility of the hedging choices and incentive-compatibility. In addition, the multiplier $\tilde{\lambda}$ may depend on $\rho(x)$ and $\phi(x)$. The associated Lagrangian is denoted L_2.

Differentiating L_2 with respect to $\rho(x)$ when $x \in X'$,

$$\frac{\partial L_2}{\partial \rho(x)} = g(x|a)(-s_1 + s_2\tilde{\lambda}\frac{g_a(x|a)}{g(x|a)}) \tag{A-7}$$

This term is increasing in x and therefore $\rho(x)$ can be set equal to θ for $x \leq x_0$ and Equation (A-7) binds for $x > x_0$ (when either s_2 or $\tilde{\lambda}$ are zero $x_0 \in \{\theta, +\infty\}$). The statement follows readily.

Let me show that when $x_0 = \theta$, $\tilde{\lambda} = \psi'(a)$. Rewriting the incentive-compatibility:

$$\begin{aligned}
\psi'(a) &= \int g_a(x|a)u(\phi(x))dx + \tilde{\lambda}\left(1 - \int g_a(x|a)\rho(x)dx\right) \\
&= \tilde{\lambda} + \int g_a(x|a)u(\phi(x))dx - \tilde{\lambda}\int g_a(x|a)\rho(x)dx \\
&= \tilde{\lambda} + \int g_a(x|a)u(\phi(x))dx - \tilde{\lambda}\int g_a(x|a)\left(\min X + \frac{u(\phi(x)) - u(\underline{w})}{\tilde{\lambda}}\right)dx \\
&= \tilde{\lambda}
\end{aligned}$$

□

The optimal contract takes two forms. For low realizations of the unhedgeable signal, the contract does not use the hedgeable signal, i.e. the compensation does not depend on y. Then, the contract is given by the standard Equations which link the ratio of marginal utilities to the likelihood ratio (as in Holmström (1979)). For higher realizations of the unhedgeable signal, the principal uses the managed signal by setting a performance target $\rho(x)$ for y. That is, the agent is paid $w(x, y) = \phi(x)$ above the limited liability only when y is equal to $\rho(x)$.

In the model, using the managed signal may require to give a rent to the agent and thus is valuable only if the likelihood ratio is sufficiently large. Surprisingly, the agent is paid (utility-wise) a fixed proportion of the realized hedged performance $y - \theta$.[47] If the hedgeable is used for all possible realizations of x, I show that the performance-pay coefficient is equal to the marginal cost of effort.

Proposition A. 2. *If the utility of the agent becomes linear, conditional on* $y > \theta$, $w(x, y) = u(\underline{w}) + \tilde{\lambda}(y - \theta)$ *with probability one.*

Proof: Note that $u(\phi(x)) - u(\underline{w}) = \tilde{\lambda}(\rho(x) - \theta)$ can be rewritten: $u(w(x, y)) - u(\underline{w}) = \tilde{\lambda}(y - \theta)$ because by construction y is always equal to $\rho(x)$ and $\phi(x) = w(x, y)$. As $u(.)$ becomes linear, this expression can be written: $w(x, y) = u(\underline{w}) + \tilde{\lambda}(y - \theta)$, which is linear in y.□

[47] Note that if the principal receives y instead of x *and* the principal is risk-averse, $\rho(x)$ may be greater than θ for $x \leq x_0$. Second, if the principal cannot bound the maximum effort done by the agent, the two regions may be inverted (i.e., $u(\phi(x)) - u(\underline{w}) = \tilde{\lambda}(\rho(x) - \theta)$ occurs for $x \leq x_0$).

I present another linear contract as a (partial) solution to the problem. I give first an economic rationale for linearity as obtained here. By hedging, the agent can reduce the project's net-present value for a personal gain. In the model, the ratio $(u(\phi(x))-u(\theta))/(\rho(x)-\theta)$ captures the ratio at which project value is converted into personal utility. If it falls too low for some x, the agent will prefer to reduce effort and produce outcome $y=\theta$. If it is too high for any x, this will cause a compensation that is too high or a threshold that is too low to provide incentives efficiently. The solution to these two forces is to make this ratio constant which leads to a compensation that is essentially linear in y. In intuitive terms, linearity minimizes the cost of providing incentives not to speculate when speculation is "cheap."

This form of linearity is conceptually different from the linearity obtained using robustness arguments. The contract presented in the Corollary is only linear ex-post. That is, an outside observer running the regression of wages on the hedged signal (omitting the mass point at $y=\theta$) would observe a perfectly linear relationship. On the other hand, an observer investigating the shape of contracts offered by firms or the contract as a function of true performance x, would not necessarily obtain a linear contract.

Agent's Problem under Quadratic Cost

In this Appendix, I relax two aspects of the problem. First, under MDO, the cost of hedging is by assumption additively separable. Second, I show that under stronger assumptions on the cost function, the solution to the first-order condition in the problem of the agent is unique. Abusing on the previous notation, assume that $X=(x_1,\ldots,x_N)'$. I restrict the attention to only two possible efforts, $a\in\{0,\overline{a}\}$, and assume that $\overline{a}$ is sufficiently small so that eliciting $\overline{a}$ is optimal for the principal.

Conditional on a the probability of each outcome is $P(a)=(p_1(a),\ldots,p_N(a))'$, where $p_k(a)>0$ is the probability associated to outcome x_k and a is mean of the distribution. Denote $\hat{P}=(\hat{p}_1,\ldots,\hat{p}_N)$ the probability of each outcome after hedging has occurred. The compensation of the manager is written $W=(w_1,\ldots,w_N)'$. In vector notation, denote $\mathcal{U}=(u(w_1),\ldots,u(w_N))'$. Let θ be the total cost of hedging, defined as a function of a, $P(a)$ and $\hat{P}$. Assume that hedging is small so that θ can be approximated using the following Taylor expansion for $\hat{P}$ close to $P(a)$.

$$\theta(P(a),\hat{P})\approx\theta(P(a),P(a))+D\theta(\hat{P}-P(a))$$

$$+(\hat{P}-P(a))'D^2\theta(\hat{P}-P(a))/2 \qquad \text{(A-8)}$$

I make the following assumptions. First, there is no cost for not hedging, i.e. $\theta(P(a),P(a))=0$. Second, there is zero marginal cost for a small hedge, i.e. $D\theta=0$. Third, I assume that the hessian matrix $H=D^2\theta$ is definite positive.

$$\theta \approx (\hat{P} - P(a))' H (\hat{P} - P(a))/2 \tag{A-9}$$

Through this Section, I assume that the positivity constraint on $\hat{P}$ does not bind (i.e., the eigenvalues of H are large enough). Let $P'(a)$ denote entry-wise derivatives.

