Risk Management, Executive Compensation and the Cross-Section of Corporate Earnings*

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, risk management can reduce extraneous noise in the variables outsiders observe or create opportunities for self-dealing behavior. Solving this trade-off, the analysis characterizes which events should be hedged and how executive compensation can give the right incentives to do so. The model rationalizes several empirical features commonly described as anomalous, in an optimal-contract setting. In equilibrium, the distribution of hedged earnings is hump-shaped with asymmetric tails even though the production technology does not exhibit such features. Further, the model can predict an S-shaped response of the stock price to current earnings. The model accounts for the prevalence of linear compensation schemes and the relatively low performance-pay observed in managerial jobs. Empirical implications for risk controls and the detection of earnings management are also examined.

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

---

*This paper is based on Chapter I of my dissertation. I am indebted to my dissertation advisors Steve Spear, Pierre Liang, Rick Green, Chris Sleet and Jon Glover for their comments along several iterations of this manuscript. Many thanks to Fernando Anjos, Vincent Glode, Burton Hollifield, Aziz Lookman, Spencer Martin, Bryan Routledge, Duane Seppi, Sevin Yeltekin and participants to the FEAST Workshop, Tepper finance seminar and Stony Brook Game Theory Conference. Several examples were borrowed from final projects in the Tepper undergraduate Corporate Finance class 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: February 1st 2008.
Most companies and institutions engage in 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, while Southwest Airlines hedges most of its fuel expenses, American Airlines and Continental Airlines do not.¹

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 corporate risk management.² Quite strikingly, the first of these objectives is to “Discourage transactions and transaction structures primarily motivated by accounting and reporting concerns rather than economics.”³ A widespread idea is to formulate the basic trade-off in the following terms: risk management 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.”

¹Southwest 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.

²The 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.

³See 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 Issuers” (June 2005).
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.”

One common denominator in current debates is that discussions 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 concerns.” Several key questions are of interest in order to better understand and regulate risk management.

I. What informational frictions cause a need for corporate risk management in firms held by well-diversified investors?

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 (downside risk) 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? 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? How can risk management be detected?

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 events (II.). These conditions are derived from explicit informational frictions and not from exogenously specified capital market frictions. Then, I show how risk management 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.).

---

4The approach of this paper is with hidden actions by the manager. Other authors (e.g., Demarzo and Duffie (1991) and Demarzo and Duffie (1995)) study models in which hedging helps the principal learn the ability of the agent and focus on issues different from those discussed here.
To further motivate some new insights provided by the model, I describe next three stylized facts commonly viewed as puzzling or unexplained from the perspective of standard theory.

Reviewing compensation practices across different industries, Murphy (1999) reports that most of the pay received by executives is driven by equity compensation. Bebchuk and Grinstein (2005) explain that the share of equity-based compensation (options and restricted stock) represented 59% of the total compensation of SP500 CEOs in 2003, peaking at 78% in 2000. Recently issued options are convex (because of their time-value) in realized performance and become linear for large performance.⁵ Quantitatively, convexity is hard to reproduce in standard agency problems with risk-aversion (see Haubrich and Popova (1998) and Figure 3.3 in this paper). Intuitively, unless large realizations of earnings become very informative on the actions of the agent, concave contracts have more desirable risk-sharing properties.⁶ To summarize, a first concern is whether risk management can account for the benefits of using convex compensation schemes which become linear when performance is large.

A second empirical regularity is the **hump-shape or divot** observed in the cross-section of corporate earnings (Hayn (1995), Burgstahler and Dichev (1997), Degeorge, Patel and Zeckhauser (1999)). Degeorge, Patel and Zeckhauser show that there are “threshold” effects in realized corporate earnings: few firms report no or slightly negative earnings growth while many firms report slightly positive earnings growth, creating a hump-shape in the cross-section of earnings. Their results are represented in the top of Figure 1 and reveals a clear-cut threshold at zero earnings growth. In addition, their analysis suggests that the distribution is skewed in the tails: more firms report very large losses than very large gains. To summarize this second fact, the following asymmetries are distinctive in the cross-section: (i) for intermediate realizations of earnings, relatively good earnings are more likely than relatively bad ones (i.e., hump-shape), (ii) for extreme realizations of earnings, very bad earnings are more likely than very good ones (i.e., asymmetry in the tails).

⁵Even bonus payments often have convex properties. For example, Murphy finds that 27% of the firms use convex annual incentive schemes versus only 15% using concave schemes (Table 5).

⁶To see this, recall from Holmström (1979) that, if the likelihood ratio converges, the wage should also become flat as earnings become large which would necessarily contradict an increasing convex compensation. From a practical perspective, many observers note that very large earnings, i.e. outside of the “incentive zone,” should be fairly unrelated to effort (i.e., the likelihood ratio may become constant or even decreasing).
Figure 1.

1-year change in EPS
(from Degeorge, Patel and Zeckhauser 1999)

1Q stock Response to Earnings Surprises
(from Skinner and Sloan, 2004)

Figure 1.
Finally, several studies suggest that the response of stock prices to the difference between reported earnings and current consensus is non-linear and exhibits an S-shape pattern (Freeman and Tse (1992), Sloan and Skinner (2004)). As shown in the bottom of Figure 1, stock prices seem to be very responsive to near-median earnings, but not very responsive to large gains or losses. This empirical finding is at odds with the predictions of a learning model with Normal updating (since the update should be essentially linear) and, although more complex earnings processes may rationalize it, researchers often interpret it as evidence that CEOs actively manage their earnings.

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 for now that the firm is liquidated after this period ends and yields a net cash flow \( y \in X \). The manager privately chooses an action \( a \in [a, \bar{a}] \) and then can manage risk by selecting a distribution \( \hat{F}(\cdot) \) 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.

\( \square \) Risk Management

I define \( F(\cdot|a) \in \Gamma(a) \) as the distribution of \( y \) when the agent does not manage risk and assume that it has mean \( a \).\(^7\) In order to reflect the idea that effort increases the value of the firm, I assume that \( F(\cdot|a) \) first-order stochastically dominates \( F(\cdot|a') \) 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 \( a \), \( F(\cdot|a) \) maximizes the expected value of the firm.

\( \square \) Preferences and Technology

\(^7\)The statement is without loss of generality when the mean of \( F(\cdot|a) \) is continuous in \( a \). For example, even if the mean of \( F(\cdot|a) \) is not \( a \), one may always relabel effort \( A = \int ydF(y|a) \).
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 4, 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 Inada conditions and: \( u' > 0, u'' < 0, \psi(a) = \psi'(a) = 0, \lim_{a \to -\infty} \psi'(a) = +\infty, \psi'', \psi''' > 0 \) except possibly at \( a = a \). Finally, the contract must prescribe a minimum reserve utility equal to \( b \).

\[ \square \text{Contracting Problem} \]

A contract \( (w(\cdot), \hat{F}(\cdot), a) \) is incentive-compatible if for a given \( w(\cdot) \) 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.

\[
(P) \quad \max_{a,w(\cdot),\hat{F}(\cdot)} \int (y - w(y))d\hat{F}(y) \\
\text{s.t.} \quad \int u(w(y))d\hat{F}(y) - \psi(a) \geq b \tag{1}
\]

\[
(a, \hat{F}(\cdot)) \in \arg \max_{a,F(\cdot) \in \Gamma(a)} \int u(w(y))dF(y) - \psi(\hat{a}) \tag{2}
\]

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

\[ ^8 \text{Note that the model can also be equivalently stated in the generic formulation of Holmström and Milgrom (1987) (Equations (1)-(3) p.307). In their language, the agent chooses an action from a set } p \in P \text{ (here: } \{(a, \hat{F}) \in [a, \pi] \cup \Gamma(a)\} \text{ and pays a cost } c(p) \text{ (here: } \psi(a)). \]
compensation scheme is linear when \( w(y) = h_0 + h_1 y \). Finally, let \( \theta = \inf X \) denote the maximum loss that the agent may make.

