The Dissertation Committee for Jeremy Bertomeu

certifies that this is the approved version of the following dissertation:

Risk Management and the Design of Efficient Incentive Schemes

Committee:

Stephen E. Spear, Supervisor

Richard C. Green

Pierre J. Liang

Christopher Sleet

Jonathan Glover (external reader)

Risk Management and the Design of Efficient Incentive Schemes

by

Jeremy Bertomeu, M.S.

Dissertation

Presented to the Faculty of the Graduate School of

Carnegie Mellon University

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

Carnegie Mellon University

May 2008

...

Acknowledgments

I thank my wife Edwige Cheynel for her support and dedication through the various stages of my dissertation. This research was funded in part by the William Larimer fellowship (2003-2006) and the Center for Analytical Research in Technology (2007-2008). I thank my chair Steve Spear for his valuable advice, and my committee members Jonathan Glover, Rick Green, Pierre Liang and Chris Sleet. Feedback is acknowledged from Nuray Akin, Bruno Biais, Arch Chakraborty, Gilles Chemla, Jeff Coles, Sean Crockett, Ron Dye, Helyette Geman, Michel Gendron, Burton Hollifield, Dave Kelly, Aziz Lookman, Robert Marquez, David Martimort, Spencer Martin, Bryan Mittendorf, Richard Phillips, Guillaume Plantin, Jean-Charles Rochet, Bill Rogerson, Bryan Routledge, Duane Seppi, Sri Sridhar, Ajay Subramanian, Shyam Sunder, Jean Tirole, Sevin Yeltekin and other seminar participants at Arizona State, Baruch College, CMU, Georgia State, Imperial College, Laval, University of London, Northwestern, University of Miami, Stony Brook Game Theory Conference, Yale and Toulouse. Many conversations with fellow PhD students Fernando Anjos, Vincent Glode, Vineet Kumar TJ Liu and Richard Lowery were beneficial to my work. Finally, I would like to thank our program coordinator, Lawrence Rapp, for his help.

JEREMY BERTOMEU

Carnegie Mellon University

May 2008

Risk Management and the Design of Efficient Incentive Schemes

Publication No. ____________________

Jeremy Bertomeu, Ph.D.
Carnegie Mellon University, 2008

Supervisor: Stephen E. Spear

Models of incentives generally place strong restrictions on how the actions of the agent affect output. In particular, it is typically assumed that effort increases the distribution of output in the first-order stochastic sense (e.g., monotone likelihood ratio property). In practice, however, managers can also privately affect the riskiness of their output. For example, a CEO may invest in activities with known or uncertain returns, a fund manager may purchase hedging instruments or a lawyer may decide to settle or go to trial. Such problems are not well-captured in the standard framework.

My dissertation attempts to bring together two separate areas of the literature: agency theory and risk management. In the former, the risk of the distribution is typically restricted by tight conditions on the set of distributions available to the agent. In the latter, the informational frictions that make risk management necessary are modeled as exogenous frictions. My dissertation reveals that giving incentives to work hard may conflict with giving incentives to manage risk. I discuss how risk management will affect the nature of the optimal contract and the types of risk the managers will take. Then, I derive implications for executive compensation practices and asset prices. In particular, I answer the following questions:

(i) Why should managers manage risk given well-diversified (risk-neutral) owners?

(ii) What patterns on the cross-section of corporate earnings are indicative of risk management?

(iii) What types of contracts perform well in aligning incentives to manage risk and incentives to provide effort?

(iv) Why should optimal contracts reward managers for market movements that they do not control and may hedge on their own?

The first chapter of my dissertation 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. The model parsimoniously reproduces three stylized facts commonly described as anomalous. First, executive compensation contracts are convex with option-like linear components. Second, the cross-section of corporate earnings is hump-shaped with asymmetric tails even though the production technology does not seem to exhibit such features. Third, the response of stock returns to current earnings is S-Shaped. Empirical implications for the detection of earnings management are also examined.

In Chapter I, the choices of the agent over the distributions are penalized, in the sense that changing the distribution requires to decrease expected output. My second chapter extends the results of Chapter I to the case in which the cost for the agent to change the distribution of output is zero. In this case, I show that managers are given incentives to manage risk, removing any risk that can be hedged. When managers can take large (but bounded) gambles, eliciting effort requires to give a rent to the manager, even if the output signal is very informative on the actions of the agent. Finally, I discuss how risk controls can affect the incentive problems. When managers can take arbitrarily large gambles, only the minimum effort can be elicited and reducing the maximum loss feasible by the manager increases the payoffs to the principal. Under certain conditions, a value-at-risk constraint increases the profit of the principal.

In Chapter III, I rationalize the dependence of individual pay on systematic shocks in a general equilibrium economy with moral hazard, in which risk-averse agents face a common productivity shock. I show that the optimal contract prescribes higher total pay and pay-for-performance coefficient during expansions than recessions. Further, in the HARA class of utility functions, risk premia do not depend on the informational asymmetry. If the agent may trade outside of the contract, a relative performance contract (i.e., which filters out systematic risk) is optimal if and only if the agent has constant absolute risk-aversion. The idea that firms should better "insure" agents against aggregate shocks is, however, disproved: any zero-value change to the exposure of the agent to the systematic shock would be fully unraveled by the agent's trading decisions and would have no consequences on effort or consumption.

The last chapter of my dissertation presents an application of the results to fund management and studies the informational determinants of arbitrage capital flows. In the model, outside investors

have imperfect information about the potential risk taken by arbitrageurs, which creates an informational asymmetry to be resolved. I explore the relationship between primitive characteristics of the arbitrage (such as execution cost, potential risk and arbitrage return) and capital flows. Risk and cost decrease the amount of capital received by arbitrageurs, as expected. Counter-intuitively, however, high-return arbitrages may attract less capital than low-return ones. This situation may create multiple competitive equilibria: one of them price efficient, and others with prices far away from fundamentals. Implications for asset price instability, managerial compensation (high-water marks) and hedge fund capital flows are also discussed.

Contents

List of Figures

Chapter I

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

1 Introduction

Risk management practices have come under new scrutiny by regulators and corporate boards. 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. 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). Many observers note that managers often use risk management for reporting purposes and risk exposure is neither fully transparent nor well-controlled by shareholders.

Regulators are designing a new conceptual framework to understand risk management and "discourage transactions and transaction structures primarily motivated by accounting and reporting concerns rather than economics" (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 Issuers - June 2005). A widespread idea is to formulate the basic trade-off in the following terms: risk management can improve economic risk-sharing but can be distorted by managers for pure reporting reasons.

To better understand what is meant by "accounting and reporting concerns" and discuss the cost and benefits of risk management, several key questions are of interest.

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?

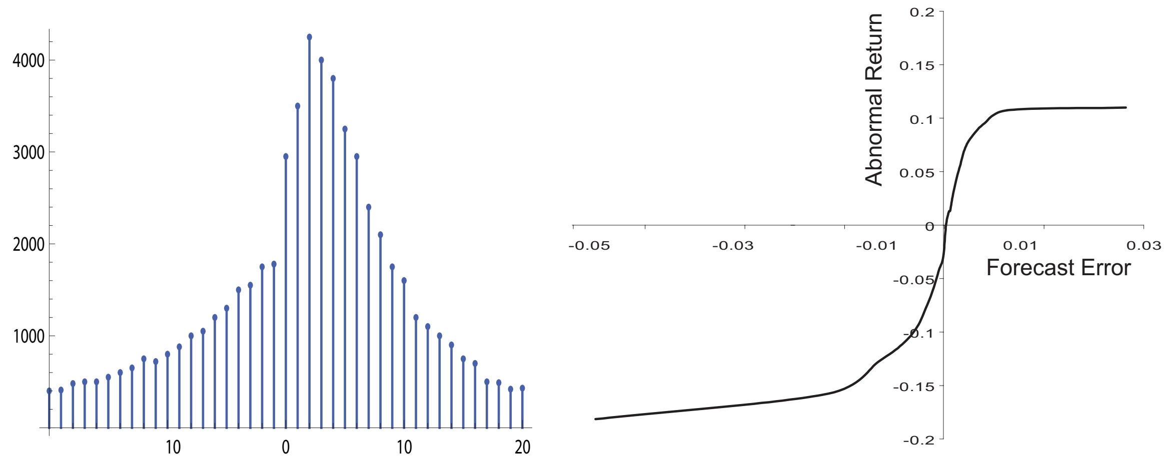
III. How can one design executive compensation contracts that give incentives to hedge in the best interest of shareholders? How does risk management affect performance-pay coefficients? What type of contracts will perform well in situations where risk management is important?

IV. What anomalous features of the cross-section and time series of earnings are consistent with risk management? How can risk management be detected?

This chapter presents a framework in which the agent manipulates earnings by privately managing risk. The risk decision is imperfectly observed by the owners of the firm which creates an informational asymmetry between owners and managers.[1] 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.).

[1] The approach of this chapter is with hidden actions by the manager. Other papers (e.g., 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.



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

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

Figure I.1. Empirical Evidence

My analysis is related to an extensive literature on risk management and project selection. Diamond (1998) describes the problem of an agent who can take zero-value gambles, and shows that the optimal contract becomes linear when the size of the firm is large. Biais and Casamatta (1999) consider an agent with access to one additional risky "bad" project. Palomino and Prat (2003) and Parlour, Purnanandam and Rajan (2006) frame the problem of fund managers who can choose among a set of distributions. The main modeling difference between these papers and this one is that the manager can choose any project given a decrease expected earnings. I do not restrict the attention to only two projects (as in Biais and Casamatta (1999)) and obtain new predictions on the optimal contract. Finally, Hirschleifer and Suh (1992) and Meth (1996) discuss problems in which the manager chooses the variance of a Normally distributed distribution and derive conditions under which a linear or convex contract is optimal. I extend their approach to the case in which risk management is over distributions that are not necessarily Normal and such that not only the variance can be altered.

Three Stylized facts

The model further leads to some new insights into the forms of earning managements made possible by the control of managers over risk.[2] I describe next several stylized facts commonly viewed as puzzling or unexplained from the perspective of the standard theory.

Many observed compensation schemes seem to be structured in ways that induce the manager to strategically manage their earnings. Most executive pay packages feature a combination of equity

[2]Among academics, it is often argued that some earnings management may be beneficial to incentives (Arya, Glover and Sunder (2003)) - a key aspect of the analysis. However, in other theories, earnings management may be fully unraveled by investors (Arya, Glover and Sunder (1998), Dechow and Skinner (2000)).

compensation (stock and options) and targeted bonus payments. The predominance of such compensation schemes is surprising for two reasons. One, excess convexity (with options) and discontinuous compensation (with bonuses) may induce the manager to manage earnings to beat performance threshold (Murphy (1999) - p.15-18) and/or take risky projects (Rajgopal and Shevlin (2002), Coles, Naveen and Naveen (2006)).[3] Two, the widespread linearity of the compensation is hard to reproduce in standard agency problems (see Haubrich and Popova (1998) and Figure 5. 2).[4] The objective of this chapter is to help understand when some linear and/or option-like components can be optimal in the compensation, and whether it is desirable to tolerate some earnings management.

Perhaps as a direct consequence of such compensation practices, it has often been observed that the cross-section of corporate earnings features a *hump-shape* or *divot* (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 I.1 and reveal 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. Specifically, 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).

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 Figure I.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). While several theoretical rationales have been proposed, existing theories are unable to jointly account for the S-shape, the hump-shape observed in the cross-section and the form of executive compensation contracts.[5] This chapter attempts to reconcile these three facts in a parsimonious common

[3]There exists considerable evidence that managers manipulate earnings to attain bonuses (see for example Healy (1985) as well as a rich follow-up literature such as Matsumoto (2002)) - note that here, in contrast, the manager manipulates the risk of the project.

[4]To 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). In Innes (1990), linearity is obtained by assuming that a limited liability by the principal binds - an unlikely assumption for the analysis of contracts between large corporations and their CEOs.

[5]Guttman, Kadan and Kandel (2006) present a model in which managers can misreport earnings for a cost, and show that

theory, showing how the need to provide incentives can determine the optimal compensation, as well as features of the cross-section and time series of earnings.

2 Binomial Information Structure

2. 1 The Model

I present first a simplified version of the model. A firm is owned by a risk-neutral principal (or owner) and operated for a single period by an agent (or manager). Risk neutrality is assumed 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). To keep the model simple, I assume for now that the firm is liquidated after this period ends and yields access to a project with earnings $x \in [0, 1]$.

□ **Production Technology**

The manager privately chooses an action $a \in \{0, 1\}$ that can affect the distribution of x. Effort $a = 1$ represents high effort and, in this case, x is drawn from $\mathcal{U}[0, 1]$. When $a = 0$, x is drawn from $\mathcal{U}[0, .5]$ with probability $\theta > .5$ and drawn from $\mathcal{U}[.5, 1]$ with probability $1 - \theta$. Under this parametrization, the information contained in x is binomial: outcomes below .5 are indicative of low effort while those above .5 are indicative of high effort. Effort $a = 1$ requires a cost $B > 0$ to the agent while effort $a = 0$ is costless.

□ **Risk Management**

The agent can implement x, or another project with earnings y but different risk characteristics. The density of y is denoted $\hat{f}(.)$. Assume that x is always the most preferred project from the perspective of value-maximization, and thus choosing y requires to reduce expected earnings.

To pursue further, I borrow from non-parametric estimation the idea of using the Integrated Squared Error or ISE (Pagan and Ullah (1999), p. 24) as a metric to measure the distance between project x and project y. I assume in this essay that this metric also captures the economic distance between the two projects and, as such, how much it is necessary to decrease earnings to implement y

this game has partially pooling equilibria with reporting kinks. Crocker and Huddart (2006) propose a model in which the manager is privately informed about value and can manipulate a report. Li (2008) argues that investors are uncertain about the mean *and* variance of future earnings, thus leading to S-shaped updates.

instead of x. Using the ISE, the average earning under y is the average earning under x (1 with effort) minus the distance between the two projects $c\int(\hat{f}(y)-f(y|a))^2dy/2$ where $c>0$ is a scaling variable. In other words, I take the standard ISE metric distance between two distributions as a measure of how much the agent needs to alter the project (reducing earnings) to produce y instead of x.