Denote $\Delta = (\delta_i)_{i=1}^n$ where $\Delta = \hat{P} - P(a)$. The Problem of the Manager can be stated as follows:

$$(A) \max_{\Delta \geq -P(a), a \in A} \Delta' \mathcal{U} + P(a)' \mathcal{U} - \psi(a)$$

s.t.

$$\mathbb{1}' \Delta = \quad 0 \qquad (\lambda) \tag{A-10}$$

$$\Delta' X \leq \quad -\Delta' H \Delta / 2 \qquad (\mu) \tag{A-11}$$

Differentiating with respect to Δ and rearranging: $\mu H \Delta = \mathcal{U} - \lambda - \mu X$. Pre-multiplying by Δ' and using Equation (A-11), $-2\mu \Delta' X = \Delta'(\mathcal{U} - \lambda - \mu X)$. Simplifying and substituting Δ yields: $(\mathcal{U} - \lambda + \mu X)' H^{-1} (\mathcal{U} - \lambda - \mu X) = 0$. If instead, one pre-multiplies by $\mathbb{1}' H^{-1}$ and use Equation (A-10), $\mathbb{1}' H^{-1} (\mathcal{U} - \lambda - \mu X) = 0$. One obtains a system of two Equations in two unknowns which yields the following second-order polynomial for μ,

$$\mu^2 ((\mathbb{1}' H^{-1} X)^2 + X' H^{-1} X \mathbb{1}' H^{-1} \mathbb{1}) + (\mathcal{U}' H^{-1} \mathcal{U} \mathbb{1}' H^{-1} \mathbb{1} - (\mathbb{1}' H^{-1} \mathcal{U})^2) = 0$$

This system has two real roots. One is negative and thus cannot be optimal and the other yields the following characterization:

$$\lambda = \frac{\mathbb{1}' H^{-1} \mathcal{U}}{\mathbb{1}' H^{-1} \mathbb{1}} - \frac{\mathbb{1}' H^{-1} X}{\mathbb{1}' H^{-1} \mathbb{1}} \sqrt{\frac{(\mathcal{U}' H^{-1} \mathcal{U} \mathbb{1}' H^{-1} \mathbb{1} - (\mathbb{1}' H^{-1} \mathcal{U})^2}{X' H^{-1} X \mathbb{1}' H^{-1} \mathbb{1} - (\mathbb{1}' H^{-1} X)^2}} \tag{A-12}$$

$$\mu = \sqrt{\frac{\mathcal{U}' H^{-1} \mathcal{U} \mathbb{1}' H^{-1} \mathbb{1} - (\mathbb{1}' H^{-1} \mathcal{U})^2}{X' H^{-1} X \mathbb{1}' H^{-1} \mathbb{1} - (\mathbb{1}' H^{-1} X)^2}} \tag{A-13}$$

$$\mu \Delta = H^{-1} (\mathcal{U} - \lambda - \mu X) \tag{A-14}$$

A simple application of the Cauchy-Schwarz inequality yields that μ is strictly positive if and only if $\mathcal{U}$ is not colinear to one, as in the previous case. The parameter μ has a simple geometric interpretation as the (scaled) angle between $\mathcal{U}$ (compensation) and $\mathbb{1}$ (constant contract). Intuitively the magnitude of the hedging choice is related to how much the compensation offered to the agent is congruent to the marginal payoff of the firm (i.e., \$1 in each state). Recall that the parameter μ represents how much the agent is willing to reduce the likelihood of high-output versus low-output outcomes for an equal wage.

Note that the following holds: $\mathbb{1}' \partial \lambda / \partial \mathcal{U} = 1$, $\mathbb{1}' \partial \mu / \partial \mathcal{U} = 0$ and $\mathcal{U}' \partial \mu / \partial \mathcal{U} > 0$. That is, shifting the compensation by adding a constant does not change the slope of the hedging threshold (as in the separable case). However, changing the utility received by the agent proportionately increases the slope of the threshold. The choice of hedging Δ and the multipliers are unique (and thus the first-order approach is valid).

Suppose $H = h\mathbb{1}\mathbb{1}' + D$ with D diagonal. Then, hedging is linear in utility. For all i,

$$\delta_i = (u(w_i)/\mu - \lambda/\mu - x_i)/D_{i,i} \tag{A-15}$$

Then, $\hat{p}_k \geq p_k(a)$ if and only if $u(w_k) \geq \lambda + \mu x_k$.

One can also verify that Δ is zero if and only if compensation is linear. The linear threshold featured earlier is recovered given a weaker restriction on cross-effects, that is, all off-diagonal terms must be the same. This assumption is a symmetry restriction on the effect of changing the likelihood of one event on the marginal cost of other outcomes.

Appendix B: Omitted Proofs

Proof of Lemma 2.1: The method for this proof is to construct a new contract that yields weakly more utility to both contracting parties and elicits perfect hedging. To do so, I verify that this new contract is desirable to the agent and does not generate deviations from the previous effort.

Let $\hat{F}$ (resp. a) be the distribution (resp. effort) chosen by the manager in response to a contract $w(.)$. I construct the compensation $\hat{w}(a) = \int w(y)d\hat{F}(y)$ and $\hat{w}(y) = \underline{w}$ for $y \neq a$. Let $\tilde{F}$ (resp. a) denote the hedging choice (resp. effort choice) of the agent in response to $\hat{w}(.)$.

Claim 1: the agent achieves weakly more utility under $\hat{w}(.)$. With $\hat{w}$, the agent may choose effort $\tilde{a} = a$ and set $y = a$. This generates an expected utility $u(\hat{w}(a)) - \psi(a)$. Then:

$$\begin{aligned} \int u(w(y))d\hat{F}(y) &\leq u(\int w(y)d\hat{F}(y)) \\ &\leq u(\int w(y)d\hat{F}(y)) \\ &\leq u(\hat{w}(a)) \end{aligned}$$

Claim 2: Under $\hat{w}$, the distribution $\tilde{F}$ must have its support included in $\{\theta, a\}$. Suppose not. Define an alternative hedging strategy G as follows: $\int_{y=a} dG(y) = \int (y-\theta)/(a-\theta)d\hat{F}(y)$ and $\int_{y=\theta} dG(y) = \int (a-y)/(a-\theta)d\hat{F}(y)$. It follows that:

$$\begin{aligned} \int y dG(y) &= a\int \frac{y-\theta}{a-\theta}d\tilde{F}(y) + \theta \int \frac{a-y}{a-\theta}d\tilde{F}(y) \\ &= \int y d\tilde{F}(y) \end{aligned}$$

It follows that G is feasible for the agent if $\tilde{F}$ is.

$$\begin{aligned}\int u(\hat{w}(y))d\tilde{F}(y) &= u(\hat{w}(a))\int_{y=a} d\tilde{F}(y) + u(\underline{w})(1-\int_{y=a} d\tilde{F}(y)) \\ &< u(\hat{w}(a))\int_{y=a} dG(y) + u(\underline{w})(1-\int_{y=a} dG(y)) \end{aligned} \tag{A-16}$$

Thus, the agent would be strictly better under G, a contradiction.