\( \square \) **Mean-Distance Ordering**

Putting more structure on the set of available distributions \( \Gamma(a) \), I introduce the following assumption which I call Mean-Distance Ordering (MDO). I suppose an agent exerting effort \( a \) can achieve any distribution \( P(a) + \Delta \) (where \( \Delta \equiv (\delta_k)_{k=1}^N \)) for \( y \) satisfying:

\[
\sum_{k=1}^N \delta_k = 0 \tag{3}
\]

\[
\sum_{k=1}^N \delta_k y_k \leq - \sum_{k=1}^N C(a, y_k, \delta_k) \tag{4}
\]

Equation (3) is required for \( P(a) + \Delta \) to be valid probability distribution. Equation (4) can be equivalently written as \( \mathbb{E}(y|a) = \mathbb{E}(x|a) - \sum_{k=1}^N C(a, y_k, \delta_k) \) and captures the cost of managing risk. The function \( C(a, y, \delta) \) represents a cost function and satisfies standard restrictions (i.e., \( C_{\delta,\delta}(a, y, \delta) > 0 \), \( C(a, y, 0) = C_\delta(a, y, 0) = 0 \), \( C_\delta(a, y, \delta) < 0 \) for \( \delta < 0 \) and \( C_\delta(a, y, \delta) > 0 \) for \( \delta > 0 \)). Further, to avoid negative probability weights, I assume the cost of managing risk becomes large if the manager attempts to change the support of \( x \), i.e. \( \lim_{\delta \to -p_k(a)} C(a, y_k, \delta) \geq \lim_{n \to N} y_n - y_1 \).

Here, I interpret managing risk as moving the distribution from \( P(a) \) to \( P(a) + \Delta \). As the agent manages risk more, in that the distribution becomes more distinct from the original distribution, the cost \( \sum_{k=1}^N C(a, y_k, \delta_k) \) increases.\(^9\) I interpret this assumption as capturing risk management from an operational perspective, as the manager must trade off between choosing projects with the maximum value (which would generate \( y = x \)) and choosing projects with different risk characteristics but lower value (when \( \Delta \) is not zero).\(^{10}\)

\(^9\)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).

\(^{10}\)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
A useful feature of MDO is that it does not restrict the shape of the distributions \( P(a) + \Delta \) feasible by the agent and thus it is well equipped to capture anomalies such as the hump-shape in the cross-section of corporate earnings. From a theoretical perspective, it also captures a reasonable aspect of risk management since it implies that any shape for the distribution of \( y \) can be realized for a cost (i.e., by reducing the mean cash flow sufficiently). In other words, by increasing effort sufficiently, the manager is able to realize almost any distribution \( P(a) + \Delta \).

Adapting the contract design problem to MDO yields:

\[
(P) \quad \max_{a, \Delta} \sum_{k=1}^{N} (p_k(a) + \delta_k)(y_k - w(y_k))
\]

s.t. \[ \sum_{k=1}^{N} (p_k(a) + \delta_k)u(w(y_k)) - \psi(a) \geq b \quad (5) \]

\[
(P_a) \quad (a, \Delta) \in \arg \max_{a, \Delta} \sum_{k=1}^{N} (p_k(a) + \tilde{\delta}_k)u(w(y_k)) - \psi(\tilde{a})
\]

s.t. (3) and (4)

2. Analysis of the Contract

2.1. Problem of the Agent

Taking the wage as given, I solve first the problem of the agent. In this problem, let \( \lambda \) (resp. \( \mu \)) denote the multiplier associated to Equation (3) (resp. (4)). Denote \( C(\delta, a, y, \delta) \) the inverse of \( C(\delta, y, \delta) \) in \( \delta \). The next Proposition characterizes the optimal choice of the agent.

**Proposition 2.1.** Suppose \( \sum_{k=1}^{N} |u(w(y_k))/(y_k - y_1)| < +\infty \) and \( \inf_k y_k \) finite. Then, a solution to \( (P_a) \) exists. If, in addition, \( a > a^* \) is elicited, the solution satisfies \( \mu > 0 \) and

\[
\delta_k = C(y_k, a, \frac{u(w(y_k)) - \lambda}{\mu} - y_k)
\]

**Corollary 2.1.** The agent does not hedge if and only if \( u(w(y)) = h_0 + h_1 y \).

which the decision to hedge is private (see Morellec and Smith (2007) for a model in which the principal manages risk).
\textbf{Proof:} First, $\delta_k = 0$ for all $k$ implies that $u(w(y_k)) = \lambda + \mu y_k$ (necessity). Second, let $u(w(y_k)) = h_0 + h_1 y_k$. Then, the agent must achieve $h_0 + h_1a$, which can be achieved with no risk management (sufficiency). $\square$

In the model, non-linearities (in utility terms) induce the agent to strategically hedge to align the contract payments with his self interest. In particular, a concave or linear wage schedule will induce some risk management with, as a result, a positive deadweight loss $a - \mathbb{E}(y|a)$. Equation (7) also provides several preliminary comparative statics: 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. A simple example is given next to illustrate these properties.

**Corollary 2.2.** Assume that $N$ is finite, $C(a,y,\delta) = c(a)\delta^2/2$ with $c(a) > 0$ sufficiently large and $u(x) = x$.\textsuperscript{11} Then:

\begin{align*}
\lambda &= \mathbb{E}(\hat{w}(y)|a) - \mathbb{E}(\hat{y}|a) \frac{\sigma(\hat{w}(y)|a)}{\sigma(\hat{y}|a)} \\
\mu &= \frac{\sigma(\hat{w}(y)|a)}{\sigma(\hat{y}|a)} \\
\delta_k &= \frac{w(y_k) - \lambda - \mu y_k}{\mu c(a)}
\end{align*}

where $\hat{w}(y) = w(y)/p(y)$ (resp. $\hat{y} = y/p(y)$) is the scaled wage (resp. output) and $p(y)$ is the probability of outcome $y$ and $\sigma(.)$ is the standard deviation of the random variable.

Under quadratic cost, the slope of the threshold $(\lambda + \mu y)$ 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 (9)). This aspect is intuitive because an agent with greater wage variability is more sensitive to output shocks. A second force may go against this intuition. In Equation (10), $|\delta_k|$ is decreasing in $\mu$, that is the magnitude of the risk management choice will be decreasing in wage variability. This is because a variable wage

\textsuperscript{11}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.
makes the agent more sensitive to the cost and thus incentivizes the agent to manage risk less. This second part of the trade-off gives some preliminary intuition for why providing incentives to work hard (with high performance-pay coefficients) may not be well-aligned with giving incentives to manage risk (with high values of $|\delta_k|$).

Although I will soon endogenize the contract, it is helpful at this point to illustrate the predicted risk management in response to a simple Call option. By Proposition 2.1, to obtain which events will be hedged, it is sufficient to represent on the same graph: 1. the utility received by the agent, $u(w(y))$, 2. a linear threshold, $\lambda + \mu y$. The manager will then increase (decrease) the likelihood of outcomes with utility above (below) the threshold. In the case of a Call option, the location of the linear threshold $\lambda + \mu y$ can be recovered graphically. In Figure 2, I plot the utility received by a risk-averse manager compensated with options.