The main idea of this essay is that the set of projects potentially available to a manager in a firm is large and thus, in the context of this model, I assume that the agent can implement any such y project. Specifically, for a given effort, the agent can privately implement any project y such that $\mathbb{E}(y)=\mathbb{E}(x)-c\int(\hat{f}(y)-f(y|a))^2dy/2$. Since any $\hat{f}(.)\neq f(.|a)$ strictly decreases earnings and alters the information received by the principal, I view choosing y as a form of costly earnings management - although via risk and not ex-post earning manipulation or misreporting. In the model, the assumption represents operational risk management and corresponds to an abstract view of many choices available to the manager: (i) producing for a cost additional information to make earnings more precise on actions, (ii) distorting the selection of projects within the firm, or (iii) investing to alter the production technology.

□ **Problem of the Agent**

The agent chooses effort a and which project to implement $\hat{f}(.)$ (possibly $\hat{f}(.)=f(.|a)$, i.e. no risk management) to maximize his expected utility. Define $\delta(y)=\hat{f}(y)-f(y|a)$. Taking the wage $w(.)$ as given, the problem of the agent is given as follows:

$$(P_A)\quad (a,\delta(.))\quad \in \arg\max_{\tilde{a},\tilde{\delta}(.)\geq -f(y|\tilde{a})}\int(f(y|\tilde{a})+\tilde{\delta}(y))w(y)-\tilde{a}B$$

$$\text{s.t.}\qquad \int\tilde{\delta}(y)dy=0 \qquad (\lambda) \qquad \text{(I.1)}$$

$$\int y\tilde{\delta}(y)dy=-c\int\tilde{\delta}(y)^2dy/2 \qquad (\mu) \qquad \text{(I.2)}$$

In (P_A), the agent chooses $\delta(y)$ (or, equivalently, $\hat{f}(.)$). Equation (I.1) is a feasibility constraint that requires $\hat{f}(y)$ to be a valid probability distribution satisfying $\int\hat{f}(y)dy=1$. Equation (I.2) corresponds to $\mathbb{E}(y)=\mathbb{E}(x)-c\int(\hat{f}(y)-f(y|a))^2dy/2$ rewritten after substituting $\delta(y)$. To avoid corner solutions, c is chosen sufficiently large (greater than 10) so that the positivity constraint $\delta(y)\geq -f(y|a)$ does not bind in (P_A). To avoid cases in which incentive problems are mild, assume in addition that $B>.01$.

□ **Problem of the Principal**

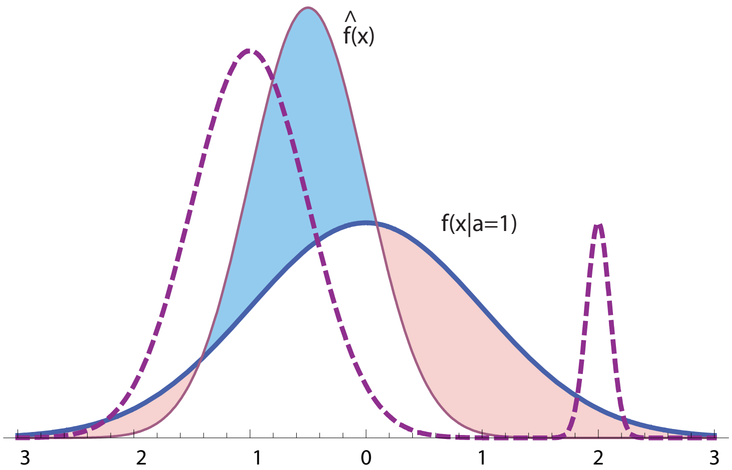


Figure I.2. Examples

The agent has limited liability and must be paid $w(.) \geq 0$; further, in this Section, I restrict the attention to non-decreasing wage schedules.[6] Suppose that effort $a = 0$ will lead to a large non-contractible loss in the future and should never be elicited. The principal chooses $(w(.), \delta(.))$ to maximize $V \equiv \int (f(y|1) + \delta(y))(y - w(y))dy$ subject to $(1, \delta(.))$ being a solution to (P_A).

2. 2 Discussion

In this model, I interpret managing risk as moving the distribution from $f(.|a)$ to $\hat{f}(.)$. As the agent manages risk more, in that the distribution becomes more distinct from the original distribution, the loss in value $c \int (\hat{f}(y) - f(y|a))^2 dy/2$ becomes more important. 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.[7]

A feature of this approach is that it builds on some of the standard metric used in statistics to measure the distance between distributions. Further, the ISE metric does not restrict the shape of the distributions $\hat{f}(.)$ feasible by the agent and thus it is equipped to capture anomalies such as the hump-shape in the cross-section of corporate earnings. From a theoretical perspective, the assumption implies that any shape for the distribution of y can be realized for a cost (i.e., by reducing mean earnings sufficiently), while standard approaches constrain the manager to a family of specified distributions that does not fully capture the large choice set over projects that one should observe in practice.

[6] For example, the manager may costlessly reduce earnings ex-post.

[7] An important aspect of the definition should be emphasized. The assumption 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 (see Morellec and Smith (2007) for a model in which the principal manages risk).

To give some more economic intuition on what the assumption can capture, I give next a simple illustrative example. Suppose that an airline company is considering whether or not to hedge some of its fuel expenses, by writing a contract with an intermediary that pays for some of its expenses. In the left-hand side of Figure I.2, I represent the profit achieved by the airline company if it does not hedge. It depends on the realization of oil prices and is assumed to be approximately normal. The airline company can also write a contract such that a proportion of its fuel expenses will be insured by a bank. This leads to a less volatile distribution $\hat{f}(.)$. As is well-known, however, a fairly-priced insurance can only be given if there is perfect information between the firm and the intermediary and thus, in general, the airline company may have to forfeit some expected profit to buy this insurance.[8] In this model, this informational cost is captured as the shaded area between $\hat{f}(.)$ and $f(|a)$. Finally, suppose that the airline manager has incentives to beat competitors and, to do so, over-insures fuel expenses. Simultaneously, he needs to dissimulate these transactions, or else investors would categorize these profits as pure financial transactions and not operations (defeating their purpose). This dissimulation should have a cost (e.g., auditor fees, financial engineering, litigations), further reducing expected profits. However, it creates an asymmetric distribution that features a higher likelihood of a large gain (dashed curve).

The simplified model is useful to deliver the main intuitions of the framework. Most of the assumptions, however, are relaxed in the next sections. In particular, many of the results can be extended to a less parameterized incentive problem. In the general model, I extend the approach to risk-aversion by the agent, arbitrary distributions and a non-quadratic loss function.

2. 3 Hump-Shape in the Cross-Section of Earnings

I solve first the problem of the agent (P_A). In this problem, it is convenient to rewrite the incentive-compatibility constraint for $a = 1$ as follows:

$$\int_{.5}^{1} w(y)dy - \frac{B}{2\theta - 1} \geq \int_{0}^{.5} w(y)dy \tag{I.3}$$

Following Equation (I.3), the manager must be paid more in the states indicative of high effort than those that are more likely under low effort. In the rest of this Section, I assume that the wage $w(y)$ offered to the agent satisfies Equation (I.3). In Problem (P_A), the first-order condition with respect to $\delta(y)$ is given by:

[8] Many thanks to Tom Lys for suggesting the following example: such a contract may make the airline company overconsume oil when prices are high and under-consume when prices are low (since it does not fully internalize prices after this insurance has been taken). One may of course argue that not all contracts feature a dependence on the operations of the firm; however, many do, including bank financing contracts and structured financing vehicles.

$$w(y) - \lambda - \mu y = c\mu\delta(y) \tag{I.4}$$

On the left-hand side of Equation (I.4), the wage $w(y)$ is compared to a linear threshold $\lambda + \mu y$ (to be recovered endogenously). When the agent is paid more that this threshold, $\delta(y) > 0$ and the agent increases the likelihood of outcome y.

There are two main economic intuitions at play here. First, given greater pay for an outcome, the agent is willing to raise the probability of this outcome even at the cost of reducing expected cash earnings. Second, for a given compensation, the agent prefers the reduce the likelihood of large gains: intuitively, if two outcomes $y' > y$ yield similar pay $w(y') = w(y)$, then reducing the likelihood of y' reduces mean earnings more and thus gives the manager more freedom to increase the likelihood of other outcomes. Taken together, these two observations imply that the manager is willing to increase the likelihood of outcomes with a large pay and decrease the likelihood of outcomes with high earnings.

In the next Proposition, I derive the threshold and the risk management decisions for an arbitrary compensation scheme $w(y)$.

Proposition 2.1. *For a given non-constant $w(.)$ such that $a = 1$ is elicited, $(\lambda, \mu, \delta(y))$ is unique and given as follows:*

$$\lambda = \mathbb{E}(w(x)) - 2\sqrt{3}\sigma(w(x)) \tag{I.5}$$

$$\mu = 2\sqrt{3}\sigma(w(x)) \tag{I.6}$$

$$c\delta(x) = \frac{(w(x) - \mathbb{E}(w(x))}{2\sqrt{3}\sigma(w(x))} - (x - .5) \tag{I.7}$$

where $\sigma(.)$is the standard deviation.

From Equations (I.5) and (I.6), $\lambda + \mu\mathbb{E}(x)$ is equal to $\mathbb{E}(w(x))$: there is no risk management at the mean outcome only if the agent is paid his expected compensation when not managing risk. As a result, any compensation that features above-average compensation at the average outcome will induce the agent to increase the likelihood of this outcome.[9] In addition, the no-hedging threshold is equal to λ at $y = 0$, and $\lambda + \mu$ at $y = 1$, i.e. if the agent is paid below $2\sqrt{3}$ standard deviations of average effort at the minimum outcome, he will reduce its likelihood (and vice-versa). The analysis thus yields additional "benchmark" bounds on risk management: the manager will increase (resp. decrease) the likelihood of

[9]Note that here that the mean and standard deviation used in Proposition are taken with respect to the original distribution x, not y. However, the difference should not be too large if c is large also.

any outcome that prescribes a compensation above (resp. below) approximately 3.5 standard deviations from the mean.

Finally, Equation (I.7) shows that the optimal amount of risk management can be derived by comparing the (normalized) "abnormal" compensation $w(y) - \mathbb{E}(w(y)))$ to the "abnormal" output $y - .5$. In simple terms, an agent paid given a negative abnormal compensation for a positive abnormal output would manage risk to reduce the likelihood of this event. More specifically, note from Equation (I.7) that the abnormal wage must be strictly greater (in absolute value) to the abnormal output - due to the term $.5\sqrt{3}$ - for $\delta(y)$ to have the sign of the abnormal wage. Intuitively, the abnormal output is given more weight than the wage in the determination of $\delta(y)$, essentially capturing the control of the agent over the production technology.

Bonus Contract

I describe next, in graphical terms, how the linear threshold (and thus $\delta(y)$) can be derived for simple wage schedules. Consider a simple bonus contract: $w_b(y) \equiv 1_{y \geq .5} e$ where $e > 0$ is a performance bonus given when the agent reports high outcomes. In the left-hand side of Figure I.3, the contract $w_b(y)$ is represented in dotted lines and the threshold $\lambda + \mu y$ is represented in bold.[10] Comparing the linear threshold to the compensation yields the following risk management decisions:

Proposition 2.2. *If a bonus contract $w_b(.)$ is offered and elicits $a = 1$, there exists $\theta \in (0, .5)$ such that:*

(i) If $y \in [0, .5 - \theta) \cup [.5, .5 + \theta)$ (i.e., low outcomes or "beating" the bonus threshold), $\delta(y) > 0$.

(ii) If $y \in [.5 - \theta, .5) \cup [.5 + \theta, 1)$ (i.e., high outcomes or "falling short" of the bonus threshold), $\delta(y) < 0$.

The agent responds to a bonus contract by trying to beat the bonus, shifting probability mass from outcomes in which y falls short of target to outcomes in which the bonus target is attained or slightly beaten. To gain as much as leverage as possible to increase the latter, the manager also increases the likelihood that the bonus is met by increasing the likelihood of large loss and reducing the likelihood of a large gain. In the right-hand side of Figure I.3, I represent how the distribution of x is altered, producing a hump-shape located at .5 indicative of the hump-shape in the cross-section of earnings observed empirically.

[10] The location of the threshold can be easily obtained from Proposition 2.1. The threshold $\lambda + \mu y$ must intersect at the expected compensation, i.e. at the point A. Suppose, by contradiction, that $\lambda + \mu y$ does not intersect $w_b(y)$. In this case, the agent would be increasing the likelihood of outcomes with $y \geq .5$, and decrease the likelihood of outcomes with $y < .5$. This would contradict $\mathbb{E}(y) < \mathbb{E}(x)$ and thus is not feasible.

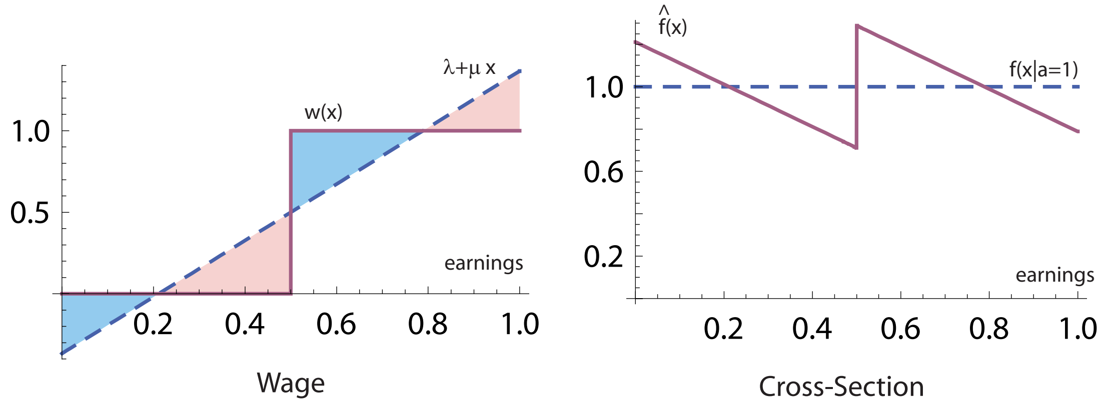


Figure I.3. Response to a Bonus

2. 4 Linearity of the Optimal Contract

The principal elicits effort $a = 1$, taking into account incentive-compatibility and the incentives to manage risk (and decrease value) that the wage schedule may produce.