Claim 3: a must be incentive-compatible under $\hat{w}$. Suppose not. The agent must be choosing $\tilde{a} < a$ and $\tilde{F}$ with support $\{\theta, a\}$. Define the distribution G' as follows: for any $X' \subset X\backslash\{\theta\}$, $\int_{X'} dG'(y) = \int_{y=a} d\tilde{F}(y)\int_{X'} d\hat{F}(y)$ and $\int_{y=\theta} dG'(y) = \int_{y=\theta} d\tilde{F}(y) + \int_{y=a} d\tilde{F}(y)\int_{y=\theta} d\hat{F}(y)$. Suppose that the agent follows G' and $\tilde{a}$ instead of $\hat{F}$ and a with $w(.)$.

$$\begin{aligned}\int ydG'(y) &= \theta(\int_{y=\theta} d\tilde{F}(y) + \int_{y=a} d\tilde{F}(y)\int_{y=\theta} d\hat{F}(y)) + \int_{y=a} d\tilde{F}(y)\int_{y\neq\theta} yd\hat{F}(y) \\ &= \theta\int_{y=\theta} d\tilde{F}(y) + \int_{y=a} d\tilde{F}(y)\int yd\hat{F}(y) \\ &= \theta\int_{y=\theta} d\tilde{F}(y) + a\int_{y=a} d\tilde{F}(y) \\ &= \tilde{a}\end{aligned}$$

Therefore G' is feasible with effort $\tilde{a}$.

$$\begin{aligned}\int u(w(y))dG'(y) - \psi(\tilde{a}) &= (\int_{y=\theta} d\tilde{F}(y) + \int_{y=a} d\tilde{F}(y)\int_{y=\theta} d\hat{F}(y))u(w(\theta)) \\ &\quad + \int_{y=a} d\tilde{F}(y)\int_{y\neq\theta} u(w(y))d\hat{F}(y) \\ &= u(w(\theta))\int_{y=\theta} d\tilde{F}(y) + \int u(w(y))d\hat{F}(y)\int_{y=a} d\tilde{F}(y) - \psi(\tilde{a}) \\ &\geq \int u(w(y))d\hat{F}(y) - \psi(a) + \int u(w(y))d\hat{F}(y)\int_{y=a} d\tilde{F}(y) - u(w(a))\int_{y=a} d\tilde{F}(y) \\ &\geq \int u(w(y))d\hat{F}(y) - \psi(a) + \int_{y=a} d\tilde{F}(y)\int (u(w(y)) - u(w(a)))d\hat{F}(y) \\ &\geq \int u(w(y))d\hat{F}(y) - \psi(a)\end{aligned}$$

This is a contradiction to $(a, \hat{F})$ incentive-compatible under $w(.)$.

It follows that the principal achieves weakly more under $\hat{w}(.)$ than under $w(.)$. Note finally that the inequality obtained in claim 1 is strict when the agent is risk-averse. The previous claims remain true using $\hat{w}(y) = \int w(y)d\hat{F}(y) - \epsilon$ for ϵ small enough. However, this contract will strictly increase the utility of the principal.□

Proof of Proposition 2.1: Suppose the contract is not agent-unbounded and let $(\hat{F}^n, a^n)_{n=1}^{\infty}$ be a sequence of actions for the agent such that $\int yd\hat{F}^n(y) \leq a^n$ and $\int u(w(y))d\hat{F}^n(y) - \psi(a^n)$ converges to $\sup_{a,\hat{F}(.)\in\Gamma(a)} \int u(w(y))d\hat{F}(y) -$

$\psi(a) < +\infty$. I need to show that necessarily a_n must converge to $\underline{a}$.

For $n > 1$, let G^n be a sequence of distributions constructed as follows:

$$\int_{y=-n+\underline{a}} dG^n(y) = \frac{a^n - \underline{a}}{n + a^n - \underline{a}} + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int_{y=-n+\underline{a}} dF^n(y)$$

and for any $X' \subset X \backslash \{-n\}$,

$$\int_{X'} dG^n(y) = (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int_{y=-n} dF^n(y)$$

In intuitive terms, when hedging according to G^n, the manager samples between $-n - \overline{a}$ and $F^n(.)$.

First, I argue that $(\underline{a}, G^n)$ is feasible by the agent. To see this,

$$\begin{aligned} \int y dG^n(y) &= \frac{a^n - \underline{a}}{n + a^n - \underline{a}}(-n + \underline{a}) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int y dF^n(y) \\ &\leq \frac{a^n - \underline{a}}{n + a^n - \underline{a}}(-n + \underline{a}) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) a^n \\ &\leq \underline{a} \end{aligned}$$

Second, calculating the utility obtained by the agent on this sequence:

$$\begin{aligned} \int u(w(y)) dG^n(y) &= \frac{a^n - \underline{a}}{n + a^n - \underline{a}} u(w(-n - y)) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int u(w(y)) d\hat{F}^n(y) \\ &\geq \frac{a^n - \underline{a}}{n + a^n - \underline{a}} u(\underline{w}) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int u(w(y)) d\hat{F}^n(y) \end{aligned} \quad \text{(A-17)}$$

Taking the limit on n,

$$\lim \int u(w(y)) dG^n(y) = \lim \int u(w(y)) d\hat{F}^n(y)$$

And therefore, since F^n converges to an supremum of the problem, it must that $\psi(a^n)$ converges to zero, i.e. a^n converges to $\underline{a}$. Therefore, the contract is incentive-free.□

Proof of Lemma 2.2: Under the prescribed actions, the manager obtains $u(w) - \psi(a)$. Assume a deviation to $a - \epsilon \leq a$ and $\hat{F}$. The best possible distribution maximizes the probability that $y = a$ is attained which can only be achieved if $\hat{F}$ has support over a and θ. Let p denote the probability that $y = \theta$ and $1 - p$, the probability that $y = a$. Since hedging must have zero NPV, $p\theta + (1 - p)a = a - \epsilon$. Therefore: $p = \epsilon/(a - \theta)$. Reinjecting in the problem of the manager: $\frac{\epsilon}{a-\theta} u(\underline{w}) + (1 - \frac{\epsilon}{a-\theta}) u(w) - \psi(a - \epsilon)$. This problem is concave. For $\epsilon = 0$ optimal, the first-order condition in Equation (4) must prescribe $\epsilon \leq 0$.□

Proof of Proposition 2.2: I show first that $a^* \leq a^{**}$ if and only if first-best is incentive-compatible. Note that $a^* \leq a^{**}$ implies that $\psi'(a^*) \leq \psi'(a^{**})$, i.e.:

$$\psi'(a^*) \leq \frac{b + \psi(a^{**}) - u(\underline{w})}{a^{**} - \theta}$$

Define the function $\phi(a)$ as follows:

$$\phi(a) = \frac{b + \psi(a) - u(\underline{w})}{a - \theta}$$

Differentiating with respect to a,

$$\phi'(a) = \frac{u(\underline{w}) - b - \psi(a) + (a - \theta)\psi(a)}{a - \theta}$$

This expression is negative under the regularity condition assumed earlier and thus:

$$\psi'(a^*) \leq \frac{b + \psi(a^*) - u(\underline{w})}{a^* - \theta}$$

And thus (W^*, a^*) is incentive-compatible. The case with $a^* > a^{**}$ is analogous.