**Proposition 2.2.** Suppose: $|\inf_k y_k| = \sup_k y_k = +\infty$, $\lim_{x \to +\infty} u'(x) = 0$ and $w(y) = \max(0,y)$ (standard Call option). Then, there exists four regions $\theta_1 < 0 < \theta_2 < \theta_3$ such that:

(i) For $y_k < \theta_1, \delta_k > 0$, i.e. the manager increases the likelihood that the option matures out-of-the-money.

(ii) For $y_k \in (\theta_1, \theta_2)$, $\delta_k < 0$, i.e. the manager reduces the likelihood that the option matures at-the-money.

(iii) For $y_k \in (\theta_2, \theta_3)$, $\delta_k > 0$, i.e. the manager increases the likelihood that the option matures in-the-money.

(iv) For $y_k > \theta_3, \delta_k < 0$, i.e. the manager reduces the likelihood that the option matures far in-the-money.

**Proof:** The statement is proved graphically. I omit the case $\mu = 0$ which can be easily disproved. Clearly, because $u'(x)$ converges to zero, the threshold is above $u(w(y))$ for $y$ large and, because $u(w(y))$ is bounded from below, the threshold is below $u(w(y))$ for $y$ small. There are two cases consider. Case 1: $\lambda + \mu y$ intersects $u(w(y))$ only once.\(^\text{12}\) But then this would imply that the manager increases (decreases) the likelihood of large (low) $y_k$, a contradiction to

\(^{12}\)To ease the exposition, by $u(w(y))$, I mean the linear interpolation across values of $y_k$. 

12
Equation (4). Case 2: $\lambda + \mu y$ intersects $u(w(y))$ only twice. In particular, it intersects at least once above the strike price. But then by concavity, it must intersect twice, and thus $u(w(y))$ must stay above the threshold for all $y$ small. To conclude, note that $u(w(y))$ cannot intersect $\lambda + \mu y$ more than three times. □

In the left-hand side of Figure 2, I plot the response of a risk-averse manager. Under MDO, the manager behaves in the manner documented empirically, i.e. (i) hedges against the option maturing at-the-money or far in-the-money and (ii) increases the likelihood that the option matures out-of-the-money or slightly in-the-money.

I develop this argument further with an example showing whether the resulting distribution of $y$ may feature the extreme bunching or asymmetry found empirically. I assume that the agent has a utility function $u(x) = 2\sqrt{x}$. 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 one. Finally, I assume that the cost of managing risk is $C(y, y, \delta) = e^{-1/2(y-1)^2}$ for any $y$.

The resulting distribution of $y$ is plotted against that of $x$, on the right-hand side of Figure 2. When the agent manages earnings, the distribution features a kink at the strike price and then a large number of firms reporting earnings to beat zero output. This distribution is similar to the cross-sectional evidence in Figure 1.

2.2. Problem of the Principal

I analyze now the full contract design problem faced by the principal. The optimal contract must be mindful of two trade-offs. One, the contract must trade off incentives to increase value and risk-aversion by the agent (Holmström 1979). Two, the contract must trade off the cost of risk management and the benefits of eliciting a distribution $P(a) + \Delta$ that is more informative on the actions of the agent.

Proposition 2.3. The agent receives a contract eliciting $a > a$.

---

13 Since the model is stated discretely, I use here a version of the model with an arbitrarily fine grid to approximate the Normal distribution.

14 A two-step bonus scheme will have the same features as a Call option (which can also be established with simple graphical arguments).
Risk Management in Response to a Call

Cross-Section of Earnings

Figure 2.
By choosing \( a = a \), the principal can offer a flat contract that does not induce costly risk management. This may be described as an extreme solution to the problem if the principal believes that the cost associated to risk management is too important. Proposition 2.3 establishes this extreme solution is not chosen and some incentives are still preserved in the model. Intuitively, the principal can use contract that is linear in utility \( u(w(y)) \) linear and that will not induce costly risk management.\(^{15}\)

An important technical aside is required at this point. In my problem, the first-order approach may not be valid and, as is well-known, finding sufficient conditions for it to hold in non-standard problems is difficult.\(^{16}\) Fortunately, the first-order approach is not required for the results presented here. To see this, note that by Equation (7), for a given \( (a, \lambda, \mu) \), the choice of \( \Delta \) is unique. Then, taking \( (a, \lambda, \mu) \) as chosen optimally, the risk management choices will always be a solution to the following reduced problem \((P')\):

\[
(P') \qquad \max_{\Delta} \sum_{k=1}^{N} (\delta_k + p_k(a))(y_k - u^{-1}[C_\delta(a, y_k, \delta_k)\mu + \lambda + \mu y_k])
\]

s.t.

\[
\sum_{k=1}^{N} \delta_k = 0 \quad (\alpha) \tag{11}
\]

\[
\sum_{k=1}^{N} \delta_k y_k = -\sum_{k=1}^{N} C(a, y_k, \delta_k) \quad (\beta) \tag{12}
\]

\[
\mu = \psi'(a) - \mu \sum_{k=1}^{N} p_k'(a)C_\delta(a, y_k, \delta_k) \quad (\gamma) \tag{13}
\]

\[
b - \lambda - \mu a = -\psi(a) + \mu \sum_{k=1}^{N} [(\delta_k + p_k(a))C_\delta(a, y_k, \delta_k) - C(a, y_k, \delta_k)] \quad (\tau) \tag{14}
\]

In Problem \((P')\), I simply state that \( a, \lambda \) and \( \mu \) are chosen optimally but I do not use first-order conditions (which would require the first-order approach) to characterize their choice.\(^{17}\)

Equations (11) and (12) are the MDO feasibility conditions. Equation (13) is the agent’s op-

---

\(^{15}\)Note also that by incentive-compatibility, the compensation of the agent \( w(y) \) may not be constant and thus there must be imperfect risk-sharing between the agent and the (risk-neutral) principal.

\(^{16}\)The only paper that solves for an optimal contract in a fairly unspecified agency problem is Palomino and Prat (2003); but their analysis requires the first-order approach.

\(^{17}\)Moreover, these additional first-order conditions do not have a simple interpretation here.
timality condition on effort. Equation (14) is the participation of the agent (once wages have been substituted out). The multipliers associated to each constraint are denoted in parenthesis.

Proposition 2.4. Let $a$ be the elicited effort. The optimal contract satisfies:

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

$$- \alpha + (1 - \beta)y_k - \beta C_\delta(a, y_k, \delta_k) - \mu \gamma C_{\delta,a}(a, y_k, \delta_k)$$  \hspace{1cm} (15)

I interpret next this optimality condition. Equation (15) decomposes the main contract design problem faced by the firm into two aspects.

First, the contract must solve the trade-off between incentives to create value and efficient risk-sharing. This is captured here by the term $S = 1/u'(w(y_k)) - \tau - \gamma p'_k(a)/(p_k(a) + \delta_k)$.

It should be set at zero in a problem without risk management (see Equation (7) in Holmström (1979)). In this problem, $S$ positive means that the wage is too large as compared to what is required in the standard model. Then, Equation (15) shows that this force will work to reduce the wage offered by the firm.