Lemma 2.1. *The optimal compensation is linear by parts, and takes the following form:* $w(y) = \max(0, a_0 + a_1 y + 1_{y \geq .5} e)$*, where* $a_1 \geq 0$ *and* $e \geq 0$*.*

To interpret the result of Lemma 2.1, consider first the standard model in which the agent cannot manage risk. In this case, outcomes above .5 are indicative of high effort. Therefore, any compensation such that $w(y) = 0$ for $y < .5$ and $\mathbb{E}(w(y)|y \geq 0.5) = 2B/(2\theta - 1)$ is optimal. In other words, the shape of the contract over outcomes greater than .5 is irrelevant.

In comparison, I show that, in this model, the trade-off between incentives and the cost of risk management makes it desirable to offer a compensation that is linear in parts. While all outcomes above .5 are equally informative on effort, non-linearities induce the agent to manage risk and reduce value. To offset these problems, the principal offers a compensation that is linear for a given level of information on effort. In the next Proposition, I characterize how the principal can best exploit differences of informativeness between low and high realizations of earnings.

Proposition 2.3. *The optimal contract takes the following form:*

(i) Simple Bonus: $w(y) = \frac{2B}{2\theta - 1} 1_{y \geq .5}$.

(ii) Option with a bonus above the strike: $w(y) = a_1 \max(0, y - y_0) + 1_{y \geq .5} e$*, where* $y_0 \in (0, .5]$ *and* $e > 0$*.*

The optimal contract is kinked and exhibits aspects such as a bonus with possibly an option-like component. The bonus payment helps increase pay for outcomes that indicate high effort. However, it also induces some value-decreasing risk management, which is detrimental to incentives and total value created. To mitigate this problem, it may be optimal to offer an option component which brings the contract closer to linearity. In this case, the strike of the option is always weakly below the bonus threshold (case (ii)), and thus the option generates some payments for low outcomes: the linearity of the option component mitigates incentives to beat the bonus threshold.[11]

Welfare Implications

I analyze next the welfare properties of the model when the principal is constrained to a simple bonus contract. In this problem, denote V the expected utility of the principal, V_A the expected utility of the agent and L the reduction in value due to risk management.

Corollary 2.1. *Suppose the simple bonus contract (case (i)) is used,*

$$V = \frac{1}{2} - \frac{B}{2\theta - 1}(1 - \frac{3 - 2\sqrt{3}}{12c}) - \frac{2 - \sqrt{3}}{24c} \quad \text{(I.8)}$$

$$V_A = \frac{B}{2\theta - 1}(1 + \frac{2\sqrt{3} - 3}{12c}) - B \quad \text{(I.9)}$$

$$L = \frac{\sqrt{3} - 2}{24c} \quad \text{(I.10)}$$

In particular, the expected utility of the agent is decreasing in c, and the utility of the principal and the total loss due to risk management are decreasing in c.

The ability to manage risk increases the expected payments received by the agent. The chosen probability distribution is non-uniform and implies that $y \geq .5$ with probability greater than one half. As it becomes more difficult to manage risk (i.e., c increases), the expected utility of the agent decreases. Counter-intuitively, however, a greater cost of c decreases the total value lost to risk management. Here, the reduction in risk management by the agent for greater values of c dominates the increase in the ex-post loss for a given level of risk management. As a result, greater efficiency and greater profits for the principal are achieved when manipulating the distribution is more costly.

Non-Uniform Case

[11] See also Zhou and Swan (2003) for a discussion of the optimality of a bonus within the Linear-Exponential-Normal model; their model, however, does not capture any cost associated to the use of the bonus.

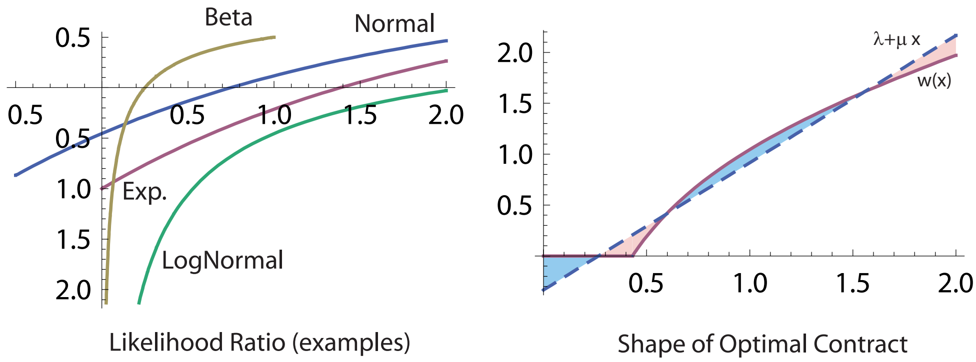


Figure I.4. Non-Uniform Distributions

To test the robustness of the findings, I extend next the analysis to more general distributional assumptions. Assume that, conditional on $a \in \{0,1\}$, the distribution of x is given by $f(x|a)$ strictly positive and bounded away from zero on a compact support that does not depend on a. As before, incentive-compatibility does not depend on risk management and can be written:

$$\int f(y|1)w(y)dy - B \geq \int f(y|0)w(y)dy \tag{I.11}$$

The next result is an immediate corollary to Proposition 2.3.

Proposition 2.4. *The optimal contract takes one of the following forms (up to sets with zero probability):* $w(y) = \max(0, a_0 + a_1 y + a_2 \frac{f(x|1)-f(x|0)}{f(x|1)})$.

In Proposition 2.4, the term $(f(x|1)-f(x|0))/f(x|1)$ is the likelihood ratio and captures whether an outcome is consistent with high effort or low effort (i.e., whether x is more likely under $a = 1$ or $a = 0$). In this model, the first-order conditions select a contract that is linear in the likelihood ratio.

Compare this characterization to the standard model in which the agent cannot manage risk. In the standard model, the optimal contract prescribes $w(y) = 0$ except for outcomes with the maximum likelihood ratio (Tirole (2006), p.130-135). This can create two problems. One, the resulting optimal contract will generally be extremely volatile, thus casting some doubts on the match between model and data, as well as whether it is appropriate to assume risk-neutrality by the agent over the (large) range of payoffs from the contract. Two, if the likelihood ratio is strictly increasing, an optimal contract does not exist. This problem can be prevented by assuming that the principal has limited liability (Innes 1990); however, the idea that in certain states the principal makes a transfer of all of its cash reserves to the executive is unappealing, in particular for large corporations.

For most standard distributions, the likelihood ratio is a concave function of x. For example,

Figure I.4 plots the likelihood ratio when $x|a$ is: (a) Normal $N(0.5a+0.5, 1)$, (b) Exponential $\mathcal{E}(1/(1-0.5a))$, (c) Beta $\beta(0.5a+0.5, 1)$, (d) LogNormal $LN(0.5a+0.5, 1)$. By Proposition 2.4, this suggests wages of the form as represented in the right-hand side of Figure I.4 which feature manipulation to beat the threshold at which the limited liability no longer binds and a hump-shape similar to that of Figure I.3.[12]

2. 5 S-Shape Response to Current Earnings

To capture the time series predictions of the model, I extend the model to two periods. The effort choice problem is assumed to remain static (so that the generic framework is essentially unchanged) although the risk management problem over multiple periods is now explicitly modeled. The manager chooses effort a ex-ante but produces period 1 return z and period 2 return x, for total earnings over both periods $z+x$. Each variable is uniformly distributed on $[0, .5]$ and assumed to be independent to clear out any effect driven by the correlation between periods and not the incentive problem. The distribution of $2x$ depends on effort and satisfies the previous assumptions, while the distribution of z does not depend on effort (the results are similar if z depends on effort).

The joint distribution of (z, x) is denoted $f(.,.|a)$. Conditional on no risk management, expected earnings over the two periods are given by $\mathbb{E}(z+x) = 1$. The manager can choose any other earnings (y_1, y_2) with distribution $\hat{f}(.,.)$ such that the mean of this distribution is lower when $\hat{f}(.,.)$ is more distinct from $f(.,.|a)$. The principal does not observe per-period earnings and can offer a wage $w(\pi)$ that can only depend on $\pi = y_1+y_2$. The difference between the two distributions is denoted $\delta(y_1, y_2) = \hat{f}(y_1, y_2) - f(y_1, y_2|a)$. Using the Integrated Squared Error as the metric yields that $\delta(.,.)$ can be achieved if $\delta(.,.)$ integrates to one and:

$$\int\int \delta(y_1, y_2)(y_1+y_2)dy_1dy_2 = -c\int\int \delta(y_1, y_2)^2 dy_1dy_2/2 \tag{I.12}$$

The first-order condition with respect to $\delta(.,.)$ takes a form that is similar to the one-period model:

$$w(y_1+y_2) - \lambda - \mu(y_1+y_2) = \mu c\delta(y_1, y_2) \tag{I.13}$$

The analogue of Proposition 2.1 in the multi-period model is given as follows:

[12] In Figure I.4, I represent four regions - the two regions in the middle generating the hump-shape. There may also be cases with only the first three regions if the limited liability binds over only very low outcomes.

Proposition 2.5. *For a given non-constant $w(.)$ such that $a = 1$ is elicited, the solution $(\lambda, \mu, \delta(.))$ is unique and given as follows:*

$$\begin{aligned}
\lambda &= 22\mathbb{E}(w(x+z)) - \sqrt{6}\sqrt{62\mathbb{E}(w(x+z))^2 + \mathbb{E}((w(x+z))^2)} && \text{(I.14)}\\
\mu &= -36\mathbb{E}(w(x+z)) + \sqrt{6}\sqrt{62\mathbb{E}(w(x+z))^2 + \mathbb{E}((w(x+z))^2)} && \text{(I.15)}\\
c\delta(x) &= \frac{w(x+z) - 22\mathbb{E}(w(x+z))}{-36\mathbb{E}(w(x+z)) + 2\sqrt{6}\sqrt{62\mathbb{E}(w(x+z))^2 + \mathbb{E}((w(x+z))^2)}}\\
&\quad -(y - \frac{\sqrt{6}\sqrt{62\mathbb{E}(w(x+z))^2 + \mathbb{E}((w(x+z))^2)}}{-36\mathbb{E}(x+z) + 2\sqrt{6}\sqrt{62\mathbb{E}(w(x+z))^2 + \mathbb{E}((w(x+z))^2)}}) && \text{(I.16)}
\end{aligned}$$

Comparing μ from Proposition 2.1 and in Equation (I.15) reveals that, for a given wage, μ is always greater in the two-period model. In addition, consider the level of y such that the agent paid the expected wage $\mathbb{E}(w(x+z))$ would not manage risk. From Equation (I.16), the level of y can be verified to be greater than .5, versus always .5 in the one-period model, indicating a stronger bias to manage against large earnings. These findings are intuitive. In the single period model, the agent can only manage earnings over outcomes in the same period. In the two-period model, the manager can shift earnings across periods as well, thus magnifying incentives to manipulate.

Proposition 2.6. *A linear contract (i.e., no risk management) is suboptimal. The optimal contract is linear over $(0, .25)$ and over $(.75, 1]$ and convex over $(.25, .75)$.*

Proposition 2.6 shows that the optimal contract has features similar to those shown earlier. First, the optimal contract is linear for extreme outcomes, since these outcomes do not become more indicative of high effort. Here, for $\pi \in [0, .25]$, the principal knows that x was in $[0, .25]$ as well (and vice-versa for $\pi \geq .75$). As before, the linearity is helpful to avoid excessive risk management by the manager. Second, the contracts features some non-linearities for intermediate outcomes in $[.25, .75]$. Over these outcomes, the signal z jams information about x and makes inference more difficult. I interpret here the convexity as an option-like feature within the contract. As a result, a fully linear contract that does not elicit risk management is suboptimal, and the principal tolerates some risk management in an optimal incentive scheme.

Bonus Contract and S-shape

As a simplifying case, consider next the case of a simple bonus scheme that pays $2K$ when the final return is above .75 and zero else. This corresponds to a simple example of a wage of the form described in Proposition 2.6.

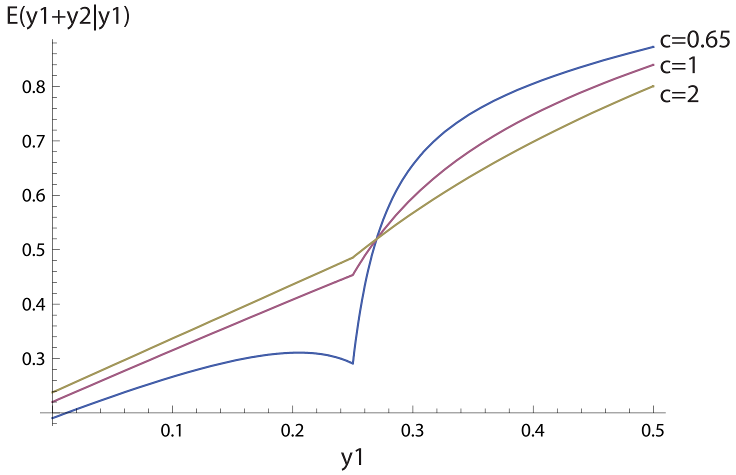


Figure I.5. $\mathbb{E}(y_1 + y_2|y_1)$ as a function of y_1

Proposition 2.7. *For a simple bonus scheme,*

If $y_1 < .25$,

$$\mathbb{E}(y_1 + y_2|y_1) = y_1 + \frac{1}{-48c + 48y_1 + 4\sqrt{13}} + \frac{1}{4} \tag{I.17}$$

If $y_1 \geq .25$,

$$\mathbb{E}(y_1 + y_2|y_1) = \frac{1}{4}\left(4y_1 + \frac{64(y_1 - 1)y_1 + \sqrt{13} + 9}{-12\left(-3 + \sqrt{13}\right)c + 12\sqrt{13}y_1 - 68y_1 - 3\sqrt{13} + 21} + 1\right) \tag{I.18}$$

To capture the stock response to current earnings, I plot in Figure I.5 $\mathbb{E}(y_1 + y_2|y_1)$. Note that, under no manipulation, y_2 and y_1 should be independent. Here, however, the manager will strategically shift earnings across periods to increase the likelihood that the bonus is paid. When earnings are close to average in period 1, the manager increases the probability of large gains in period 2 to beat the bonus threshold. On the other hand, given large gains in period 1, it is likely that the threshold will be attained, and thus there are fewer incentives to make gains in period 2. As shown in Figure I.5, the response flattens for large realizations of y_1. As a result, for a given realization of period 1 earnings, expectations about period 2 are S-shaped, even though the unmanaged earnings are independent.