For the final part of the statement, I show that the elicited effort is above $\underline{a}$. Note that the contract $(u^{-1}(b), \underline{a})$ implies that the incentive-compatibility and the reservation binds, so it is sufficient to check that $a^{**} > \underline{a}$. Plugging $a^{**} = \underline{a}$ into Equation (5) ensures that this is indeed the case.□

Proof of Proposition 2.3: Note first that $W^{***} \geq \underline{w}$, so that it is only necessary to verify that $a^{***} \geq \underline{a}$. To do so, it is sufficient to plug $a = \underline{a}$ into the first-order condition corresponding to the program $a - u^{-1}(\psi'(a)(a - \theta) + u(\underline{w}))$. This yields the following expression: $1 - \frac{\psi''(\underline{a})(\underline{a}-\theta)}{u'(\underline{w})}$ which, simplified, yields Equation (7). When Equation (7) is true (resp. false), this term is positive (resp. negative), and thus $a^{***} > \underline{a}$ (resp. $a = \underline{a}$).□

Proof of Proposition 2.4: The first part of the argument is similar to the proof of Lemma 2.1 (and not repeated here). The contract $u(w(y))$ must be (weakly) concave, or else an agent could achieve the same utility as the concavification of $u(w(y))$ by taking gambles but, because the agent is risk-averse, this would be more costly to the principal than offering the concavification directly. In formal terms, for any non-concave function, there exists a concave function (its concavification) that does strictly better for the principal.

To simplify notations, let $u(\underline{w})$ be normalized to zero and omit the j exponent on the utility of the agent. To show that the optimal compensation is linear, I rewrite first Equation (6) as a function of W and a but not c. First, incentive-compatibility implies that:

$$u(W) = c\psi'(a)(a - \theta)$$

Second, one may substitute this expression in Equation (6):

$$\frac{u'(W)}{u(W)} = \frac{1}{a-\theta} + \frac{\psi''(a)}{\psi'(a)}$$

Denoting $\eta(W) = u'(W)/u(W)$,

$$W = \eta^{-1}(\frac{1}{a-\theta} + \frac{\psi''(a)}{\psi'(a)})$$

By simple differentiation, it is easily verified that, for two smooth functions f and g, $f(g(z))$ is convex if $f(.)$ is decreasing and convex, and $g(.)$ is concave. It follows that when $u(.)$ becomes linear (so that $\rho(.)$ becomes decreasing and convex), W becomes convex in a. Note finally that as c varies on $\mathbb{R}\backslash\{0\}$, a must vary on $[\underline{a}, \overline{a})$ so that the plot $W(a)$ is convex in a.

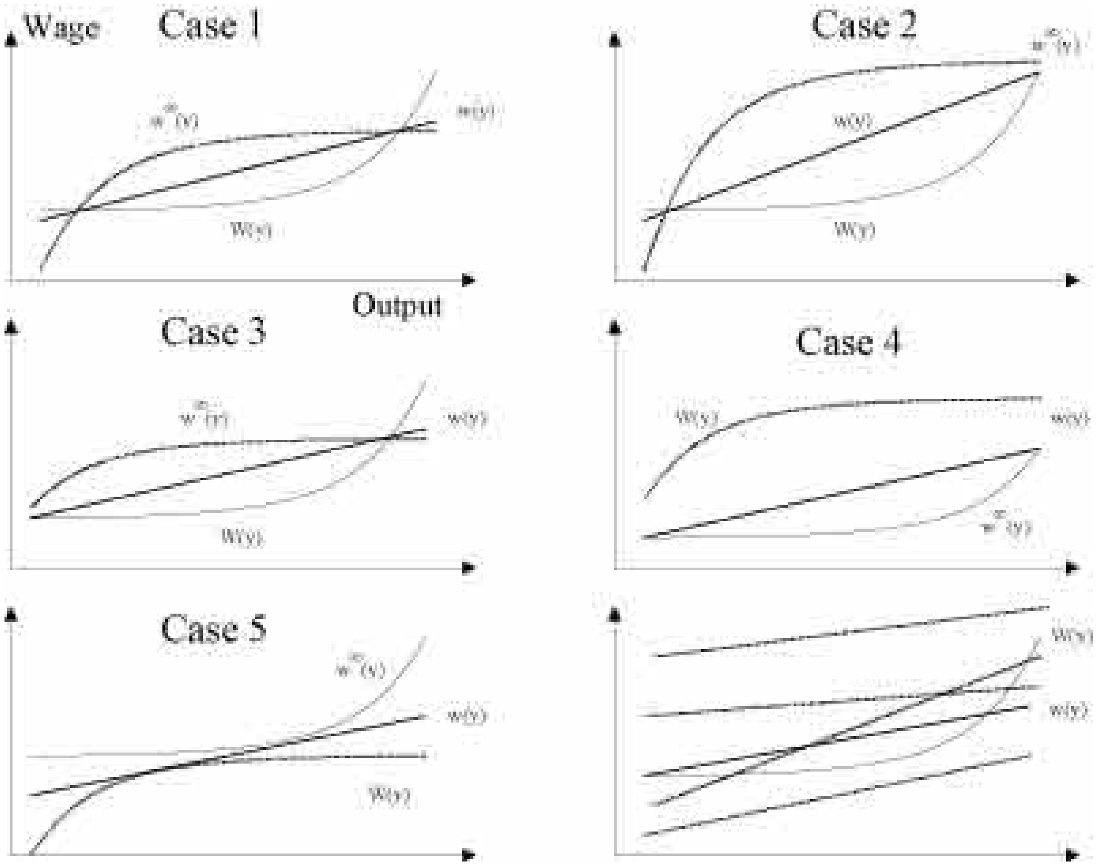

Figure 6. Linear Contract - Cases

As j becomes large the optimal contract associated to u^j must converge on $[\underline{a}, \overline{a}]$. Suppose the limit is not linear almost everywhere and denote $w^\infty(.)$ the limiting contract. I consider next several cases which are represented in Figure 6. In each of these cases, it should be noted that any compensation schedule $w(.)$ that is still weakly concave but is everywhere closer to $W(a)$ than $w^\infty(.)$ is (in the sense that $W(a) - w^\infty(a)$ and $W(a) - w(a)$ have the same sign but $|W(a) - w^\infty(a)| > |W(a) - w(a)|$ for all a) will be preferred by the principal. This is because a^{***} is the solution to a concave program for any c.