Second, since the presence of a wage makes the principal effectively non-risk-neutral (since $y - w(y)$ is typically non-linear), the principal has additional incentives to manage risk. In the model, the first part of Equation (15), $-\alpha + (1 - \beta)y_k - \beta C_\delta(a, y_k, \delta_k)$, captures the most preferred risk management choice of a principal controlling risk. Thus, when $S = 0$, the principal behaves myopically as if taking the wage as given exogenously and managing risk optimally in response to it.

Corollary 2.3. Suppose that $C_{\delta,\delta}(a, y_k, 0) = C_{\delta,a}(a, y_k, 0) = 0$ for all $a, y_k$, then $\delta_k = 0$ for all $k$ cannot be optimal.

Proof: If not managing risk is optimal, $u(w_k) = \lambda + \mu y_k$. But then, by Equation (15), $w(y_k) = \alpha + \beta y_k$, but since $u$ is strictly concave, this is a contradiction. □

Risk management will always be desirable if the cost of managing risk decreases fast for small risk management choices. Intuitively, some risk management can always raise the useful-
ness of the output signal at very little cost.\textsuperscript{18} That is, the agency problem makes a risk-neutral principal effectively risk-averse to some of the noise in the production technology.

3. Explaining the Stylized Empirical Facts

3.1. Nature of the Optimal Contract

Leaving aside for now the observed convexity in compensation contracts, I explain first why the optimal contract should become linear over large outcomes.\textsuperscript{19}

**Proposition 3.1.** Suppose that for all $a, \delta, y$,

(i) $|p'_k(a)|$ is bounded by a number that does not depend on $k$,

(ii) $C_{\delta,\delta}(a, y_k, 0) \leq C(a, y_k, 0)$ for $k$ large enough.

(iii) $u'(w)$ converges when $w$ becomes large.

Then, $w(y_k)$ converges to a linear function of $y_k$ as $y_k$ 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 (the second side of the trade-off studied in the previous Section).\textsuperscript{20} For states in the tail of the distribution, this concern dominates any improvements in the likelihood ratio. This result may seem surprising as compared to standard agency theory. For example, if the likelihood ratio becomes constant ($p'_k(a)/p_k(a)$ converges), the optimal compensation in Holmström (1979) should become flat. Here, choosing a compensation that becomes flat may generate too strong incentives to reduce the likelihood of large earnings. This is often undesirable for the principal because it reduces the probability of

\textsuperscript{18}It should be noted that, typically, $C_\delta(a, y, 0) = 0$ is not sufficient to guarantee some equilibrium risk management. Later on, I show that if $y_k$ can grow large, the assumption that $C_{\delta,\delta}(a, y, 0) = 0$ can be lifted.

\textsuperscript{19}Note that, in the standard model, $w(y)$ may also become linear if $(u')^{-1}(1/(\lambda + \mu f(y)))$ becomes linear. This condition, however, does not map into a clear economic interpretation and would likely be violated in practice.

\textsuperscript{20}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.
other outcomes informative on the actions of the agent while simultaneously generating large risk management cost.\footnote{In fact, the knife-edge case $\beta = 1$ can be removed if the cost of managing risk in the tails is sufficiently large since $\mu > 0$ implies that the agent would do considerable risk management in the tails if the contract did become flat.}

A limiting contract can also be obtained when the cost of managing risk becomes small. To consider this case, I state a sequence of problems with a cost function $C_j(a, y_k, \delta) = C(a, y_k, \delta)/j$ ($j > 1$). As $j$ becomes large, the cost function becomes small. In addition, I make the assumption that $C(a, y_k, \delta_k)$ becomes large (resp. $\frac{C(a, y_k, \delta_k)}{C(a, y_k, \delta_k)}$ is bounded away from zero) when $\delta_k$ converges to $-p_k(a)$ or $1 - p_k(a)$ for all $(y_k, a)$. In intuitive terms, this assumption means that the elasticity of the cost of managing risk to a change in risk does not become too small.

**Proposition 3.2.** Let $a^*$ be given by: $\psi'(a^*) = u'(u^{-1}(\psi(a^*) + \tilde{b}))$ and assume that there exists $k^*$ such that $y_{k^*} = a^*$. Then, as $j$ becomes large, the contract converges to a unique (convex) contract:

$$w(y_k) = u^{-1}(\psi'(a^*)(y_k - a^*) + \tilde{b})$$

(16)

If $u'(w)$ converges to a finite non-zero constant when $w$ becomes large (risk-neutrality in the limit), then $w(.)$ becomes linear as $w$ becomes large.

I argue here that the contract should become linear in utility as the cost of managing risk becomes small, but provided that the cost of managing risk increases sufficiently fast for policies close to infeasibility. Intuitively, only small non-linearities are required to give incentives to manage risk; important non-linearities, on the other hand, may lead to cost that are unnecessarily large.\footnote{In this limiting case, the optimal contract converges to the first-best outcome (i.e., effort under the control of the principal), so that the informational friction is fully resolved. This is also the case, when normalizing by the firm’s cash flow, in Diamond (1998). As in his model, this limiting argument selects a unique limiting contract (there would be no notion of a unique optimal contract if stating the first-best problem directly).} This result shows that incentives to offset the agent’s risk-aversion require to offer a wage that is convex in money; for example, Bebchuk and Fried (2004) explain that convexity is needed to align the risk-management incentives of a risk-averse agent with those of risk-neutral shareholders: “Because managers are insufficiently diversified and risk-averse, they may hesitate to take chances that would be desirable for shareholders. Options are believed to
counteract this tendency by providing executives with a financial incentive to take risks. [...] Strike Prices that are too high or too low can cause executives to take too much or too little risk” (p. 159).

3.2. Cross-Section of Earnings

I develop next several comparative statics derived from the optimality conditions of the problem. For a particular state of the world, it is helpful first to rewrite Equations (7) and (15) in terms of their continuous analogue (omitting the indices and denoting the likelihood ratio \( LR = \frac{p_k'(a)}{p_k(a)} \)):

\[
\begin{align*}
  w & = u^{-1}(\lambda + \mu y + \mu C_\delta(a, y, \delta)) \\
  w & = -\alpha + (1 - \beta)y - \beta C_\delta(a, y, \delta) - \mu \gamma C_{\delta,a}(a, y, \delta) - \mu C_{\delta,\delta}(a, y, \delta) \left( \int \left( 1 - \frac{u'(w)}{u(w)} - \tau \right) - \gamma p LR \right)
\end{align*}
\]

In the model, there are only two endogenous variables that depend on the state of the world, \( w \) and \( \delta \). Provided the number of states is large and the distribution \( P(a) \) 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 \( y | a \) being absolutely continuous). As an approximation to a continuous problem, I take here the multipliers and effort as constants. Then, one may view Equations (17) and (18) as two Equations in two unknowns \((w, \delta)\) where all the other terms are taken as exogenous constants.

To facilitate the analysis at this point, I remove some of the cross-effects in the cost of managing risk. I make the following assumptions: For all \( a, y, \delta \), \( C_{\delta,\delta,a}(a, y, \delta) = C_{\delta,y}(a, y, \delta) = C_{\delta,a,y}(a, y, \delta) = C_{\delta,a,y}(a, y, \delta) = 0 \), \( C_{\delta,\delta,\delta}(a, y, \delta) < 0 \) and bounded away from zero. These conditions restrict the cross-effects in the model and are not all necessary for each comparative static taken separately. The following comparative statics are obtained from the Implicit Function theorem applied on Equations (17) and (18).
Corollary 3.1. The following comparative statics hold:

(i) For states such that the wage is sufficiently large, \( \partial \delta / \partial y < 0 \), i.e. the manager reduces the likelihood of states with large payoffs.