3 General Model

3. 1 Assumptions

I generalize next the static model to provide more robust conditions under which the linearity and the hump-shape can be reproduced. Let $a \in [\underline{a}, \overline{a}]$ denote the effort choices available to the managers and let $f(.|a)$ denote the distribution of the original project without risk management. This distribution is assumed to have full support over $\mathbb{R}$ with a non-vanishing density on any interval. The principal is risk-neutral. However, the manager is risk-averse and achieves a utility $u(w(y)) - \psi(a)$ satisfying standard Inada conditions and: $u' > 0$, $u'' < 0$, $\psi(\underline{a}) = \psi'(\underline{a}) = 0$, $\lim_{a \to \overline{a}} \psi'(a) = +\infty$, $\psi'', \psi''' > 0$ except possibly at $a = \underline{a}$. Finally, the contract must prescribe a minimum reserve utility equal to b. The function $\psi(.)$ represents the cost of effort and risk-aversion by the agent replaces the assumption of a limited liability.

The cost of managing risk is generalized to a non-quadratic cost function. Assume that an agent exerting effort a can achieve any distribution $\hat{f}(y) = f(y|a) + \delta(y)$ satisfying:

$$\int \delta(y)dy \quad = \quad 0 \tag{I.19}$$

$$\int \delta(y)ydy \quad \leq \quad -\int C(a, y, \delta(y))dy \tag{I.20}$$

Equation (I.19) is required for $\hat{f}(y)$ to be valid probability distribution. Equation (I.20) captures the cost of managing risk. The function $C(a, y, \delta)$ represents a U-shaped 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 corner solutions when the positivity constraint on $\hat{f}(y)$ binds, I assume that $\lim_{\delta \to -f(y|a)} C(a, y_k, \delta) = +\infty$. The contract design problem can be stated as follows:

$$(P) \qquad \max_{a, w(.), \tilde{\delta}(.)} \int (f(y|a) + \tilde{\delta}(y))(y - w(y))dy$$

$$s.t. \qquad \int (f(y|a) + \tilde{\delta}(y))u(w(y))dy - \psi(a) \quad \geq b \tag{I.21}$$

$$(P_a) \quad (a, \tilde{\delta}(.)) \in \arg\max_{A, \delta(.)} \int (f(y|A) + \delta(y))u(w(y))dy - \psi(A) \tag{I.22}$$

s.t. (I.19) and (I.20)

In the rest of the essay, I analyze this contract design problem. Some of the assumptions are further generalized in the Appendix. In Appendix A.1 I extend the model to risk-aversion by the principal and in Appendix A.2 I consider a non additively-separable cost function.

3. 2 Analysis of the Contract

Problem of the Agent

Taking the wage as given, I solve first the problem of the agent. In this problem, let λ (resp. μ) denote the multiplier associated to Equation (I.19) (resp. (I.20)). Denote $\mathcal{C}^y(a,.)$ the inverse of $C_\delta(a,y,\delta)$ in δ. The next Proposition characterizes the optimal choice of the agent.

Proposition 3.1. *If $a > \underline{a}$ is elicited, the solution satisfies $\mu > 0$ and:*

$$\delta(y) = \mathcal{C}^y(a, \frac{u(w(y)) - \lambda}{\mu} - y) \tag{I.23}$$

Corollary 3.1. *The agent does not hedge if and only if $u(w(y)) = h_0 + h_1 y$.*

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 $a - \mathbb{E}(y|a)$.

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 3.1, 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 I.6, I plot the utility received by a risk-averse manager compensated with options.

Proposition 3.2. *Suppose: $\lim_{x \to +\infty} u'(x) = 0$ and $w(y) = \max(0, y)$ (normalized Call option). Then, there exist four regions $\theta_1 < 0 < \theta_2 < \theta_3$ such that:*

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

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

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

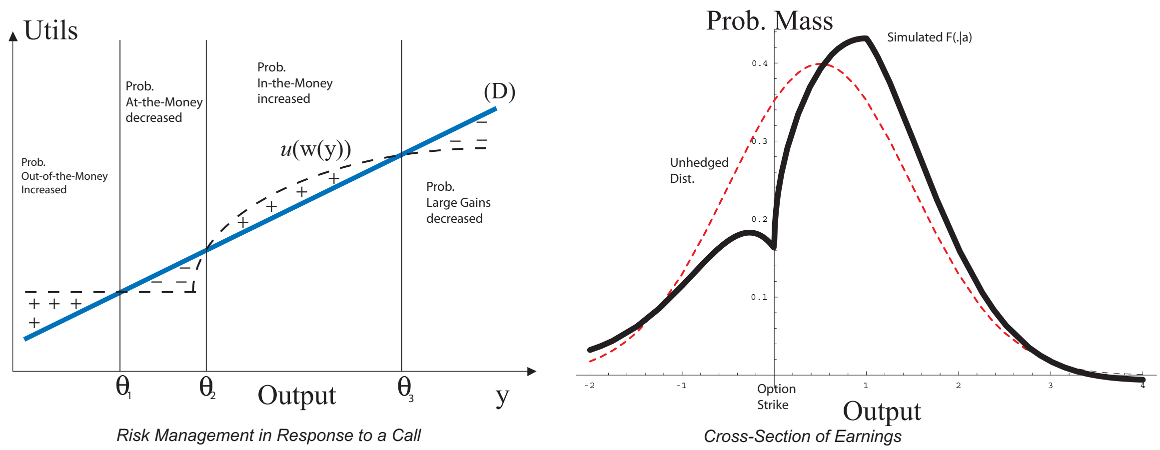


Figure I.6. Response to a Call Option

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

In the left-hand side of Figure I.6, I plot the response of a risk-averse manager. 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 $a = .5$. 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 $C(.5, y, \delta) = e^{r|y-1|}\delta^2$ for any y.[13]

The resulting distribution of y is plotted against that of x, on the right-hand side of Figure I.6. 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 I.1.[14]

Problem of the Principal

[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).

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 $\hat{f}(.)$ that is more informative on the actions of the agent.

Proposition 3.3. *The agent receives a contract eliciting $a > \underline{a}$.*

By choosing $a = \underline{a}$, 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 3.3 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. To analyze the problem theoretically, I assume here that the first-order approach is valid.

$$(P') \qquad \max_{\delta(.)} \int (\delta(y) + f(y|a))(y - u^{-1}[C_\delta(a, y, \delta(y))\mu + \lambda + \mu y])dy$$

s.t.

$$\int \delta(y)dy = 0 \qquad (\alpha) \qquad \text{(I.24)}$$

$$\int \delta(y)ydy = -\int C(a, y, \delta(y))dy \qquad (\beta) \qquad \text{(I.25)}$$

$$\mu = \psi'(a) - \mu \int f_a(y|a)C_\delta(a, y, \delta(y))dy \qquad (\gamma) \qquad \text{(I.26)}$$

$$b - \lambda - \mu a = -\psi(a) + \mu \int [(\delta(y) + f(y|a))C_\delta(a, y, \delta(y)) - C(a, y, \delta(y))]dy \qquad (\tau) \qquad \text{(I.27)}$$

Equations (I.24) and (I.25) are the feasibility conditions. Equation (I.26) is the agent's optimality condition on effort. Equation (I.27) is the participation of the agent (once wages have been substituted out). The multipliers associated to each constraint are denoted in parenthesis.

Proposition 3.4. *Let a be the elicited effort. The optimal contract satisfies:*

$$\begin{aligned} w(y) &= -\mu C_{\delta,\delta}(a, y, \delta(y))(\delta(y) + f(y|a))(\frac{1}{u'(w(y))} - \tau - \gamma \frac{f_a(y|a)}{\delta(y) + f(y|a)}) \\ &\quad -\alpha + (1 - \beta)y - \beta C_\delta(a, y, \delta(y)) - \mu\gamma C_{\delta,a}(a, y, \delta(y)) \end{aligned} \qquad \text{(I.28)}$$

I interpret next this optimality condition. Equation (I.28) 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 $S = 1/u'(w(y)) - \tau - \gamma f_a(y|a)/(f_a(y|a) + \delta(y))$. 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 (I.28) 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 (I.28), $-\alpha + (1-\beta)y - \beta C_\delta(a, y, \delta(y))$, captures the most preferred risk management choice of a principal controlling risk. Thus, when $S = 0$ (no distortions to optimal incentives), the principal behaves myopically as if taking the wage as given exogenously and managing risk optimally in response to it.

Corollary 3.2. *Suppose that $C_{\delta,\delta}(a, y, 0) = C_{\delta,a}(a, y, 0) = 0$ for all a, y, then $\delta(y) = 0$ for all k cannot be optimal.*

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 usefulness of the output signal at very little cost.[15] That is, the agency problem makes a risk-neutral principal effectively risk-averse to some of the noise in the production technology.

4 Properties of the Contract

4. 1 Linearity of the Optimal Contract

I explain first why the optimal contract should become linear over large outcomes. Note that, in the standard incentive model, $w(y)$ may also become linear if $(u')^{-1}(1/(\lambda + \mu\frac{f_a(y)}{f(y)}))$ becomes linear. This condition, however, does not map into a clear economic interpretation and would likely be violated in practice.

Proposition 4.1. *Suppose that for all a, δ, y,*

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

[15] 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.

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

(iii) $u'(w)$ *converges to a strictly positive number 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).[16] 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 other outcomes informative on the actions of the agent while simultaneously generating large risk management cost.[17]

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 $C^j(a, y, \delta) = C(a, y, \delta)/j$ ($j > 1$). As j becomes large, the cost function becomes small. In addition, I make the assumption that $\frac{C_\delta(a,y,\delta(y))}{C(a,y,\delta(y))}$ is bounded away from zero. 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 4.2. *As* j *becomes large, the contract converges to a (convex) contract:* $w(y) = u^{-1}(a_0 + a_1 y)$.

I argue here that the contract should become linear in utility as the cost of managing risk becomes small. 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.[18] 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

[16] Interestingly, this result presents an apparent similarity with Diamond (1998). Diamond explains that, as the size of (all of) the firm's earnings becomes large relative to the cost of effort, the optimal compensation scheme converges to a linear function.

[17] 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.

[18] 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 earnings, 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).

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).

4. 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 (I.23) and (I.28) in terms of their continuous analogue (omitting the dependence on y and denoting the likelihood ratio $LR = f_a(y|a)/f(y|a)$):

$$
\begin{aligned}
w &= u^{-1}(\lambda + \mu y + \mu C_\delta(a, y, \delta)) && \text{(I.29)}\\
w &= -\alpha + (1-\beta)y - \beta C_\delta(a, y, \delta) - \mu\gamma C_{\delta,a}(a, y, \delta)\\
&\quad -\mu C_{\delta,\delta}(a, y, \delta)((\delta + p)(\frac{1}{u'(w)} - \tau) - \gamma p LR) && \text{(I.30)}
\end{aligned}
$$

In the model, there are only two endogenous variables that depend on the state of the world, w and δ. Because the other endogenous variables are global in the problem and are unaffected by a point-wise variation, I take here the multipliers and effort as constants. Then, one may view Equations (I.29) and (I.30) as two Equations in two unknowns (w, δ) 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, δ, $C_{\delta,\delta,a}(a, y, \delta) = C_{\delta,y}(a, y, \delta) = C_{\delta,a,y}(a, y, \delta) = C_{\delta,a,y}(a, y, \delta) = 0$, $C_{\delta,\delta,\delta}(a, y, \delta) < 0$ 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 (I.29) and (I.30).

Corollary 4.1. *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.[19]

[19]One may guarantee that $\gamma > 0$ if the effort choice may take only two values. Further, by continuity, γ should be positive

(iii) If $\gamma > 0$ and $|LR|$ is sufficiently large with $Sign(LR) > 0$ (resp. $Sign(LR) < 0$), $\partial\delta/\partial f > 0$ (resp. $\partial\delta/\partial f < 0$), 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).[20] Then, this comparative static will work to reduce the likelihood of negative likelihood ratio states and increase the likelihood of positive likelihood 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 $\partial\delta/\partial f < 0$. This will work to flatten the distribution. As the likelihood ratio changes sign before the mode, y will lie in a region where $\partial\delta/\partial f > 0$, 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.

---

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).

[20] An intuition for an S-shaped likelihood ratio is that most of the gains in informativeness are realized over intermediate outcomes.