Case 1: Suppose $w^\infty(.)$ intersects $W(.)$ twice, at (W_1, a_1) and (W_2, a_2). Then, the linear compensation $w(y) = \frac{W_2 - W_1}{a_2 - a_1}(y - a_2) + W_2$ (plotted in the upper-left-hand side) will do strictly better than $w^\infty(.)$, a contradiction. The next cases will be similar to this case although with different definitions of a_i and W_i $(i = 1, 2)$.

Case 2: Suppose $w^\infty(.)$ intersects $W(.)$ once, at (W_1, a_1) and then $w^\infty(y) \geq W(y)$ for $y \geq a_1$. Then, the previous argument holds but defining $a_2 = \overline{a}$ and $W_2 = w^\infty(\overline{a})$.

Case 3: Suppose $w^\infty(.)$ intersects $W(.)$ once, at (W_1, a_1) and then $w^\infty(y) \leq W(y)$ for $y \geq a_1$. Then, the previous argument holds but defining $a_2 = \underline{a}$ and $W_2 = w^\infty(\underline{a})$.

Case 4: Suppose $w^\infty(.)$ is always greater than $W(.)$. Then, one may apply the same argument by defining $a_1 = \underline{a}$ and $W_1 = w^\infty(\underline{a})$, and $a_2 = \overline{a}$ and $W_2 = w^\infty(\overline{a})$.

Case 5: Suppose $w^\infty(.)$ is always smaller than $W(.)$. Define $V_1 = \{(y, \tilde{w})/\tilde{w} \geq w^\infty(y)\}$ and $V_2 = \{(y, \tilde{w})/\tilde{w} \leq W(y)\}$. V_1 and V_2 are convex sets which may intersect only on their boundary, therefore by the separating hyperplane theorem, there exists a separating hyperplane (in this case a compensation) $w(.)$ such that $w^\infty(y) \leq w(y) \leq W(y)$. This compensation is strictly preferred to w^∞.

Corollary 2.2 follows by the same argument (the proof is the same as the five cases considered earlier), so that the linear function must be of the form exhibited in the lower-right hand side of Figure 6 (in bold is plotted the linear function that is most preferred versus in dotted ones that are not as preferred by the principal). To obtain Corollary 2.3, let $H(.)$ converge to a mass point. By continuity, the solution of the model must converge to the optimum of the previous problem. However, since the solution implies $w(y) = h_0 + h_1 y$ linear.□

Proof of Proposition 3.1: Claim 1: Problem (A) has a solution. The result is obvious when N is finite. Suppose N is not finite. Let $((\delta_k^n)_{k=1}^\infty, a^n)$ be a sequence of feasible actions converging to an optimum. Define $a^\infty = \liminf a^n$ and, for all k, $\delta_k^\infty = \liminf \delta_k^n$. I shall show that S is feasible for the agent.

$$\begin{aligned}
\sum_{n=1}^{\infty}(\delta_k^\infty x_k + C(a^\infty, x_k, \delta_k^\infty)) &= \sum_{n=1}^{\infty}(\liminf \delta_k^n x_k + C(\liminf a^k, x_k, \liminf \delta_k^n)) \\
&= \sum_{n=1}^{\infty}(\liminf(\delta_k^n x_k + C(a^k, x_k, \delta_k^n))) \\
&\leq \liminf \sum_{n=1}^{\infty}(\delta_k^n x_k + C(a^k, x_k, \delta_k^n)) \qquad \text{(Fatou's Lemma)} \\
&\leq 0
\end{aligned}$$

Then, $(\delta_k^\infty)_{k=1}^\infty$ satisfies Equation (13). For any k, one may define $I_k \subset [1, \infty]$, the set of indices such that $(\delta_{k'}^n)_{n \in I_k}$ converges to $\delta_{k'}^\infty$ for all $k' \leq k$.

$$\sum_{n=1}^{\infty} \delta_k^\infty = \sum_{n=1}^{n_0} \delta_k^\infty + \sum_{n=n_0+1}^{\infty} \delta_k^\infty$$

For any $\epsilon > 0$ there exists n_0 such that:

$$\sum_{n=1}^{\infty} \delta_k^{\infty} \leq \sum_{n=1}^{n_0} \delta_k^{\infty} + \epsilon/2$$

Then, there exists k such that for all $k' \in [k, \infty) \cap I_{n_0}$, $|\delta_k^{\infty} - \delta_k^n| < \epsilon/(2n_0)$. It follows then that:

$$\sum_{n=1}^{\infty} \delta_k^{\infty} \leq \epsilon/2 + \epsilon/2 = \epsilon$$

Then, $(\delta_k^{\infty})_{k=1}^{\infty}$ satisfies Equation (12). Thus the action S is feasible.

Finally, one needs to show that S is utility-maximizing for the agent. To see this, note that $\delta_k^{\infty} x_k + (1-\delta_k^{\infty})\theta \leq \overline{a}$, and therefore: $\delta_k^n u(w(x_k)) \leq (\overline{a}-\theta)/(x_k-\theta)u(w(x_k))$. Note that the the function $g(x_k) = |u(w(x_k))/(x_k - \theta)|$ dominates $\delta_k^n u(w(x_k))$. And thus one may apply the dominated convergence theorem to obtain:

$$\liminf \sum_{k=1}^{\infty} \delta_k^n u(w(x_k)) = \sum_{k=1}^{\infty} \delta_k^{\infty} u(w(x_k))$$

Claim 2: $\mu > 0$ Suppose not. The first-order condition in δ_k yields that $u(w(x_k)) = \lambda$ which is a contradiction to $\sum_{k=1}^{N} u(w(x_k))p_k'(0) > 0$. Equations (14), (15) and (16) are the first-order conditions of the problem.□

Proof of Corollary 3.2: By Equation (14),

$$\sum_{k=1}^{n} u(w(x_k)) - \lambda - \mu \sum_{k=1}^{n} x_k = 0$$

By Equation (15),

$$\frac{1}{c(a)} \sum_{k=1}^{n} (\frac{u(w(x_k)) - \lambda}{\mu} - x_k)x_k = -\sum_{k=1}^{n} \frac{c(a)}{2} \frac{1}{c(a)^2} (\frac{u(w(x_k)) - \lambda}{\mu} - x_k)^2$$

$$\begin{aligned} 0 &= \sum_{k=1}^{n} (\frac{u(w(x_k)) - \lambda}{\mu} - x_k)(\frac{u(w(x_k)) - \lambda}{\mu} + x_k) \\ &= \sum_{k=1}^{n} ((\frac{u(w(x_k)) - \lambda}{\mu})^2 - x_k^2) \\ &= \sum_{k=1}^{n} u(w(x_k))^2 + \lambda^2 N - 2\lambda\mu \sum_{k=1}^{n} u(w(x_k)) - \sum_{k=1}^{n} x_k^2 \end{aligned}$$