(ii) If \( \gamma LR \geq 0 \) is sufficiently large, \( \partial \delta / \partial LR > 0 \), i.e. the manager reduces the likelihood of states with lower likelihood ratio.\(^{23}\)

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

First, I show that optimal contract elicits some risk management against states with large payoffs. This result confirms the preliminary comparative statics in the problem of the agent. In the model, reducing the likelihood of states with large payoffs allows the manager to increase the likelihood of many states with lower payoffs and (relatively) high wages. While this motive can be offset by promising a very high compensation conditional on large earnings, such an arrangement would be very inefficient from the perspective of risk-sharing. In other words, the principal must trade off paying the agent more for large earnings and tolerating some costly risk management over large earnings. This is one aspect of the cross-section documented in Degeorge et al. (1999).

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 risk management strategy that increases their likelihood. This aspect is intuitive and shows that, indeed, the principal is eliciting risk management choices that raise the informativeness of the output signal. An initial motivation for a hump-shape can be obtained at this point. Suppose that the likelihood ratio changes fast close to the mode of the distribution (e.g., the likelihood ratio is S-shaped).\(^{24}\) Then, this comparative static will work to reduce the likelihood of negative likelihood ratio states and increase the likelihood of positive likelihood

---

\(^{23}\)One may guarantee that \( \gamma > 0 \) if the effort choice may take only two values. Further, by continuity, \( \gamma \) should be positive provided the cost of managing risk is sufficiently large (since the property is true in Holmström (1979) and the current model becomes equivalent to his when the cost of managing risk is large).

\(^{24}\)An intuition for an S-shaped likelihood ratio is that most of the gains in informativeness are realized over intermediate outcomes.
ratio states. Jointly with the fact that $\partial \delta / \partial y < 0$, this will produce a hump-shape close to the point where the likelihood ratio changes.

Third, the last comparative static explains why the hump-shape should be, in fact, located close to the mode of the original distribution of $x$. To see this, suppose that the distribution $P(a)$ is bell-shaped and the likelihood ratio follows an S-shape (as before) and changes sign close to (but slightly before) the mode of the distribution of $x$. In this case, as $p$ increases as $y$ moves toward the mode, $y$ will lie in a region where $\partial \delta / \partial p < 0$. This will work to flatten the distribution. As the likelihood ratio changes sign before the mode, $y$ will lie in a region where $\partial \delta / \partial p > 0$, producing a peak in the distribution. Finally, $y$ will reach the mode, at which point $p$ will start decreasing. This will sharpen the peak produced in the previous region. Thus, this last comparative static explains why regions close to the mode of a bell-shaped distribution reinforce the S-shape.

### 3.3. A Numerical Example

To illustrate the findings, I develop a numerical example inspired from the LEN framework. The model is discretized in two effort choices $a \in \{0, \bar{a}\}$. Conditional on $a \in \{0, \bar{a}\}$, the distribution of outputs is normally distributed with mean $a$ and variance $\sigma^2$. The set of outputs is discretized over $N = 200$ points $(y_k)_{k=1}^N$ and such that each point has equal probability conditional on $F(x|\bar{a})$. The manager has a utility function $u(x) = -e^{-rx}$. Finally, the cost function is set equal to $C(a, x, \delta) = c_h e^{-|x-a|\delta^2/2}$.

The benchmark parameters are set as follows: $\sigma = 1$, $\bar{a} = .5$, $r = 1$, $\psi(\bar{a}) = .1$, $c_h = 1$ and $b = -1$. The results are shown in Figure 3.3 under four treatments: 1. First-Best Contract, 2. Second-Best without risk management (i.e., standard agency problem), 3. Second-Best with risk management, 4. Second-Best with risk management but constrained to a Linear contract. The optimal contract with risk management (3.) is surprisingly close to a linear contract, in sharp contrast with the optimal contract when the manager cannot manage risk (2.). Thus, the model performs well in matching convex and then linear compensation schemes. On the other hand, in this Normal specification, the change in the likelihood ratio is not fast enough close to the mode (i.e. it is not S-shaped as required in the comparative statics) to produce a clear
To facilitate a hump-shape even for slow changes in the likelihood ratio, I assume that the agent faces a limited liability constraint and must be paid more than $-1$. In the right-hand side of Figure 3.3, I show that the wage is remarkably close to linear whenever the limited liability does not bind and exhibits a familiar option-like shape. In contrast, the standard model would generate a concave optimal wage. In the left-hand side, I show that this option generates a hump-shape in the distribution of earnings even though the original distribution is bell-shaped.
4. Extensions

4.1. Income Smoothing and S-Shape

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.

For notational convenience, I rewrite first the model in terms of a problem in which the
distribution of \( y \) is absolutely-continuous.\(^{25}\) I assume now that \( x \) has an absolutely-continuous distribution with density \( f(\cdot|a) \). The concept of MDO is simplified and adapted to reflect a continuous set of outcomes. I assume that there exists a positive convex cost function \( C(\delta) \) such that \( C(0) = C'(0) = 0 \), \( C \) is positive, \( C''(\delta) < 0 \) (resp. \( C''(\delta) > 0 \)) for \( \delta < 0 \) (resp. \( \delta > 0 \)) and \( \lim_{\delta \to -\infty} \max_a f(x|a) C'(\delta) = +\infty \).

Next, I follow the standard methodology developed in Debreu (1972) to map the static model into a multi-period one with a simple change of notations. Let \( \gamma = (y^1, ..., y^T) \) where \( y^i > 0 \) for all \( i \), be a sequence of realizations of returns from period \( t = 1, ..., T \); \( f(\gamma|a) \) denotes the (multivariate) density of \( \gamma \) when the agent does not manage risk while \( \hat{f}(\gamma) \) denotes the distribution of returns when risk is managed. As before, I assume that the manager chooses \( a \) for a personal cost \( \psi(a) \). Explicitly writing the MDO restriction in this case yields:

\[
\int \hat{f}(\gamma) \prod_{t=1}^{T} y^t dy \leq a - \int C(f(\gamma|a) - \hat{f}(\gamma))dy
\]  \((19)\)

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

\[
\int \hat{f}(\gamma) dy = 1
\]  \((20)\)

In addition, to keep intact the contract design problem studied earlier, I assume that the wage may only depend on \( \prod_{t=1}^{T} y^t \) (and not on each individual \( y^i \)). It should be clear that, by expressing the model in multiple period in this manner, the previous framework is essentially unchanged. Let \( \delta(\gamma) = f(\gamma|a) - \hat{f}(\gamma) \). Stating the first-order condition in the problem of the agent yields:

\[
u(w(\prod_{t=1}^{T} y^t)) - \lambda \prod_{t=1}^{T} y^t - \mu C(\delta(\gamma)) = 0
\]  \((21)\)

The graphical analysis of the dynamic case is the same as in the static case. Note that the threshold governing whether the manager reduces or increases the likelihood of an event

\(^{25}\) All the results can also be stated in the discrete case but at the cost of substantially burdening the exposition. This is because the support of per-period returns must be tied to the support of total earnings. I avoid the continuous case in the previous Section for technical reasons, since there are cases such that even \( x|a \) is absolutely-continuous, \( y|a \) may not be under MDO (i.e., it may have mass points).
is a linear function of $\prod_{t=1}^{T} y^t$, a product function. This function will work to encourage the manager to increase the likelihood of realizations of $y$ such that each component of $(y^i)_{i=1}^{T}$ lies close together and thus produce income smoothing motives.