5 Numerical Examples

5. 1 S-Shape with Options

To conclude, I present several numerical examples. I assume now that x has an absolutely-continuous distribution with density $f(.|a)$. I show first how the framework can generate a S-shaped response of stock prices to current earnings. 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 $\lim_{\delta \to -\max_a f(x|a)} C(\delta) = +\infty$. Let $\mathbb{y} = (y_1, y_2)$ where $y_t > 0$ corresponds to return on assets for period t; $f(y_1, y_2|a)$ denotes the (multivariate) density of (y_1, y_2) when the agent does not manage risk while $\hat{f}(y_1, y_2)$ denotes the distribution of returns when risk is managed. In this example, I assume that total earnings are multiplicative in past earnings, for example as in returns on assets; for example, a successful investment in previous periods may magnify how much is invested in future periods (similar results can be obtained in the additive case considered earlier).

$$\int\int \hat{f}(y_1, y_2) y_1 y_2 dy_1 dy_2 \leq a - \int\int C(f(y_1, y_2|a) - \hat{f}(y_1, y_2)) dy_1 dy_2 \tag{I.31}$$

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

$$\int\int \hat{f}((y_1, y_2)) dy_1 dy_2 \tag{I.32}$$

In addition, to keep intact the contract design problem studied earlier, I assume that the wage may only depend on $y_1 y_2$ (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. Stating the first-order condition in the problem of the agent yields:

$$u(w(y_1 y_2)) - \lambda - \mu y_1 y_2 - \mu \mathcal{C}(\delta(y_1, y_2)) = 0 \tag{I.33}$$

I turn now to the S-Shaped response of the market price to current returns. In this model, expectations about y_2 given y_1 are captured by $MV(y_1)$:

$$MV(y_1) = \frac{\int y_2 \hat{f}(y_1, y_2) dy_2}{\int \hat{f}(y_1, y_2) dy_2} \tag{I.34}$$

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

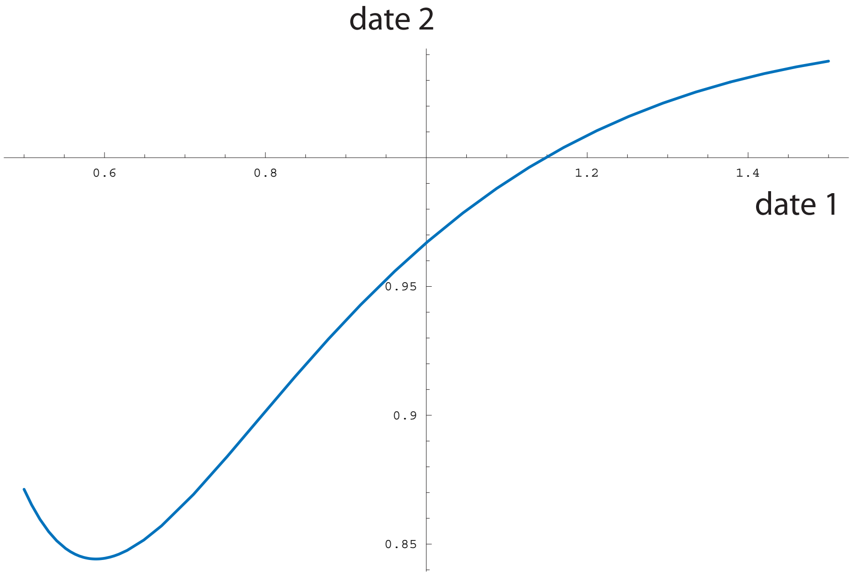


Figure I.7. Earnings Response

previous Section. Suppose that the manager is risk-neutral and compensated with a simple Call option with maturity date 2 and strike 1.

In Figure I.7, the density $\delta(y_1, y_2)$ 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 (I.34) 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.[21]

5. 2 Cross-Section and Contracts

To further illustrate the findings, I develop a numerical example inspired from the LEN framework. 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 $(y_k)_{k=1}^N$ and such that each point has equal probability conditional on $F(x|\overline{a})$. 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: $\sigma = 1$, $\overline{a} = .5$, $r = 1$, $\psi(\overline{a}) = .1$, $c_h = 1$ and $b = -1$. The results are shown in Figure 5. 2 under four treatments: 1. First-Best Contract, 2. Second-

[21] 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).

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., the likelihood ratio is linear and not S-shaped as required in the comparative statics) to produce a clear hump-shape.

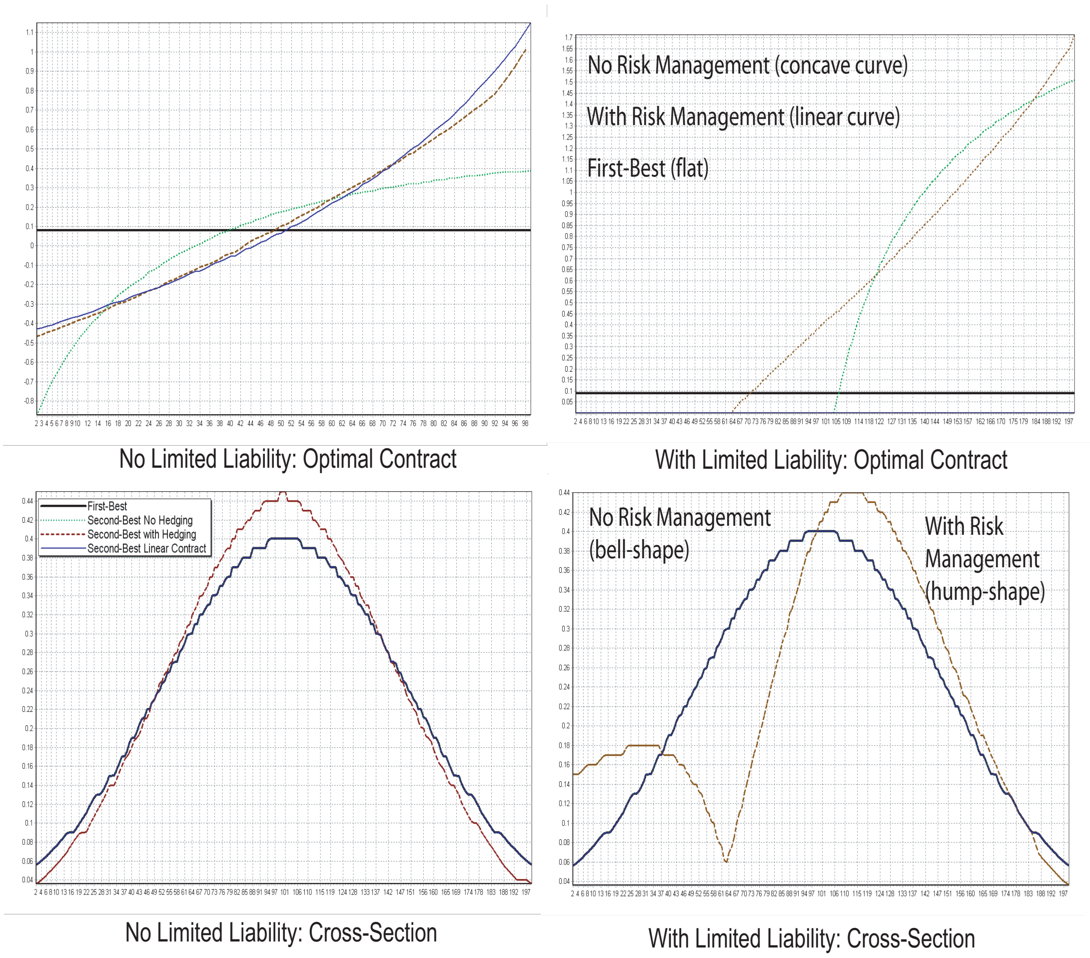


Figure I.8. Optimal Contracts and Risk Management

To facilitate a hump-shape even for a linear 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 5. 2, 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.

6 Appendix

6. 1 Appendix A: 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. 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 risk 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.

$$\int \delta(y) dy = \quad 0 \qquad (\tilde{\alpha}) \qquad \text{(I.35)}$$

$$\int \delta(y) y dy = \quad -\int C(\delta(y)) dy \qquad (\tilde{\beta}) \qquad \text{(I.36)}$$

$$\psi(a) + b \leq \quad \int u(w(y))(\delta(y) + f(y|a)) dy \qquad (\tilde{\tau}) \qquad \text{(I.37)}$$

Remark 1: In first-best, no risk management cannot be optimal.

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

$$\frac{v'(y - w(y))}{u'(w(y))} = \tilde{\tau} \qquad \text{(I.38)}$$

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:

$$v(y - w(y)) - \tilde{\alpha} - \tilde{\beta}(y + C'(\delta(y))) + \tilde{\tau} u(w(y)) = 0 \qquad \text{(I.39)}$$

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 $v'(y - w(y)) > 0$.□

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 3.3 (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. 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 risk choices in this case.[22] 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 $v(.)$.

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

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

s.t.

$$\int \delta(y)dy = \quad 0 \qquad (\alpha) \qquad \text{(I.40)}$$

$$\int \delta(y)ydy = \quad -\int C(\delta(y))dy \qquad (\beta) \qquad \text{(I.41)}$$

$$\mu = \quad \psi'(a) - \mu \int f_a(y|a)C_\delta(\delta(y))dy \qquad (\gamma) \qquad \text{(I.42)}$$

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

$$-\psi(a) + \mu \int \quad [(\delta(y) + f(y|a))C'(\delta(y)) - C(\delta(y))]dy \quad (\tau) \qquad \text{(I.43)}$$

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.

Remark 2: The ratio of marginal utility $v'(y - w(y))/u'(w(y))$ cannot be constant, i.e. risk-sharing is imperfect.

Proof: 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 (I.28) that $v(y - w(y))$ is constant, a contradiction to perfect risk-sharing.$\square$

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 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.

Remark 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, 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.

Proof: 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: $\frac{\partial v'(y-w(y))/u'(w(y))}{\partial y} \geq 0$.$\square$

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.

6. 2 Appendix B: Non-Separable 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. It is now helpful to use a discrete version of the model, so that the outcomes must be realized on $(x_k)_{k=1}^N$, denoted as a vector $X = (x_1, \ldots, x_N)'$. 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 $P(a)$.

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

$$+(\hat{P} - P(a))' D^2\theta(\hat{P} - P(a))/2 \tag{I.44}$$

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.

$$\theta \approx (\hat{P} - P(a))' H (\hat{P} - P(a))/2 \tag{I.45}$$

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.

$$\begin{aligned} \mathbb{1}'\Delta = & \quad 0 & (\lambda) \quad & \quad \text{(I.46)} \\ \Delta' X \leq & \quad -\Delta' H \Delta/2 & (\mu) \quad & \quad \text{(I.47)} \end{aligned}$$

Differentiating with respect to Δ and rearranging: $\mu H \Delta = \mathcal{U} - \lambda - \mu X$. Pre-multiplying by Δ' and using Equation (I.47), $-2\mu\Delta' X = \Delta'(\mathcal{U} - \lambda - \mu X)$. 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 (I.46), $\mathbb{1}' H^{-1} (\mathcal{U} - \lambda - \mu X) = 0$. 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:

$$\lambda = \frac{\mathbb{1}'H^{-1}\mathcal{U}}{\mathbb{1}'H^{-1}\mathbb{1}} - \frac{\mathbb{1}'H^{-1}X}{\mathbb{1}'H^{-1}\mathbb{1}} \sqrt{\frac{(\mathcal{U}'H^{-1}\mathcal{U}\mathbb{1}'H^{-1}\mathbb{1} - (\mathbb{1}'H^{-1}\mathcal{U})^2}{X'H^{-1}X\mathbb{1}'H^{-1}\mathbb{1} - (\mathbb{1}'H^{-1}X)^2}} \tag{I.48}$$

$$\mu = \sqrt{\frac{\mathcal{U}'H^{-1}\mathcal{U}\mathbb{1}'H^{-1}\mathbb{1} - (\mathbb{1}'H^{-1}\mathcal{U})^2}{X'H^{-1}X\mathbb{1}'H^{-1}\mathbb{1} - (\mathbb{1}'H^{-1}X)^2}} \tag{I.49}$$

$$\mu\Delta = H^{-1}(\mathcal{U} - \lambda - \mu X) \tag{I.50}$$

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,

$$\delta_i = (u(w_i)/\mu - \lambda/\mu - x_i)/D_{i,i} \tag{I.51}$$

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.

6. 3 Appendix C: Omitted Proofs

Proof of Proposition 2.1: Substitute $\delta(y)$ from Equation (I.4) in Equation (I.1).

$$\int \delta(y)dy = \int [\frac{w(y)}{\mu c} - \frac{\lambda}{\mu c} - \frac{y}{c}]dy = 0 \tag{I.52}$$

$$\frac{\int w(y)dy}{\mu c} - \frac{\lambda}{\mu c} - \frac{1}{2c} = 0 \tag{I.53}$$

Similarly,

$$\int \delta(y)ydy = -\frac{\lambda}{2c\mu} + \frac{\int yw(y)dy}{c\mu} - \frac{1}{3c} \tag{I.54}$$

and:

$$c\int \delta(y)^2 dy/2 = c\int [\frac{w(y)}{\mu c} - \frac{\lambda}{\mu c} - \frac{y}{c}]^2 dy/2 \quad \text{(I.55)}$$

$$= \frac{3\lambda^2 + 3\lambda\mu - 6\lambda \int w(y)dy + 3\int w(y)^2 dy - 6\mu \int yw(y)dy}{6c\mu^2} \quad \text{(I.56)}$$

Substituting Equations (I.54) and (I.56) into Equation (I.2), the following must then hold:

$$-13\mu^2 + 12\lambda(\lambda - 2\int w(y)dy) + 12\int w(y)^2 dy = 0 \quad \text{(I.57)}$$

Substituting λ from Equation (I.53) in Equation (I.57) yields:

$$\mu^2 + 12(\int w(y)dy)^2 - 3\int w(y)^2 dy = 0 \quad \text{(I.58)}$$

Solving this Equation yields a unique positive root for μ (the other root is negative and is not a maximum) as given by (I.6). Substituting for λ and $\delta(y)$ yields Equation (I.5).□

Proof of Lemma 2.1: Suppose $a = 1$ is elicited and denote V the expected profit of the principal.[23]

$$V = \int (1 + \delta(y))(y - w(y))dy \quad \text{(I.59)}$$

$$= \frac{\int w(y)dy\lambda}{c\mu} - \frac{\lambda}{2c\mu} - \int w(y)dy - \frac{\int w(y)^2 dy}{c\mu} + \frac{\int yw(y)dy}{c} + \frac{\int yw(y)dy}{c\mu} - \frac{1}{3c} + \frac{1}{2} \quad \text{(I.60)}$$

$$= -\frac{6c(2\int w(y)dy - 1)}{12c}$$
$$+\frac{\int w(y)dy \left(-2\sqrt{3}\int w(y)dy + 6\sqrt{\int w(y)^2 dy - \int w(y)dy^2} + \sqrt{3}\right) + 2\sqrt{3}(\int w(y)^2 dy - \int yw(y)dy)}{12c\sqrt{\int w(y)^2 dy - \int w(y)dy^2}}$$
$$-\frac{\int yw(y)dy}{c} \quad \text{(I.61)}$$

Let $K_w = \int w(y)dy$, $K_{w^2} = \int w(y)^2 dy$ and $K_{wy} = \int yw(y)dy$, and replacing the terms with integrals of $w(y)$ in V, let V' denote the new objective.