Solving for μ yields the following polynomial:

$$\mu^2((\sum_{k=1}^{N} x_k/N)^2 + \sum_{k=1}^{N} x_k^2/N) + (\sum_{k=1}^{N} u(w(x_k))^2/N - (\sum_{k=1}^{N} u(w(x_k))/N)^2 = 0$$

This Equation has a unique positive real root.

$$\begin{aligned}\mu &= \sqrt{\frac{\sum_{k=1}^N u(w(x_k))^2/N - (\sum_{k=1}^N u(w(x_k))/N)^2}{\sum_{k=1}^N x_k^2/N - (\sum_{k=1}^N x_k/N)^2}} \\ &= \frac{\sigma(u(w(y))/p(y))}{\sigma(y/p(y))}\end{aligned}$$

Reinjecting yields the expression for λ.$\square$

Proof of Proposition 4.1: (i) If $\underline{a}$ is elicited and the principal is risk-neutral, a constant contract is optimal among the unrestricted class of contracts. Therefore, it is also optimal among the following subclass of linear contracts, i.e. $w(y) = u^{-1}(\lambda + \mu y)$ with $\mu \geq 0$. Note first that since the contract is linear, the agent does not hedge. The agent maximizes:

$$\sum_{k=1}^N p_k(a)u(w(x_k)) - \psi(a) = \mu a - \lambda - \psi(a)$$

Thus, $\psi'(a) = \mu$. Note also that if (λ, μ) is optimal, it must be optimal to bind the participation of the agent, that is $\lambda = b - \psi'(a)a$. The principal maximizes the following objective:

$$\max_a \sum_{k=1}^N p_k(a)(x_k - u^{-1}(\psi'(a)(x_k - a) + b))$$

The first-order condition for this problem is:

$$\sum_{k=1}^N p_k'(a)(x_k - w(x_k)) - \psi''(a)\sum_{k=1}^N p_k(a)\frac{1}{u'(w(x_k))}(x_k - a)$$

If $a = \underline{a}$ is optimal, it is optimal to set $\mu = 0$. The first term in the above Equation is strictly positive by first-order stochastic dominance. Now note that:

$$\begin{aligned}\sum_{k=1}^N p_k(a)(x_k - a) &= \sum_{x_k < \underline{a}} p_k(a)(x_k - a) + \sum_{x_k \geq \underline{a}} p_k(a)(x_k - a) \\ &\leq \sum_{x_k < \underline{a}} p_k(a)(x_k - a) + \sum_{x_k \geq \underline{a}} p_k(a)(x_k - a) \\ &\leq 0\end{aligned}$$

It follows that $M_a(\underline{a}) > 0$.$\square$

Proof of Lemma 4.1: By Equation (17), $\mathfrak{u}(x_k) = \lambda + \mu x_k + \mu C_\delta(a, x_k, \delta_k)$. Therefore the agent achieves: $U = \sum_{k=1}^N (p_k(a) + \delta_k)(\lambda + \mu x_k + \mu C_\delta(a, x_k, \delta_k)) - \psi(a)$. Simplifying this Equation yields the left-hand side of Equation (22). I argue then that the participation is binding. First, one can replace $\mathfrak{u}(x_k)$ in Equation (16) and eliminate λ since $\sum_{k=1}^N p_k'(a) = 0$. Just like in the standard model any fixed change in the level of compensation does not affect incentives to work diligently or hedge. Second, it follows that by reducing λ the principal can reduce expected payments without affecting incentives. Thus, the participation of the manager must bind.□

Proof of Proposition 4.3: Note that δ_k must converge to zero, and therefore since $p_k(a)$ also converges to zero, $(p_k(a) + \delta_k)v'(x_k - w(x_k))/u'(w(x_k))$ must go to zero unless $v'(x_k - w(x_k))/u'(w(x_k))$ is unbounded. Case 1: $u'(w(x_k))$ is not bounded away from zero. This implies that there is a subsequence with indices I such that $\lim u(w(x_{k'})) = +\infty$ with $k' \in I$. Therefore v' must be bounded away from zero as k' becomes large. But, by Equation (30), v must then be large, which is a contradiction. Case 2: $v'(x_k - w_k)$ becoming large presents a similar contradiction. Taking the limit over Equation (30) using (i)-(ii) yields the desired result.□

Proof of Proposition 4.4: Under costless hedging, the principal can achieve first-best: $\hat{F}$ assigns probability one to a^*. As j becomes large, the contract under costly hedging can generate a surplus that is arbitrarily close to first-best. But, it must then hold that $\lim_{k\to+\infty} \int C(a, x_k, \delta_k)/j = 0$. This implies that $C(a, x_k, \delta_k)/j$ goes to zero for all x_k. But then $C_\delta(a, x_k, \delta_k)/j$ also converges to zero. And by Equation (17), it must then be that $u(w(x_k))$ converges to $\lambda + \mu x_k$. The optimal contract follows as the (unique) linear contract solution to first-best.□

Proof of Corollary 4.2: Each part of the statement is proved separately. To ease notations, the function $C(a, x, \delta)$ (and its derivatives) are denoted C, omitting the variables. Since the derivations can be long, algebraic steps are executed in the companion Mathematica notebook.

(i) For $\partial\delta/\partial x$, differentiating Equations (32) and (33) in x,

$$\begin{aligned}
\frac{\partial w}{\partial x} &= (\mu + \mu C_{\delta,\delta}\frac{\partial \delta}{\partial x} + \mu C_{\delta,x})/u'(w) \\
\frac{\partial w}{\partial x} &= (1-\beta) - \beta\frac{\partial \delta}{\partial x} C_{delta,\delta} - \beta C_{\delta,x} - \mu\gamma C_{\delta,a,x} - \mu\gamma C_{\delta,\delta,a}\frac{\partial \delta}{\partial x} - \mu(C_{\delta,\delta,x} + C_{\delta,\delta,\delta}\frac{\partial \delta}{\partial x})((d+p)(1/u'(w) - \tau) - \gamma pLR) \\
&\quad -\mu C_{\delta,\delta}(\frac{\partial \delta}{\partial x}(1/u'(w) - \tau) + (\delta+p)(-\frac{\partial w}{\partial x}u''(w)/u'(w)^2)) \qquad \text{(A-18)}
\end{aligned}$$