The model also yields several predictions on the dynamics of returns. Let $1 < t' < T$ be a time period, and let $\tilde{y}^{t'} = \prod_{t=1}^{t'} y^t$ (resp. $\tilde{x}^{t'} = \prod_{t=1}^{t'} x^t$) be the managed (resp. original) 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 managed (resp. original) return after date $t'$.

**Proposition 4.1.** The following holds:

$$\text{cov}(y^{t'}, \overline{y}^{t'}) \leq \text{cov}(\tilde{x}^{t'}, \overline{x}^{t'}) - (\mathbb{E}(\tilde{x}^{t'})\mathbb{E}(\overline{x}^{t'}) - \mathbb{E}(y^{t'})\mathbb{E}(\overline{y}^{t'}))$$

(22)

The inequality is strict when the manager manages risk.

Comparing the covariance between periods under risk management to the original covariance, one of the following facts must be true when risk management occurs: the manager reduces the covariance of returns across periods (i.e. $\text{cov}(y^{t'}, \overline{y}^{t'}) < \text{cov}(\tilde{x}^{t'}, \overline{x}^{t'})$) or changes expected returns per period (the term $\mathbb{E}(\tilde{x}^{t'})\mathbb{E}(\overline{x}^{t'}) - \mathbb{E}(y^{t'})\mathbb{E}(\overline{y}^{t'})$). Both forms of risk management 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 risk management is not too intense (e.g., the cost of managing risk 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.

I turn now to the S-Shaped response of the market price to current returns. In this model, the market value of the firm at $t'$, after $y^{t'}$ has been announced, is captured by $MV(y^{t'})$:

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

(23)

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 $y^1$ and $y^2$ are i.i.d. and distributed uniformly on $[0, 2]$ (the location of the support does not affect the results). Finally, the cost of risk management is set as in the previous Section. Suppose that the manager is risk-neutral and compensated with a simple Call option with maturity date 2 and strike 1. I use a Call option because the previous results reveal that a Call option will be close to optimal in settings with risk management and a limited liability.

In Figure 5, the density $\delta(y)$ 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 (23) as a function of $y^1$, the following response to $y^1$ which exhibits an S-Shaped profile where the response is steeper near the median return.\footnote{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).}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{expected_return.png}
\caption{Earnings Response}
\end{figure}
4.2. Risk-Aversion by the Principal

I extend 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.

In this extension, I use the continuous formulation of MDO developed in earlier. Let \( v(\cdot) \) 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 risk are chosen by the principal).

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

s.t.

\[
\int \delta(y)dy = 0 \quad (\hat{\alpha})
\]

\[
\int \delta(y)ydy = -\int C(\delta(y))dy \quad (\hat{\beta})
\]

\[
\psi(a) + b \leq \int u(w(y))(\delta(y) + f(y|a))dy \quad (\hat{\tau})
\]

**Proposition 4.2.** In first-best, no risk management cannot be optimal.

In comparison to the previous Section, some risk management 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 2.3 (when the principal is risk-neutral), it may now be optimal to offer a flat contract and elicit \( a \). First, the distribution \( f(y|a) \) 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 \( a \)).

Suppose first that a flat contract is optimal. Then, one needs to substitute in Problem \( (P_{fb}) \), \( w(y) \) by \( u^{-1}(\psi(a) + b) \) (so that the participation binds) and set \( a = \bar{a} \). 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. 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 $a = a$ is not optimal and characterize the optimal risk choices in this case.\textsuperscript{27} 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. The contract design is restated by introducing the utility $v(.)$.

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

s.t.

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

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

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

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

This problem is the same as Problem $\mathcal{P}_a$ 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 4.3.** Suppose the first-order approach is valid. Then, 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

\textsuperscript{27}I omit the case in which $a$ is optimal but not with a flat contract since this case is not particularly interesting.
required for incentive purposes and risk-sharing must remain imperfect.

I discuss next whether risk management is desirable when the principal is risk-averse. As suggested in first-best, since risk management is now used without informational frictions, a preliminary intuition would suggest that risk management should remain optimal here.

**Proposition 4.4.** 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, managing risk is optimal. Else, if not managing risk is optimal and $\lim_{z \to +\infty} u'(z) = 0$, $v'(y - w(y))/u'(w(y))$ is strictly increasing for $y$ sufficiently large.

Unlike with a risk-neutral principal, not managing risk may be optimal. 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 risk management. Here, given that risk management 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 risk management, 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 risk management 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 risk management 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. Concluding Remarks

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 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 imperfections 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.

The paper also presents other contributions to the theoretical and empirical literature on managerial risk-taking, corporate hedging, and earnings management. Several authors consider agency problems in which the agent can privately affect risk. Diamond (1998) shows that linear contracts can be optimal when the size of the firm is large relative to the agent. Biais and Casamatta (1999) consider an agent who has access to two projects. Palomino and Prat (2003) further discuss problems in which the agent has access to a set of distributions indexed by a risk metric. They show that, under certain conditions, a bonus-like contract is optimal. Prescott (2004) and Liang and Nan (2007) consider versions of a Linear-Exponential-Normal (LEN) model in which risk can be reduced by the agent. Other authors consider manipulation of the firm cash flows, although not by risk-taking. Dye (1988) and Fischer and Verrecchia (2000) consider settings in which the agent can alter reported cash flows and examine implications for valuation risk. Goldman and Slezak (2006) consider a LEN model with a costly upward bias and suppose that this may involve an opportunity cost; in comparison, in this model, the agent will be able to change the complete distribution of cash flows, also possibly for a cost. Their focus is on the pay-for-performance coefficients and accounting bias, while mine is on the shape of the optimal contract and the distribution of reported cash flows.

This paper contributes to this literature in several aspects. First, the model accommodates risk-aversion and continuous effort and nests the standard framework of Holmström (1979) as a

---

limiting case. Second, most of this literature focuses on the optimal contract and not the shape of the distribution chosen by the agent. In comparison, the model delivers intuitive predictions with respect to the cross-section of earnings. In particular, this model is the first to reproduce in a common theory the three stylized facts described previously. \(^{29}\)

The results also relate to an extensive literature on executive contracts and risk-taking. Since the effect of giving options on managerial risk-taking is, in theory, ambiguous (Ross (2004), Braido and Ferreira (2006)), several authors discuss whether options increase or reduce risk-taking behavior. Hanlon, Rajgopal and Shevlin (2003), Covas (2004) and Coles, Daniel and Lalitha (2006) relate stock options received by CEOs to more risky investment choices. Similarly, the risk management choices of mutual fund managers seem to be closely related to the industry’s incentive structure (Chevalier and Ellison (1999), Carhart, Kaniel, Musto and Reed (2002), Elton, Gruber and Blake (2003), Golec and Starks (2004)). The model shows why option-like contract shapes have desirable risk-management properties and describes what risks one would expect managers to be taking. In particular, I show that option-like payments can lead to a project choice that is more informative on managerial effort, but should also induce the manager to take some (small) risks close to the option’s strike price.

Appendix A: Agent’s Problem under Quadratic Cost

I relax now 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, \(a \in \{0, \pi\}\), and assume that \(\pi\) is sufficiently small so that eliciting \(\pi\) is optimal for the principal.