The problem of the principal can be written as follows:

$$\max_{w(.), K_w, K_{w^2}, K_{wy}} V'$$

[23] To save space, some simple algebraic steps have been omitted and are available on request or from a separate Mathematica notebook supplied next.

s.t.

$$\int_{.5}^{1} w(y)dy = \int_{1}^{1.5} w(y)dy - \frac{B}{2\theta - 1} \tag{I.62}$$

$$K_w = \int w(y)dy \tag{I.63}$$

$$K_{w^2} = \int w(y)^2 dy \tag{I.64}$$

$$K_{wy} = \int yw(y)dy \tag{I.65}$$

The multipliers associated to each of these constraints are denoted, respectively, γ, L_w, L_{w^2} and L_{wy}. Differentiating pointwise with respect to $w(y)$, if the limited liability does not bind:

$$L_w + 2w(y)L_{w^2} + yL_{wy} + 1_{y\geq 1}\gamma = 0$$

Therefore $w(y)$ has the linear form shown in the Lemma. To conclude, note that if $a_1 < 0$ in the relaxed problem, the constraint imposing $w(.)$ non-decreasing will bind, which will imply $a_1 = 0$.□

Proof of Proposition 2.3: Consider a wage of the form: $w(y) = a_1 \max(0, y - y_0) + 1_{y\geq 1}e$. There are two cases to exclude: a. $y_0 = 0$ and $e = 0$, i.e. simple equity, b. $e = 0$ and $y_0 > 0$, i.e. option but no bonus.

Case a. Clearly, the limited liability must bind for $y = 0$. Consider an incentive-compatible wage schedule of the form: $w(y) = a_1 y + e1_{y\geq 1}$, where e represents a (possibly zero) bonus payment.

For this wage to be incentive-compatible, it must hold that:

$$\int_{.5}^{1} (a_1 y + e)dy = \frac{B}{2\theta - 1} + \int_{0}^{.5} a_1 y dy \tag{I.66}$$

Therefore:

$$a_1 = \frac{4B}{2\theta - 1} - 2e \tag{I.67}$$

This implies that:

$$\int w(y)dy = \frac{2B}{2\theta - 1} - \frac{e}{2} \tag{I.68}$$

$$\int w(y)^2 dy = \frac{1}{24}\left(\frac{4B}{1 - 2\theta} + 2e\right)^2 + \frac{28B^2 + 10e(1 - 2\theta)B + (e - 2e\theta)^2}{6(1 - 2\theta)^2} \tag{I.69}$$

$$\int yw(y)dy = \frac{4B}{3(2\theta - 1)} - \frac{7e}{24} \tag{I.70}$$

Substituting these expressions in V,

$$V = -\frac{c\left(\frac{24B}{2\theta-1} - 6e - 6\right) + \frac{\left(2\theta+\sqrt{16B^2+e(4-8\theta)B+(e-2e\theta)^2}-1\right)\left(-8B-e+2e\theta+2\sqrt{16B^2+e(4-8\theta)B+(e-2e\theta)^2}\right)}{2(2\theta-1)\sqrt{16B^2+e(4-8\theta)B+(e-2e\theta)^2}}}{12c} \tag{I.71}$$

The derivative of V in e is verified to be always strictly positive (using an algebraic solver such as Mathematica or Maple), which implies the boundary solution $a_1 = 0$ and $e = 2B/(2\theta - 1)$, that is a simple bonus.

Case b. Consider an option: $w(y) = a_1 \max(y - y_0, 0) + 1_{y \geq .5} e$. Since the intermediate algebraic steps are similar to case a, they are omitted to save space.

Case b.1. Suppose in addition that $y_0 > .5$. In this case, it must be that $e = 0$. For this to be incentive-compatible, it must hold that:

$$a_1 = \frac{2B}{(2\theta - 1)(1 - y_0)^2} \tag{I.72}$$

Again, V is written after a_1 and the function form are substituted in. This leads to the following:

$$\frac{\partial V}{\partial y_0} = -\frac{3y_0^2(-2\theta(y_0-1) + y_0 - 1) - 2B(3y_0+1)\left(\sqrt{-\frac{3y_0+1}{y_0-1}}y_0^2 - 2\sqrt{-\frac{3y_0+1}{y_0-1}}y_0 + \sqrt{\frac{3y_0+1}{1-y_0}} - 1\right)}{6c(2\theta-1)\left(-3y_0^2 + 2y_0 + 1\right)^{3/2}} \tag{I.73}$$

The term above is strictly negative, and thus $y_0 > .5$ is suboptimal (i.e., the principal would gain by reducing the strike of the option).

Case b.2. Consider next a strike below .5. Incentive-compatibility for $a = 1$ implies:

$$a_1 = \frac{2(2B + e - 2e\theta)}{(2\theta - 1)(2y_0^2 - 1)} \tag{I.74}$$

One can substitute this expression in V and verify that the derivative in e, evaluated at $e = 0$, is strictly positive (to save space, the expression is provided in a separate Mathematica notebook).□

Proof of Proposition 2.5: Substituting $\delta(y)$ from Equation (I.4) yields:

$$\int \delta(y)dy = \frac{\int w(y)dy}{c\mu} - \frac{2\lambda + \mu}{8c\mu} = 0 \tag{I.75}$$

The second constraint can similarly be written as:

$$-7\mu^2 + 24\lambda(\lambda - 2\int w(y)dy) + 24\int w(y)^2 dy = 0 \tag{I.76}$$

Solving for λ and μ yields the desired result.□

Proof of Proposition 2.6: 1. Calculate first the distribution of $\pi = x + z$ conditional on no effort $F(.)$ and the corresponding density $g(\pi|a = 0)$. The density of $g(\pi|a = 1)$ is then obtaining by setting $\theta = .5$. Finally, I compute the likelihood ratio associated to the problem.

(i) If $\pi < .25$,

$$\begin{aligned} F(\pi) &= \int_0^{\pi}\int_0^{\pi-x} 8\theta dz dx & \text{(I.77)}\\ &= 4\theta\pi^2 & \text{(I.78)}\\ g(\pi|0) &= 8\theta\pi & \text{(I.79)}\\ g(\pi|1) &= 4\pi & \text{(I.80)}\\ \frac{g(\pi|1) - g(\pi|0)}{g(\pi|1)} &= 1 - 2\theta & \text{(I.81)} \end{aligned}$$

(ii) If $\pi \in [.25, .5)$,

$$\begin{aligned} F(\pi) &= \int_0^{.25}\int_0^{\pi-x} 8\theta dz dx + \int_{.25}^{\pi}\int_0^{\pi-x} 8(1-\theta) dz dx & \text{(I.82)}\\ &= \frac{1}{4} - \frac{\theta}{2} + (4\theta - 2)\pi + (4 - 4\theta)\pi^2 & \text{(I.83)}\\ g(\pi|0) &= 8(1-\theta)\pi + 4\theta - 2 & \text{(I.84)}\\ g(\pi|1) &= 4\pi & \text{(I.85)}\\ \frac{g(\pi|1) - g(\pi|0)}{g(\pi|1)} &= (2\theta - 1)(1 - 1/\pi) & \text{(I.86)} \end{aligned}$$

(iii) If $\pi \in [.5, .75)$,

$$F(\pi) = \int_0^{\pi-.5} 4\theta dx + \int_{\pi-.5}^{.25}\int_0^{\pi-x} 8\theta dz dx + \int_{.25}^{0.5}\int_0^{\pi-x} 8(1-\theta) dz dx \quad \text{(I.87)}$$

$$= 2\pi + \theta\left(-4(\pi-1)\pi - \frac{1}{2}\right) - \frac{3}{4} \quad \text{(I.88)}$$

$$g(\pi|0) = 4\theta + 2 - 8\pi \quad \text{(I.89)}$$

$$g(\pi|1) = 4(1-\pi) \quad \text{(I.90)}$$

$$\frac{g(\pi|1) - g(\pi|0)}{g(\pi|1)} = \frac{(2\pi-1)(2\theta-1)}{2(\pi-1)} \quad \text{(I.91)}$$

(iv) If $\pi \geq .75$,

$$F(\pi) = \theta + \int_{.25}^{\pi-0.5} 4(1-\theta)dx + \int_{\pi-.5}^{.5} 8(1-\theta) dz dx \quad \text{(I.92)}$$

$$= -3 + 4\theta + 8(1-\theta)\pi - 4(1-\theta)\pi^2 \quad \text{(I.93)}$$

$$g(\pi|0) = 8(1-\theta)(1-\pi) \quad \text{(I.94)}$$

$$\frac{g(\pi|1) - g(\pi|0)}{g(\pi|1)} = 2\theta - 1 \quad \text{(I.95)}$$

Note that the likelihood ratio is continuous, constant on $(0, .25)$ and $(.75, 1)$, convex on $(.25, .5)$ and on $(.5, .75)$. Note finally that the likelihood ratio is also differentiable at .5, therefore it is convex on the region $(.25, .75)$.

2. Calculate next the expected utility of the principal:

$$V = \frac{24(18c+11)K_w^2 - 6\left(4c\left(\sqrt{6}\sqrt{62K_w^2 + K_{w^2}} + 9\right) + 72K_{wy_1} + 72K_{wy_2} + 2\sqrt{6}\sqrt{62K_w^2 + K_{w^2}} + 1\right)K_w}{24c\left(\sqrt{6}\sqrt{62K_w^2 + K_{w^2}} - 18K_w\right)}$$
$$+ \frac{-12K_{w^2} + 12(K_{wy_1} + K_{wy_2}) + 12\sqrt{6}\sqrt{62K_w^2 + K_{w^2}}(c + 2(K_{wy_1} + K_{wy_2})) - \sqrt{6}\sqrt{62K_w^2 + K_{w^2}}}{24c\left(\sqrt{6}\sqrt{62K_w^2 + K_{w^2}} - 18K_w\right)} \quad \text{(I.96)}$$

where: $K_w = \int\int w(y_1 + y_2)dy_1dy_2$, $K_{wy_i} = \int\int y_i w(y_1 + y_2)dy_1dy_2$ $(i = 1, 2)$ and $K_{w^2} = \int\int (w(y_1 + y_2))^2 dy_1dy_2$.

3. Consider next a simple linear wage with, possibly, a bonus at $\pi = .75$: $w(y) = a_1\pi + 1_{\pi\geq .75}e$. As before, it needs to be shown that $e = 0$ can be improved with a small bonus.

Rewriting the incentive-compatibility:

$$\int g(\pi|a=0)w(\pi)d\pi + B = \int g(\pi|a=1)w(\pi)d\pi \quad \text{(I.97)}$$

$$\frac{e(1-\theta)}{4} + \frac{a_1(5-2\theta)}{8} + B = \frac{a_1}{2} + \frac{e}{8} \quad \text{(I.98)}$$

Solving in a_1,

$$a_1 = \frac{8B}{2\theta - 1} - e \quad \text{(I.99)}$$

As before, I substitute a_1, solve for the K associated to the wage and differentiate the value to the principal in e:

$$\frac{\partial V}{\partial e}|_{e=0} = \frac{3\left(8B\left(-758\left(-703+36\sqrt{379}\right)c - 1203\sqrt{379} + 21603\right) + 3\left(-379+22\sqrt{379}\right)(2t-1)\right)}{48512\left(-18+\sqrt{379}\right)^2 Bc} \quad \text{(I.100)}$$

This term is always strictly positive, which confirms that a linear wage is suboptimal.□

Proof of Proposition 3.1: The first-order condition in $\delta(.)$ is:

$$u(w(y)) - \lambda - \mu y - \mu C_\delta(a, y, \delta(y)) = 0 \quad \text{(I.101)}$$

Note that if $\mu = 0$, $u(w(y)) = \lambda$ and therefore $a = \underline{a}$ is elicited, a contradiction. Rearranging this expression yields the result.□

Proof of Corollary 3.1: First, $\delta(y) = 0$ for all y implies that $u(w(y)) = \lambda + \mu y$ (necessity). Second, let $u(w(y)) = h_0 + h_1 y$. Then, the agent must achieve $h_0 + h_1 a$, which can be achieved with no risk management (sufficiency).□

Proof of Proposition 3.2: 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. But then this would imply that the manager increases (decreases) the likelihood of large (low) y_k, a contradiction to Equation (I.20). 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.□

Proof of Proposition 3.3: (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:

$$\int f(y|a)u(w(y))dy - \psi(a) = \mu a - \lambda - \psi(a)$$

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:

$$\max_a \int f(y|a)(y - u^{-1}(\psi'(a)(y-a) + b))dy$$

The first-order condition for this problem is:

$$\int f_a(y|a)(y - w(y))dy - \psi''(a) \int f(y|a) \frac{1}{u'(w(y))}(y-a)dy$$

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:

$$\begin{aligned} \int f(y|a)(y-a)dy &= \int_{y<\underline{a}} f(y|a)(y-a)dy + \int_{y\geq\underline{a}} f(y|a)(y-a)dy \\ &\leq \int_{y<\underline{a}} f(y|a)(y-a)dy + \int_{y\geq\underline{a}} f(y|a)(y-a)dy \\ &\leq 0 \end{aligned}$$

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

Proof of Corollary 3.2: If not managing risk is optimal, $u(w(y)) = \lambda + \mu y$. But then, by Equation (I.28), $w(y) = \alpha + \beta y$, but since u is strictly concave, this is a contradiction.□

Proof of Proposition 4.1: If, $C_{\delta,\delta}(a, y, \delta(y))$ does not converge to zero, then, by assumption, the total loss be greater or equal than $\int_{K'}^{+\infty} C_{\delta,\delta}(a, y, \delta(y))dy$, for K' chosen sufficiently large - a contradiction. By Equation (I.23), as y becomes large, $C_\delta(a, y, \delta(y))$ converges to a linear function of $(w(y)-\lambda)/\mu - y$. Substituting this in Equation (I.28), $w(y)$ converges to $-\alpha + (1-\beta)y - \beta\{(w(y) - \lambda)/\mu - y\}$.□

Proof of Proposition 4.2: As c becomes small, the principal can achieve first-best. 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 $\lim_{j\to+\infty}\int C(a,y,\delta(y))/j = 0$. This implies that $C(a,y,\delta(y))/j$ goes to zero for all y. But then $C_\delta(a,y,\delta(y))/j$ also converges to zero. And by Equation (I.23), it must then be that $u(w(y))$ converges to a linear function.□

Proof of Corollary 4.1: 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 (I.29) and (I.30) in x,

$$\begin{aligned}\frac{\partial w}{\partial x} &= (\mu+\mu C_{\delta,\delta}\frac{\partial\delta}{\partial x}+\mu C_{\delta,x})/u'(w)\\ \frac{\partial w}{\partial x} &= (1-\beta)-\beta\frac{\partial\delta}{\partial x}C_{delta,\delta}-\beta C_{\delta,x}-\mu\gamma C_{\delta,a,x}-\mu\gamma C_{\delta,\delta,a}\frac{\partial\delta}{\partial x}-\mu(C_{\delta,\delta,x}+C_{\delta,\delta,\delta}\frac{\partial\delta}{\partial x})((d+f)(1/u'(w)-\tau)-\gamma fLR)\\ &\quad -\mu C_{\delta,\delta}(\frac{\partial\delta}{\partial x}(1/u'(w)-\tau)+(\delta+f)(-\frac{\partial w}{\partial x}u''(w)/u'(w)^2))\end{aligned} \tag{I.102}$$

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

$$\frac{\partial\delta}{\partial x}=\frac{(-1+\beta+\beta C_{\delta,x})u'(w)^3+\mu u'(w)^2(1+C_{\delta,x}+C_{\delta,a,x}\gamma u'(w)+C_{\delta,\delta,x}(\delta+f-(\gamma fLR+(\delta+f)\tau)u'(w)))-C_{\delta,\delta}(1+C_{\delta,x})\mu^2(\delta+f)u''(w)}{u'(w)^2(\mu(-C_{\delta,\delta,\delta}(\delta+f)-C_{\delta,\delta,a}\gamma u'(w)+C_{\delta,\delta,\delta}(\gamma fLR+(\delta+f)\tau)u'(w))+C_{\delta,\delta,}(-\beta u'(w)+\mu(-2+\tau u'(w))))+C^2_{\delta,\delta}\mu^2(\delta+f)u''(w)} \tag{I.103}$$

Setting $u'(w)=0$,

$$\frac{\partial\delta}{\partial x}=-\frac{1+C_{\delta,x}}{C_{\delta,\delta}} \tag{I.104}$$

This term is negative under the conditions of Corollary 4.1.