Solving these Equations in $\partial\delta/\partial x$:

$$\frac{\partial \delta}{\partial x} = \frac{(-1+\beta+\beta C_{\delta,x})u'(w)^3 + \mu u'(w)^2(1 + C_{\delta,x} + C_{\delta,a,x}\gamma u'(w) + C_{\delta,\delta,x}(\delta + p - (\gamma pLR + (\delta+p)\tau)u'(w))) - C_{\delta,\delta}(1 + C_{\delta,x})\mu^2(\delta+p)u''(w)}{u'(w)^2(\mu(-C_{\delta,\delta,\delta}(\delta+p) - C_{\delta,\delta,a}\gamma u'(w) + C_{\delta,\delta,\delta}(\gamma pLR + (\delta+p)\tau)u'(w)) + C_{\delta,\delta,}(-\beta u'(w) + \mu(-2 + \tau u'(w)))) + C^2_{\delta,\delta}\mu^2(\delta+p)u''(w)} \qquad \text{(A-19)}$$

Setting $u'(w) = 0$,

$$\frac{\partial \delta}{\partial x} = -\frac{1 + C_{\delta,x}}{C_{\delta,\delta}} \qquad \text{(A-20)}$$

This term is negative under the conditions of Corollary 4.2.

(ii) For $\partial\delta/\partial LR$, differentiating Equations (32) and (33) in p,

$$\frac{\partial w}{\partial LR} = (\mu C_{\delta,\delta}\frac{\partial \delta}{\partial p})/u'(w) \tag{A-21}$$

$$\frac{\partial w}{\partial LR} = -b\frac{\partial \delta}{\partial LR}C_{\delta,\delta} - \mu\gamma C_{\delta,\delta,a}\frac{\partial \delta}{\partial LR} - \mu(C_{\delta,\delta,\delta}\frac{\partial \delta}{\partial LR})((\delta+p)(1/u'(w)-\tau)-\gamma pLR)$$
$$-\mu C_{\delta,\delta}(\frac{\partial \delta}{\partial LR}(1/u'(w)-\tau)-\gamma p+(\delta+p)(-\frac{\partial w}{\partial LR}u''(w)/u'(w)^2)) \tag{A-22}$$

Solving these Equations in $\partial\delta/\partial LR$:

$$\frac{\partial w}{\partial p} = \frac{C_{\delta,\delta}\gamma\mu p u'(w)^3}{(\mu(-C_{\delta,\delta}d(\delta+p) - C_{\delta,\delta}a\gamma u'(w) + C_{\delta,\delta,\delta}(\gamma pLR + (\delta+p)\tau)u'(w)) + C_{\delta,\delta}(-\beta u'(w) + m(-2+\tau u'(w))))u'(w)^2 + C_{\delta,\delta}^2\mu^2(\delta+p)u''(w)} \tag{A-23}$$

This term is positive under the conditions of Corollary 4.2.

(iii) For $\partial\delta/\partial p$, differentiating Equations (32) and (33) in LR,

$$\frac{\partial w}{\partial p} = (\mu C_{\delta,\delta}\frac{\partial \delta}{\partial p})/u'(w) \tag{A-24}$$

$$\frac{\partial w}{\partial p} = -\beta\frac{\partial w}{\partial p}C_{\delta,\delta} - \mu\gamma C\delta,\delta,a\frac{\partial \delta}{\partial p} - \mu(C_{\delta,\delta,\delta}\frac{\partial \delta}{\partial p})((\delta+p)(1/u'(w)-\tau)-\gamma pLR)$$
$$-\mu C_{\delta,\delta}(\frac{\partial \delta}{\partial p}(1/u'(w)-\tau)-\gamma p+(\delta+p)(-\frac{\partial w}{\partial p}u''(w)/u'(w)^2)) \tag{A-25}$$

Solving these Equations in $\partial\delta/\partial p$:

$$\frac{\partial w}{\partial p} = \frac{C_{\delta,\delta}\mu(-1+\gamma LRu'(w)+\tau u'(w))u'(w)^2}{(\mu(-C_{\delta,\delta}d(\delta+p) - C_{\delta,\delta}a\gamma u'(w) + C_{\delta,\delta,\delta}(\gamma pLR + (\delta+p)\tau)u'(w)) + C_{\delta,\delta}(-\beta u'(w) + m(-2+\tau u'(w))))u'(w)^2 + C_{\delta,\delta}^2\mu^2(\delta+p)u''(w)} \tag{A-26}$$

This term is positive or negative under the conditions of Corollary 4.2.□

Proof of Proposition 5.1: The first-order condition with respect to $w(y)$ yields that:

$$\frac{v'(y-w(y))}{u'(w(y))} = \tilde{\tau} \tag{A-27}$$

This is the standard Arrow-Borch condition for efficient risk-sharing and implies that v and u must be increasing. The first-order condition with respect to $\delta(y)$ yields that:

$$v(y-w(y)) - \tilde{\alpha} - \tilde{\beta}(y + C'(\delta(y))) + \tilde{\tau}u(w(y)) = 0 \tag{A-28}$$

Evaluating at $\delta(y) = 0$ for all y,

$$v(y - w(y)) - \tilde{\alpha} - \tilde{\beta} y + \tilde{\tau} u(w(y)) = 0$$

Differentiating this expression with respect to y,

$$(1 - w'(y))v'(y - w(y)) - \tilde{\beta} + \tilde{\tau} w'(y) u'(w(y)) = 0$$

Rearranging this expression:

$$\begin{aligned} 0 &= \frac{1}{w'(y)} \frac{u'(w(y))}{v'(y - w(y))} - \frac{u'(w(y))}{v'(y - w(y))} + \tilde{\tau} - \frac{\tilde{\beta}}{w'(y) u'(w(y))} \\ &= \frac{\tilde{\tau}}{w'(y)} - \tilde{\tau} - \frac{\tilde{\beta}}{w'(y) u'(w(y))} + \tilde{\tau} \\ &= \tilde{\tau} - \frac{\tilde{\beta}}{u'(w(y))} \end{aligned}$$

This implies that $\tilde{\beta} = \tilde{\tau} = 0$, a contradiction to $v'(y - w(y)) > 0$.□

Proof of Proposition 5.2: The first-order condition with respect to μ,

$$\int (\delta(y) + f(y|a)) \frac{v'(y - w(y))}{u'(w(y))} (-y - C'(\delta(y)))dy + \gamma + \gamma \int f_a(y|a) C'(\delta(y))dy$$

$$\tau a + \tau \int ((f(y|a) + \delta(y))C'(\delta(y)) - C(a, y, \delta(y)))dy = 0$$