Conditional on \(a\) the probability of each outcome is \(P(a) = (p_1(a), \ldots, p_N(a))^\prime\), 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 \(W = (w_1, \ldots, w_N)^\prime\). In vector notation, denote \(U = (u(w_1), \ldots, u(w_N))^\prime\). Let \(\theta\) be the total cost of hedging, defined as a function of \(a\),

\(^{29}\)There are also other authors that attempt to explain these stylized facts outside of a moral hazard problem. Crocker and Huddart (2006) use a two-period model in which earnings are imperfectly observable in the first period and the manager can misreport for a cost. Li (2007) shows that an S-shape response of the stock price to current earnings can be reproduced if capital markets use current earnings to learn both future earnings and the precision of the manager’s information. Further, if managers can also misreport earnings, the cross-section of earnings will feature a hump-shape.
$P(a)$ and $\hat{P}$. Assume that hedging is small so that $\theta$ can be approximated using the following Taylor expansion for $\hat{P}$ close to $P(a)$.

$$
\theta(P(a), \hat{P}) \approx \theta(P(a), P(a)) + D\theta(\hat{P} - P(a)) + (\hat{P} - P(a))'D^2\theta(\hat{P} - P(a))/2
$$

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.

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

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:

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

s.t.

$$
\begin{align*}
\|\Delta\| & = 0 \\
\Delta'X & \leq -\Delta'H\Delta/2
\end{align*}
$$

Differentiating with respect to $\Delta$ and rearranging: $\mu H \Delta = U - \lambda - \mu X$. Pre-multiplying by $\Delta'$ and using Equation (A-4), $-2\mu \Delta'X = \Delta'(U - \lambda - \mu X)$. Simplifying and substituting $\Delta$ yields: $(U - \lambda + \mu X)'H^{-1}(U - \lambda - \mu X) = 0$. If instead, one pre-multiplies by $\|\|H^{-1}$ and use Equation (A-3), $\|\|H^{-1}(U - \lambda - \mu X) = 0$. One obtains a system of two Equations in two unknowns which yields the following second-order polynomial for $\mu$,

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

This system has two real roots. One is negative and thus cannot be optimal and the other yields the following characterization:
\[ \lambda = \frac{\mathbb{I}'H^{-1}U - \mathbb{I}'H^{-1}X}{\mathbb{I}'H^{-1}U - \mathbb{I}'H^{-1}X} \sqrt{\left(\mathbb{I}'H^{-1}U\mathbb{I}'H^{-1}U - (\mathbb{I}'H^{-1}U)^2\right)} \quad \text{(A-5)} \]

\[ \mu = \sqrt{\left(\mathbb{I}'H^{-1}U\mathbb{I}'H^{-1}U - (\mathbb{I}'H^{-1}U)^2\right)} \quad \text{(A-6)} \]

\[ \mu \Delta = H^{-1}(U - \lambda - \mu X) \quad \text{(A-7)} \]

A simple application of the Cauchy-Schwarz inequality yields that \( \mu \) is strictly positive if and only if \( U \) is not colinear to one, as in the previous case. The parameter \( \mu \) has a simple geometric interpretation as the (scaled) angle between \( U \) (compensation) and \( \mathbb{I} \) (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 \( \mu \) 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: \( \mathbb{I}'\partial \lambda / \partial U = 1, \mathbb{I}'\partial \mu / \partial U = 0 \) and \( U'\partial \mu / \partial U > 0 \). 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 \( \Delta \) and the multipliers are unique (and thus the first-order approach is valid).

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

\[ \delta_i = \frac{(u(w_i) - \lambda / \mu - x_i)}{D_{i,i}} \quad \text{(A-8)} \]

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

One can also verify that \( \Delta \) 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 Proposition 2.1:** Claim 1: Problem (A) has a solution. The result is obvious when \( N \) is finite. Suppose \( N \) is not finite. Let \( (\delta^a_k)_{k=1}^\infty \) be a sequence of feasible actions converging to an optimum. Define
\(a^\infty = \lim \inf a^n\) and, for all \(k\), \(\delta_k^\infty = \lim \inf \delta_k^n\). I shall show that \(S\) is feasible for the agent.

\[
\sum_{n=1}^{\infty} (\delta_k^\infty x_k + C(a^\infty, x_k, \delta_k^\infty)) = \sum_{n=1}^{\infty} \left(\lim \inf \delta_k^n x_k + C(\lim \inf a^k, x_k, \lim \inf \delta_k^n)\right)
= \sum_{n=1}^{\infty} \left(\lim \inf (\delta_k^n x_k + C(a^k, x_k, \delta_k^n))\right)
\leq \lim \inf \sum_{n=1}^{\infty} (\delta_k^n x_k + C(a^k, x_k, \delta_k^n)) \quad \text{(Fatou’s Lemma)}
\leq 0
\]

Then, \((\delta_k^\infty)_{k=1}^{\infty}\) satisfies Equation (4). 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 \(k' \leq k\).

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

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 (3). Thus the action \(S\) is feasible.

Finally, one needs to show that \(S\) is utility-maximizing for the agent. To see this, note that \(\delta_k^\infty x_k + (1-\delta_k^\infty)\theta \leq \pi^r\), and therefore: \(\delta_k^n u(w(x_k)) \leq (\pi - \theta)/(x_k - \theta) u(w(x_k))\). 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:

\[
\lim \inf \sum_{k=1}^{\infty} \delta_k^n u(w(x_k)) = \sum_{k=1}^{\infty} \delta_k^\infty u(w(x_k))
\]

Claim 2: \(\mu > 0\) Suppose not. The first-order condition in \(\delta_k\) yields that \(u(w(x_k)) = \lambda\) which is a contradiction to \(\sum_{k=1}^{N} u(w(x_k))p_k'(0) > 0\). \(\square\)

**Proof of Corollary 2.2:** By Equation (3),

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

By Equation (4),

\[
\frac{1}{\epsilon(a)} \sum_{k=1}^{n} \left( \frac{u(w(x_k))}{\mu} - x_k \right) x_k = -\sum_{k=1}^{n} \frac{\epsilon(a)}{2} \left( \frac{u(w(x_k)) - \lambda}{\mu} - x_k \right)^2
\]

34
\[
0 = \sum_{k=1}^{n} \left( \frac{u(w(x_k)) - \lambda}{\mu} - x_k \right) \left( \frac{u(w(x_k)) - \lambda}{\mu} + x_k \right)
\]
\[
= \sum_{k=1}^{n} \left( \frac{u(w(x_k)) - \lambda}{\mu} \right)^2 - x_k^2
\]
\[
= \sum_{k=1}^{n} u(w(x_k))^2 + \lambda^2 N - 2\lambda \mu \sum_{k=1}^{n} u(w(x_k)) - \sum_{k=1}^{n} x_k^2
\]

Solving for \( \mu \) yields the following polynomial:
\[
\mu^2 \left( \frac{1}{N} \sum_{k=1}^{N} x_k \right)^2 + \frac{1}{N} \sum_{k=1}^{N} x_k^2 - \frac{1}{N} \sum_{k=1}^{N} u(w(x_k))^2 + \frac{1}{N} \sum_{k=1}^{N} u(w(x_k))/N = 0
\]

This Equation has a unique positive real root.
\[
\mu = \sqrt{\frac{\frac{1}{N} \sum_{k=1}^{N} s_k}{\frac{1}{N} \sum_{k=1}^{N} x_k^2}}
\]

Reinjecting yields the expression for \( \lambda \). \( \square \)