(ii) For $\partial\delta/\partial LR$, differentiating Equations (I.29) and (I.30) in p,

$$\frac{\partial w}{\partial LR} = (\mu C_{\delta,\delta}\frac{\partial\delta}{\partial f})/u'(w) \tag{I.105}$$

$$\begin{aligned}\frac{\partial w}{\partial LR} &= -b\frac{\partial\delta}{\partial LR}C_{\delta,\delta}-\mu\gamma C_{\delta,\delta,a}\frac{\partial\delta}{\partial LR}-\mu(C_{\delta,\delta,\delta}\frac{\partial\delta}{\partial LR})((\delta+f)(1/u'(w)-\tau)-\gamma fLR)\\ &\quad -\mu C_{\delta,\delta}(\frac{\partial\delta}{\partial LR}(1/u'(w)-\tau)-\gamma f+(\delta+f)(-\frac{\partial w}{\partial LR}u''(w)/u'(w)^2))\end{aligned} \tag{I.106}$$

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

$$\frac{\partial w}{\partial f}=\frac{C_{\delta,\delta}\gamma\mu fu'(w)^3}{(\mu(-C_{\delta,\delta}d(\delta+f)-C_{\delta,\delta}a\gamma u'(w)+C_{\delta,\delta,\delta}(\gamma fLR+(\delta+f)\tau)u'(w))+C_{\delta,\delta}(-\beta u'(w)+m(-2+\tau u'(w))))u'(w)^2+C^2_{\delta,\delta}\mu^2(\delta+f)u''(w)} \tag{I.107}$$

This term is positive under the conditions of Corollary 4.1.

(iii) For $\partial\delta/\partial f$, differentiating Equations (I.29) and (I.30) in LR,

$$\begin{aligned}\frac{\partial w}{\partial f} &= (\mu C_{\delta,\delta}\frac{\partial \delta}{\partial f})/u'(w) && \text{(I.108)}\\ \frac{\partial w}{\partial f} &= -\beta\frac{\partial w}{\partial f}C_{\delta,\delta} - \mu\gamma C\delta,\delta,a\frac{\partial \delta}{\partial f} - \mu(C_{\delta,\delta,\delta}\frac{\partial \delta}{\partial f})((\delta+f)(1/u'(w)-\tau)-\gamma fLR)\\ &\quad -\mu C_{\delta,\delta}(\frac{\partial \delta}{\partial f}(1/u'(w)-\tau)-\gamma f+(\delta+f)(-\frac{\partial w}{\partial f}u''(w)/u'(w)^2)) && \text{(I.109)}\end{aligned}$$

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

$$\frac{\partial w}{\partial f} = \frac{C_{\delta,\delta}\mu(-1+\gamma LRu'(w)+\tau u'(w))u'(w)^2}{(\mu(-C_{\delta,\delta}d(\delta+f)-C_{\delta,\delta}a\gamma u'(w)+C_{\delta,\delta,\delta}(\gamma fLR+(\delta+f)\tau)u'(w))+C_{\delta,\delta}(-\beta u'(w)+\mu(-2+\tau u'(w))))u'(w)^2+C_{\delta,\delta}^2\mu^2(\delta+f)u''(w)} \quad \text{(I.110)}$$

This term is positive or negative under the conditions of Corollary 4.1.□

Chapter II

Corporate Hedging and Incentives

1 Introduction

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]

Graduate textbooks generally warn managers on using derivatives without a clear statement of their objectives.

"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)

[1]Southwest hedged 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.

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)

This chapter explores a limiting case of Chapter I, specifically the case in which the cost of hedging C used earlier is zero. Formally, I assume that the manager can privately engage in any mean-preserving spread, and discuss how this will affect risk choices and the optimal contract. As expected, this limiting case allows for sharper characterization of the hedging choices and the optimal contract. Section 2 extends the notations of Chapter I to distributions that may involve mass points (since, as will be shown, such distributions may now be optimal). Section 3 analyzes the optimal contract and discusses how risk management affects the chosen level of effort. Section 4 extends the model to risk controls (e.g., a Value at Risk) and provides additional intuitions for linearity in the optimal contract.

2 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 $a \in [\underline{a}, \overline{a}]$ 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.

□ **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.[2] 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 the next sections; 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 and Technology

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 et al. 1993).

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: $u' > 0$, $u'' < 0$, $\psi(\underline{a}) = \psi'(\underline{a}) = 0$, $\lim_{a \to \overline{a}} \psi'(a) = +\infty$, $\psi'', \psi''' > 0$ except possibly at $a = \underline{a}$. I assume that the agent has limited liability and must be paid $w(y) \geq \underline{w}$ for all y. Finally, the contract must prescribe a minimum reserve utility equal to b.[3]

□ 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.

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

[2]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 $A = \int y dF(y|a)$.

[3]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 $p \in P$ (here: $\{(a, \hat{F}) \in [\underline{a}, \overline{a}] \cup \Gamma(a)\}$) and pays a cost $c(p)$ (here: $\psi(a)$).

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

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

□ **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 $\hat{F}(.)$. Finally, let $\theta = \inf X$ denote the maximum loss that the agent may make.[4]

3 Perfect Risk Management

I assume here that X is an interval including $[\underline{a}, \overline{a}]$ and the manager has access to all fair gambles and hedges and thus $\Gamma(a) = \{\hat{F}(.)/\int yd\hat{F}(y) \leq a\}$. Under perfect risk management, the choice of the manager over gambles or hedges is unrestricted provided it does not increase firm value. Perfect risk management can capture situations in which effort is observable but not contractible.[5] To ease the exposition, I assume for now that $b \geq u(\underline{w})$ (the case $b < u(\underline{w})$ is studied separately in Appendix A).

Example 1: A fund manager has access to a risk-free trading strategy (or perfect arbitrage) that, for each dollar invested, yields a certain return 10%. The manager can invest $\$10a$ but then incurs a cost equal to $\psi(a)$ (Berk and Green (2004) interpret such cost as management fees or transaction cost). 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.

[4] 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 $\int_{-\infty}^{+\infty} zdF_A(z) \geq \int_{-\infty}^{+\infty} zdF_B(z)$. 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.

[5] Even in settings where perfect risk management may seem too extreme, 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.

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 also Hart and Moore (1988)).

3. 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.

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

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.[6]

3. 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. 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.

Lemma 3.1. *An optimal contract must prescribe fully hedging.*

Lemma 3.1 is the main theoretical results supporting the results in this Section; it states that the agent should hedge away all the risk. The key observation is that any contract with speculation can be

[6] First-best effort would still be given by Equation (II.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.

replaced by another contract that elicits the same effort but elicits fully hedging. This is because fully hedging maximizes the usefulness of the output signal. Yet, in the context of a moral-hazard problem, it may seem surprising: at first sight, a deterministic compensation may seem incompatible with incentives to work hard.[7] To see why a deterministic compensation can provide 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 $y < \overline{y}$ and thus would receive an expected compensation strictly below W. Perfect risk management implies a constant wage on-equilibrium but not necessarily off-equilibrium.

3. 3 Incentive-Compatibility

Since the agent fully hedges (by Lemma 3.1), 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 II.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.

It turns out that that the agent does not choose first-best in this example. In Figure II.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.

[7] For example, Holmström (1979) shows that, in a class of moral hazard problems, risk-sharing is imperfect, which excludes a deterministic compensation when the principal is risk-neutral.

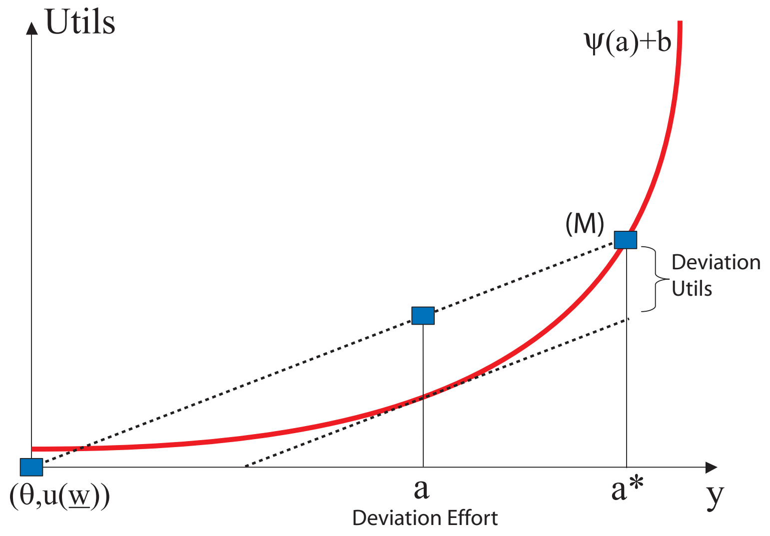


Figure II.1. Optimal Effort Choices

Proposition 3.1. *If the maximum downside is unbounded ($\theta = -\infty$), 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.[8] To give content to the contract design problem, therefore, I assume in the rest of the essay that θ is finite and meets the condition in Lemma 3.2.

Lemma 3.2. *There exists a contract eliciting effort $a > \underline{a}$ for an expected wage below w if and only if:*

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

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 mitigates

[8] 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.

incentives to speculate. Equation (II.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.

3. 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.[9]

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

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

To guarantee a unique solution to Equation (II.5), I assume that, for all a, $\psi'(a)/\psi(a) \leq 1/(a - \theta)$. 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 (II.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 be preferred by the principal to (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 $a - u^{-1}(\psi'(a)(a - \theta) + u(\underline{w}))$.[10] The (unique) solution to the first-order condition in this problem characterizes a^{***}:

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

Equation (II.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 re-

[9] 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.

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

quired compensation to avoid speculation, which increases the marginal cost of increasing effort by $\psi''(a^{***})(a^{***} - \theta) > 0$.

Proposition 3.2. *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 $b > u(\underline{w})$).

To begin with, second-best effort is increasing in θ. That is, a more important maximum loss tends to reduce total output because it makes speculation more attractive. 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 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.

In contrast, a greater limited liability works to reduce the effort elicited in the contract.[11] 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 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.[12]

The effect of a change in the outside option of the agent b on effort can be ambiguous. When contract (W^*, a^*) is elicited, elicited effort is decreasing in b. This is because the cost of effort in

[11] Note that in the standard agency model with limited liability (e.g., Innes (1990)) the relationship between elicited effort and the limited liability is more ambiguous.

[12] 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.

consumption units of effort increases when the agent is richer (as in Spear and Wang (2005)). When contract (W^{**}, a^{**}) is elicited, on the other hand, the effort is decreasing in b. As in Tirole (2006), this is because a greater personal wealth can be "pledged" and used to better punish the agent.

In terms of efficiency, effort is lower in second-best than first-best. While intuitively appealing, this property is not guaranteed in standard models with continuous effort (Holmström 1979) because the second-best effort choice will depend on the (global) informativeness of every distribution available to the agent. Here, incentives to speculate are greater when there is more value within the firm to speculate with (as captured by the left term in Equation (II.4)).[13]

The second-best contract may be Pareto-dominated by first-best (always in case (ii)). Note that in this model, first-best can always be attained if the principal controls the risk management decision. To see this, assume that the principal chooses 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 can 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.[14] It is therefore the joint interaction of effort and risk management that causes the efficiency loss. This aspect is consistent with recent regulatory changes with respect to the supervision of the risks taken by managers. The Sarbanes-Oxley Act and various implementation notes from the FASB and SEC require more 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.[15] In the financial industry, the Third Pillar of the Basel II accord increases the disclosure of financial institutions over their portfolio of investments.

4 Extensions

4. 1 Risk Controls

In the benchmark model, the principal only observes one signal on output and uses this signal to infer both actions and risk choices. In practice, however, the principal may also be able to impose other risk controls to prevent unobserved risk-taking. Formally, I discuss three types of risk controls:

[13]The other (well-known) effect is that more effort implies greater marginal cost of effort and thus increases the utility gains of a deviation.

[14]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).

[15]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.