Under perfect risk-sharing, τ is equal to $v'(y - w(y))/u'(w(y))$ for all y. Observe first that the agent will always select $a > \underline{a}$ with perfect risk-sharing. Then one may simplify the above expression as follows:

$$\gamma \psi'(a) = \tau(\lambda - b - \psi(a)) + \mu \tau = 0$$

Therefore $\gamma = 0$. Suppose that $\int C'(\delta(y)) f_a(y|a)dy = -1$. Then, the first-order with respect to $\delta(y)$ in the problem of the agent would yield:

$$\begin{aligned} u(w(y)) &= \lambda + \mu y + \mu C'(\delta(y)) \\ \int f_a(y|a) u(w(y))dy &= \mu - \mu = 0 \end{aligned}$$

This would imply that the first-order with respect to effort in the problem of the agent select $a = \underline{a}$, a contradiction to $a > \underline{a}$. If $\int C'(\delta(y)) f_a(y|a)dy \neq -1$, it must be that $\beta = 0$. But $\gamma = \beta = 0$ implies in Equation (30) that $v(y - w(y))$ is constant, a contradiction to perfect risk-sharing.□

Proof of Proposition 5.3: If no hedging is optimal, $v(y-w(y))=\alpha+\beta y$. Differentiating:

$$v'(y-w(y))-h_1\frac{v'(y-w(y))}{u'(w(y))}=\beta$$

This ytields the first part of the result. Differentiating again,

$$v''(y-w(y))(1-h_1/u'(w(y)))-\mu\frac{\partial v'(y-w(y))/u'(w(y))}{\partial y}=0$$

This yields the first part of the result. As y is large, $u'(w(y))$ converges to zero and thus: $\frac{\partial v'(y-w(y))/u'(w(y))}{\partial y}\geq 0.$□

Proof of Proposition 5.4: Develop the covariance as follows:

$$\begin{aligned}
cov(\underline{y}^{t'},\overline{y}^{t'}) &= \mathbb{E}(\underline{y}^{t'}\overline{y}^{t'})-\mathbb{E}(\underline{y}^{t'})\mathbb{E}(\overline{y}^{t'})\\
&= \int \mathrm{y}\hat{f}(\mathrm{y})d\mathrm{y}-\mathbb{E}(\underline{y}^{t'})\mathbb{E}(\overline{y}^{t'})\\
&= cov(\underline{x}^{t'},\overline{x}^{t'})-(\mathbb{E}(\underline{x}^{t'})\mathbb{E}(\overline{x}^{t'})-\mathbb{E}(\underline{y}^{t'})\mathbb{E}(\overline{y}^{t'}))+R
\end{aligned}$$

But the analogue of Constraint 15 in the multivariate case implies that:

$$R=-\int C^k(\mathcal{C}^k[\frac{u(w(\mathrm{y}))-\lambda-\mu\mathrm{y}}{\mu}])d\mathrm{y}$$

As proved earlier, no hedging cannot be optimal and thus $R\neq 0$. This concludes the argument.□

Bibliography

Arya, Anil, Jonathan Glover, and Shyam Sunder, "Are Unmanaged Earnings Always Better for Shareholders?," 2003. Yale ICF Working Paper No. 02-37.

Berk, Jonathan and Richard C. Green, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 2004, *112* (6), 1269–1294.

Biais, Bruno and Catherine Casamatta, "Optimal Leverage and Aggregate Investment," *Journal of Finance*, 1999, *54* (4), 1291–1323.

Chang, Chun, "Does Hedging Aggravate or Alleviate Agency Problems? A Managerial Theory of Risk Management," 1997. Carlson School of Management Working Paper, University of Minnesota.

Crocker, Keith and Steven Huddart, "An Optimal Contracting Approach to Earnings Management," 2006. Working Paper.

Debreu, Gerard, *Theory of Value : An Axiomatic Analysis of Economic Equilibrium*, Yale University Press, 1972.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney, "Detecting Earnings Management," *Accounting Review*, 1995, *70* (2), 193–225.

Degeorge, Francois, Jayendu Patel, and Richard Zeckhauser, "Earnings Management to Exceed Thresholds," *Journal of Business*, 1999, *72* (1), 1–33.

Diamond, Peter, "Managerial Incentives: On the Near Linearity of Optimal Compensation," *Journal of Political Economy*, 1998, *106* (5), 931–957.

Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein, "Risk Management: Coordinating Corporate Investment and Financing Policies," *Journal of Finance*, 1993, *48* (5), 1629–1658.

Glover, Jonathan, Yuji Ijiri, Carolyn Levine, and Pierre Liang, "Verifiability and Manipulability vs. Robustness to Information Asymmetries about Verifiability and Manipulability," 2007. Carnegie Mellon Working Paper.

Holmström, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

__ **and Paul R. Milgrom**, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 1987, *55* (2), 303–328.

Huang, Chi-Fu and Robert H. Litzenberger, *Foundations for Financial Economics*, North Holland, 1988.

Innes, Joseph E., "Limited Liability and Incentive Contracting with Ex-Ante Action Choices," *Journal of Economic Theory*, 1990, *52*, 45–68.

Liang, Pierre and Lin Nan, "Managerial Attention, Endogenous Performance Variance, and Incentive Contracts," 2006. Preliminary Draft.

Lowenstein, Roger, *When Genius Failed: The Rise and Fall of Long-Term Capital Management*, Fourth Estate, 2002.

Mian, Shehzad L., "Evidence on Corporate Hedging Policy," *Journal of Financial and Quantitative Analysis*, 1996, *31* (3), 419–439.

Morellec, Erwan and Clifford W. Smith, "Performance thresholds in Managerial Incentive Contracts," *Review of Finance*, 2007, *11* (1), 1–23.

Murphy, Kevin J., *Executive Compensation*, Vol. 3 of Handbook of labor economics, North Holland, 1999. Ashenfelter, Orley and Card, David.

Pagan, Adrian and Aman Ullah, *Nonparametric Econometrics* Themes in Modern Econometrics, Cambridge, UK: Cambridge University Press, 1999.

Palomino, Frederic and Andrea Prat, "Risk Taking and Optimal Contracts for Money Managers," *Rand Journal of Economics*, 2003, *34* (1), 113–137.

Parlour, Christine, Amiyatosh Purnanandam, and Uday Rajan, "Risk Control and Compensation," 2006. Preliminary Draft.

Povel, Paul and Michael Raith, "Optimal Debt with Unobservable Investments," *Rand*

Journal of Economics, 2004, *35* (3), 599–616.

Rogerson, William P., "Repeated Moral Hazard," *Econometrica*, 1985, *53* (1), 69–76.

Ross, Steven A., "Compensation, Incentives, and the Duality of Risk Aversion and Riskiness," *Journal of Finance*, 2004, *59* (1), 207–225.

Smith, Clifford W. and Rene M. Stulz, "The Determinants of Firms' Hedging Policies," *Journal of Financial and Quantitative Analysis*, 1985, *20* (4), 391–405.