**Proof of Proposition 2.3:** (i) If \( a \) 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. \( w(y) = w^{-1}(\lambda + \mu y) \) 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 \( (\lambda, \mu) \) 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 \( a = a^* \) 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:

\[
\sum_{k=1}^{N} p_k(a)(x_k - a) = \sum_{x_k \leq a} p_k(a)(x_k - a) + \sum_{x_k > a} p_k(a)(x_k - a) \\
\leq \sum_{x_k \leq a} p_k(a)(x_k - a) + \sum_{x_k > a} p_k(a)(x_k - a) \\
\leq 0
\]

It follows that \( M_a(a) > 0. \□ \)

**Proof of Proposition 3.2:** 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 \( C(\delta, a, \delta_k)/j \) also converges to zero. And by Equation (7), 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. \( \square \)

**Proof of Corollary 3.1:** 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 (17) and (18) in \( x \),

\[
\frac{\partial w}{\partial x} = (1 + \beta - \frac{\partial}{\partial x} C_{\delta, a, x} - \beta C_{\delta, x} - \mu \gamma C_{\delta, a, x} - \mu \gamma C_{\delta, x} \frac{\partial \delta}{\partial x} - \mu (C_{\delta, a, x} + C_{\delta, x}) ((d + p)/(1 + u(w) - \tau) - \gamma p LR) \nonumber \\
- \mu \gamma C_{\delta, a, x} (\frac{\partial}{\partial x} (1/u(w)) - \tau + (d + p) (-\frac{\partial w}{\partial x} u''(w)/u'(w)^2))) = A-9
\]

Solving these Equations in \( \partial \delta / \partial x \):

\[
\frac{\partial \delta}{\partial x} = \frac{(-1 + \beta + \beta C_{\delta, a, x}) u'(w)^2 + \mu u'(w)^2 (1 + C_{\delta, a, x} + C_{\delta, a, x} \gamma u'(w) + C_{\delta, a, x} (d + p - (\gamma p LR + (d + p) \tau) u'(w))) - C_{\delta, a, x} \mu (d + p) u'(w))}{u'(w)^2 (\mu (-C_{\delta, a, x} (d + p) - C_{\delta, a, x} \gamma u'(w) + C_{\delta, a, x} (d + p) \gamma LR + (d + p) \tau) u'(w)) + C_{\delta, a, x} (-\mu u''(w) + \mu (-2 + u(w) + u') + C_{\delta, a, x} \mu u''(w) + C_{\delta, a, x} \mu u'(w)))} = A-10
\]

Setting \( u'(w) = 0 \),
\[
\frac{\partial \delta}{\partial x} = -\frac{1}{C_{\delta, \delta}} \tag{A-11}
\]

This term is negative under the conditions of Corollary 3.1.

(ii) For \(\partial \delta/\partial LR\), differentiating Equations (17) and (18) in \(p\),

\[
\frac{\partial w}{\partial LR} = \frac{\mu C_{\delta, \delta}}{u'(w)} \tag{A-12}
\]

\[
\frac{\partial w}{\partial LR} = -\beta \frac{\partial w}{\partial LR} C_{\delta, \delta} - \mu C_{\delta, \delta, a} \frac{\partial \delta}{\partial LR} - \mu C_{\delta, \delta, \delta} \frac{\partial \delta}{\partial LR} \left(\frac{1}{u'(w)} - \gamma_p \right) - \gamma_p \left(\frac{\partial \delta}{\partial LR} \right) \left(\frac{1}{u'(w)u'(w)} \right) \tag{A-13}
\]

Solving these Equations in \(\partial \delta/\partial LR\):

\[
\frac{\partial w}{\partial p} = \frac{C_{\delta, \delta} \mu}{(\mu(-C_{\delta, \delta} \delta + p - C_{\delta, \delta} \alpha \gamma u'(w) + C_{\delta, \delta, \delta} (\gamma \delta + (\delta + p)x'w'(w))) + C_{\delta, \delta} (-\beta u'(w) + m(-2 + \tau u'(w)))u'(w)^2 + C_{\delta, \delta, \delta}^2 \mu^2 \delta + p u''(w))} \tag{A-14}
\]

This term is positive under the conditions of Corollary 3.1.

(iii) For \(\partial \delta/\partial p\), differentiating Equations (17) and (18) in \(LR\),

\[
\frac{\partial w}{\partial p} = \frac{\mu C_{\delta, \delta}}{u'(w)} \tag{A-15}
\]

\[
\frac{\partial w}{\partial p} = -\beta \frac{\partial w}{\partial p} C_{\delta, \delta} - \mu C_{\delta, \delta, a} \frac{\partial \delta}{\partial p} - \mu C_{\delta, \delta, \delta} \frac{\partial \delta}{\partial p} \left(\frac{1}{u'(w)} - \gamma_p \right) - \gamma_p \left(\frac{\partial \delta}{\partial p} \right) \left(\frac{1}{u'(w)u'(w)} \right) \tag{A-16}
\]

Solving these Equations in \(\partial \delta/\partial p\):

\[
\frac{\partial w}{\partial p} = \frac{C_{\delta, \delta} \mu (-1 + \gamma L \alpha e'(w) + \tau u'(w))u'(w)^2}{(\mu(-C_{\delta, \delta} \delta + p - C_{\delta, \delta} \alpha \gamma u'(w) + C_{\delta, \delta, \delta} (\gamma \delta + (\delta + p)x'w'(w))) + C_{\delta, \delta} (-\beta u'(w) + m(-2 + \tau u'(w)))u'(w)^2 + C_{\delta, \delta, \delta}^2 \mu^2 \delta + p u''(w))} \tag{A-17}
\]

This term is positive or negative under the conditions of Corollary 3.1. \(\square\)

**Proof of Proposition 4.2:** The first-order condition with respect to \(w(y)\) yields that:

\[
\frac{v'(y - w(y))}{u'(w(y))} = \tilde{\tau} \tag{A-18}
\]
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:

\[
v(y - w(y)) - \tilde{\alpha} - \tilde{\beta}(y + C'(\delta(y))) + \tilde{\tau}u(w(y)) = 0
\]

(Equation A-19)

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 \( y \),

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

Rearranging this expression:

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

\[
= \frac{\tilde{\tau}}{w'(y)} - \frac{\tilde{\beta}}{w'(y)u'(w(y))}
\]

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

**Proof of Proposition 4.3:** The first-order condition with respect to \( \mu \),

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

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

Under perfect risk-sharing, \( \tau \) is equal to \( v'(y - w(y))/u'(w(y)) \) for all \( y \). Observe first that the agent will always select \( a > a \) 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:

\[
u(w(y)) = \lambda + \mu y + \mu C'(\delta(y))
\]

\[
\int f_a(y|a) u(w(y)) dy = \mu - \mu = 0
\]
This would imply that the first-order with respect to effort in the problem of the agent select \( a = a \), a contradiction to \( a > a \). 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 (15) that \( v(y - w(y)) \) is constant, a contradiction to perfect risk-sharing. \( \square \)

**Proof of Proposition 4.4:** 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 yields 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:\n
\[
\frac{\partial v'(y - w(y))/u'(w(y))}{\partial y} \geq 0.
\]

\( \square \)

**Proof of Proposition 4.1:** Develop the covariance as follows:

\[
\text{cov}(y', \eta^t) = E(y'y') - E(y')E(\eta^t)
\]

\[
= \int y f(y)dy - E(y')E(\eta^t)
\]

\[
= \text{cov}(x', x') - (E(x')E(x') - E(y')E(\eta^t)) + R
\]

But the analogue of Constraint (4) in the multivariate case implies that:

\[
R = -\int C(k)[\frac{u'(w(y)) - \lambda - \mu \eta}{\mu}]dy
\]

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

**Bibliography**