(i) a maximum loss, (ii) a value-at-risk. Define now $f(.|a)$, the density of $F(.|a)$ (when the agent does not change risk), and assume that it has full support on X, it is differentiable in both arguments and satisfies the monotone likelihood ratio property ($f_a(y|a)/f(y|a)$ strictly increasing in y). For simplicity, assume here that b is sufficiently low so that only the incentive-compatibility condition binds (contract $(W^{*}**, a^{*}**)$).[16]

Maximum Loss

A maximum (hedging) loss control is defined as a bound $\theta' > \theta$ such that $F(.|a)$ and $\hat{F}(.)$ must coincide for any $y \leq \theta'$. In other words, the agent cannot hedge events with losses below this bound. This form of risk control generates the following trade-off. On one hand, it prevents the agent from raising the likelihood of very large losses, and thus decreases the informational rents that need to be paid (by Equation (II.5)). On the other hand, it also prevents the agent from fully-hedging, and thus makes the signal less informative on effort. In particular, if first-best effort can be elicited ($a = a^{*}$), such a risk control will always (weakly) reduce the profit of the principal.

Proposition 4.1. *Suppose θ is finite. There exists a risk control $\theta' > \theta$ that increases the profit of the principal. If $\theta = -\infty$, there exists a maximum loss such that $a > \underline{a}$ is elicited.*

Proposition 4.1 shows that small maximum risk controls are always desirable in second-best. In the model, the risk control makes it impossible to distinguish between low outcomes. However, the limited liability prevents the principal from using all the information contained in these outcomes, and thus this loss of information does not decrease efficiency. On the other hand, the risk control helps control the risk-taking choices of the agent. If θ' is set too high, however, one may reach regions in which it is optimal to pay the agent more than the limited liability and thus this loss of information may be costly to the principal.

Value at Risk

Directly controlling the support of potential outputs can be difficult in certain settings, in particular when the principal lacks the ability to sue the agent or establish precise risk controls. A possibly less intrusive risk management procedure is to use a Value at Risk constraint (VaR). The VaR approach is described in standard risk management textbooks (Jorion 2006), banking regulations and widely used by

[16]If both incentive-compatibility and participation bind, some of the results can be ambiguous, as risk controls can decrease the welfare of the agent.

financial institutions (e.g. JP Morgan RiskMetrics services). Formally, the VaR is defined as a probability $P \in (0,1)$ and V such that $\hat{F}(V) \leq P$: the probability that the hedged cash flow lies below V must be smaller than P. Assume that this constraint does not bind when there is no risk management so that $F(x|a) < P$ for all a.

Proposition 4.2. *Suppose θ is finite and let P be a fixed exogenous risk control parameter. There exists a Value at Risk (V, P) that increases the profit of the principal. In this case, $y = \theta$ (large loss) with probability P is elicited. If $\theta = -\infty$, however, no VaR makes $a > \underline{a}$ feasible.*

The VaR offers advantages that are similar to the maximum loss, although it allows the agent to remove some of the risk associated to $y < V$. Another aspect of the VaR is that, when it is effective, it induces the agent to strategically choose very large losses. In the model, these large losses are an ex-ante choice of the principal, willing to tolerate large risks to reduce informational rents. However, since the risk-taking is now under the control of the agent (and not given by an exogenous constraint) it no longer prevents excessive risk-taking when $\theta = -\infty$.[17]

4. 2 Partially Hedgeable Output

Although, in this Section, all risk can be removed at no cost, the results can be extended to the case in which only some of the risk can be removed. Formally, assume that the unhedged cash flow x can be written: $x = a + \epsilon + u$. The manager can manage risk to create any hedged cash flow $y = x + f(u)$, where $f(u)$ is an arbitrary function of u such that $\mathbb{E}(f(u)) = 0$. That is, the manager may not hedge ϵ. It is now convenient to denote $F(.|a)$ (with density $f(.|a)$) the distribution of $a + \epsilon$.

[18] In terms of the set of possible hedges $\Gamma(a)$, these assumptions can be summarized as: $\Gamma(a) = \{\hat{F}(.)/F(.|a)$ second-order stochastically dominates $\hat{F}(.)\}$. The analysis of the previous Section corresponds to the special case in which $F(.|a)$ is a point mass at $y = a$. The next statement is an immmediate corollary to Lemma 3.1.

Lemma 4.1. *If a is elicited, an optimal contract must prescribe $\hat{F}(x) = F(x|a)$.*

The intuition for Lemma 4.1 is that any contract such that $u(w(y))$ is convex will imply some risk management to concavify the compensation schedule, leading to strictly more compensation cost (and the

[17]Other problems related to the VaR are well-documented. For example, Artzner, Delbaen, Eber and Heath (1999) propose the expected tail loss as a better measure for risk. Another concern is that the VaR can be dynamically manipulated so that it is no longer binding. In our model, an expected tail loss (bounding the maximum expected tail loss instead of the probability) would make effort above $\underline{a}$ implementable while sharing the benefits of the VaR.

[18]The random variable ϵ may or may not depend on a and, as before, has mean zero.

same effort) as offering directly the concavification. Note that the statement implies an agent-analogue to the well-known informativeness principle. Under the informativeness principle, the principal should use any signal informative on the actions of the agent (Holmström 1979). Lemma 4.1 states that the agent should be given incentives to provide the maximum level of informativeness (by removing all noise) on the signal. As a benchmark, define the *no-hedging* problem as the problem in which $\Gamma(a) = \{F(.|a)\}$. Through the paper, it is also assumed, as in the standard model, that the optimal policies are smooth and differentiable.

Proposition 4.3. *The limited liability can only bind at $y = \theta$. If in the no-hedging problem, the limited liability binds at $y_0 > \theta$ then, in the problem with hedging, $w(y)$ is convex on some interval $[\theta, y']$ where $y' > \theta$.*

Corollary 4.1. *Suppose that u is either CARA or CRRA and $f_a(y|a)/f(y|a)$ is convex in y for all a, then $u(w(y))$ is linear in y.*

In the model, a contract linear in utility (and thus with a convex wage) helps remove incentives for the agent to alter risk and increases the expected compensation paid by the principal. This implies in particular that a binding limited liability can only occur at one point and, in the continuous case, the agent receives the lowest possible wage with probability zero. This finding is unlike in the case in which the limited liability does not bind; in fact, whenever the limited liability binds in the problem with no risk management, the optimal contract in the problem with risk management will feature convex portions to provide incentives to increase informativeness. Finally, Corollary 4.1 provides a simple condition such that the compensation is linear in utility for all y; namely, the informativeness content of good outcomes must increase sufficiently fast. This implies in particular that provided that either the likelihood ratio converges for y large or $u'(w)$ converges for w large, $w(y)$ becomes linear for y large. In this respect, the compensation schedule is convex and becomes linear for large outcomes, as in an option with some time value.

4. 3 Privately Observable Cost of Effort

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 (see also Appendix A for a more detailed analysis).[19]

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

$$(P') \qquad \max_{w(.)\geq \underline{w}, \hat{F}_c(.), a(c)} \int\int (y - w(y))d\hat{F}_c(y)dH(c) \tag{II.7}$$

s.t., for all c,

$$(a(c), \hat{F}_c(.)) \in \arg\max_{\tilde{a}, F(.)\in\Gamma(\tilde{a})} \int u^j(w(y))dF(y) - c\tilde{\psi}(\tilde{a}) \tag{II.8}$$

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.[20] 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.

Proposition 4.4. *As the utility of the agents becomes linear (i.e., j increases), any solution to (P') converges to a linear contract.*

Corollary 4.2. *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 if the principal could observe c.*

A linear contract is obtained as the unique solution in (P) with a small uncertainty about the cost

[19] This assumption ensures 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.

[20] 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.

of effort and almost risk-neutral agents.[21] 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.[22]

Corollary 4.3. *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 3.2 with a cost $\psi(a) = c_0\tilde{\psi}(a)$.*

I shall try to explain why linearity is optimal here. Uncertainty about the cost of effort implies that the output 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 3.1, 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 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 *performs better* at compensating the high-skilled workers while a concave compensation *is required* to avoid speculation by the low-skilled workers.

4. 4 Publicly Observable Information

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

[21] Note that by a minor notational change in the proof of Proposition 4.4, 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.

[22] 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.

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 3.1 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.[23]

Proposition 4.5. *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.

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

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 $\rho(x)$.

[23] 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.

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 $y - \theta$.[24] 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 4.6. *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.□

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.

[24] 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 $x \leq x_0$).

5 Appendix

5. 1 Appendix A: Low Reserve

I extend next the analysis of Section 2 to the case in which the limited liability is much *greater* than the outside option of the agent, i.e. $u(\underline{w}) > b + \psi(\overline{a})$ (versus $u(\underline{w}) \leq b$ previously).[25] 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 A. 1. *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^{*}.

Proof: 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 (II.9). When Equation (II.9) is true (resp. false), this term is positive (resp. negative), and thus $a^{***} > \underline{a}$ (resp. $a = \underline{a}$).□

Corollary A. 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:*

$$\theta > \underline{a} - \frac{u'(\underline{w})}{\psi''(\underline{a})} \tag{II.9}$$

In contrast to Proposition 3.2, 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 3.1 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. a symmetry restriction on the effect of changing the likelihood of one event on the marginal cost of other outcomes.

[25] The case in which $u(\underline{w}) \in (b, b + \psi(\overline{a}), b)$ is similar but involves also considering (W^{*}, a^{*}) and (W^{**}, a^{**}).

5. 2 Appendix B: Omitted Proofs

Proof of Lemma 3.1: 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: $\int_{y=a} dG(y) = \int (y-\theta)/(a-\theta)d\tilde{F}(y)$ and $\int_{y=\theta} dG(y) = \int (a-y)/(a-\theta)d\hat{F}(y)$. It follows that:

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

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

$$\begin{aligned}\int u(\hat{w}(y))d\tilde{F}(y) &= u(\hat{w}(a))\int_{y=a} d\tilde{F}(y) + u(\underline{w})(1 - \int_{y=a} d\tilde{F}(y)) \\ &< u(\hat{w}(a))\int_{y=a} dG(y) + u(\underline{w})(1 - \int_{y=a} dG(y))\end{aligned} \tag{II.10}$$

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 $w(.)$.

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

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

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

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 3.1: 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 $\int y d\hat{F}^n(y) \leq a^n$ 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:

$$\begin{aligned} \int u(w(y)) dG^n(y) &= \frac{a^n - \underline{a}}{n + a^n - \underline{a}} u(w(-n - y)) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int u(w(y)) d\hat{F}^n(y) \\ &\geq \frac{a^n - \underline{a}}{n + a^n - \underline{a}} u(\underline{w}) + (1 - \frac{a^n - \underline{a}}{n + a^n - \underline{a}}) \int u(w(y)) d\hat{F}^n(y) \qquad \text{(II.11)} \end{aligned}$$

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 3.2: 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 (II.4) must prescribe $\epsilon \leq 0$.□

Proof of Proposition 3.2: 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 (II.5) ensures that this is indeed the case.□

Proof of Proposition 4.1: To prove these results, Lemma 3.2 is extended to the case in which there is a maximum loss. Let a denote the effort of the agent and A denote the outcome chosen by the agent whenever $y > \theta'$. Note that when $\theta' > \theta$, $A > a$ (so that the mean is a). This yields:

$$F(\theta'|a) \int_{\theta}^{\theta'} f(y|a) y dy + (1 - F(\theta'|a)A) = a \tag{II.12}$$

Consider now a deviation to $a - \epsilon$; then, the agent will have to set probability p on θ' and probability $1 - p$ on A. As in Lemma 3.2, p is given by the fact that this choice must yield a mean $a - \epsilon$,

$$\int_{\theta}^{\theta'} f(y|a - \epsilon) y dy + (1 - F(\theta'|a - \epsilon))(p\theta' + (1 - p)A) \tag{II.13}$$

Reinjecting A from Equation (II.12) in this expression and solving for p:

$$p = \frac{\epsilon + \int_{\theta}^{\theta'} (f(y|a-\epsilon) - f(y|a))ydy}{a - \theta' + \theta' \int_{\theta}^{\theta'} f(y|a-\epsilon)dy - \int_{\theta}^{\theta'} f(y|a)ydy} \tag{II.14}$$

The expected utility achieved by the agent for this deviation is:

$$U = F(\theta'|a-\epsilon)u(\underline{w}) + (1 - F(\theta'|a-\epsilon))(pu(\underline{w}) + (1-p)u(W)) \tag{II.15}$$

Reinjecting p from Equation (II.14), differentiating with respect to ϵ and evaluating at $\epsilon = 0$,

$$\frac{W - \underline{w}}{a - \theta - \int_{\theta}^{\theta'} f(y|a)ydy}(1 - \int_{\theta}^{\theta'} f_a(y|a)ydy - (1 - (a - \theta - \int_{\theta}^{\theta'} f(y|a)ydy)) \int_{\theta}^{\theta'} f_a(y|a)dy) \geq \psi'(a) \tag{II.16}$$

The second term in Equation (II.16) captures the effect of the maximum loss constraint. It is then necessary to show that this term is greater than one. Rearranging this term, one must show that:

$$\frac{-\int_{\theta}^{\theta'} f_a(y|a)ydy}{\int_{\theta}^{\theta'} f_a(y|a)dy((a - \theta - \int_{\theta}^{\theta'} f(y|a)ydy))} < 0 \tag{II.17}$$

As θ' converges to θ, the right-hand side converges to $-\theta/(a-\theta)$, strictly negative. The second part of the statement follows immediately by noticing that the incentive-compatibility constraint is slack for θ' sufficiently small and $a > \underline{a}$ is elicited for any finite maximum loss controlled by the agent. □

Proof of Proposition 4.2: The first part of the statement is a Corollary to Proposition 4.1. Suppose the principal induces the agent to choose θ with probability P (binding the VaR) ; clearly, if the VaR is not binding in equilibrium, it would not be helpful in the contract design problem. It is easy to show that, in this case, the agent must be paid $\underline{w}$ when reporting θ. Next, note that this problem now maps to Proposition 4.1, by taking $F(.|a)$ to be a distribution that has a point mass at θ, and thus the same properties apply. The second part follows from the proof of Lemma 3.1, since the strategy with $\theta = -\infty$ prescribes choosing $\hat{F}$ such that the probability to report $y \neq a$ goes to zero (thus, the VaR would not bind).□

Proof of Proposition 4.3: The problem with hedging is given by the problem with no hedging subject to $u(w(y))$ concave. It is convenient here to solve for $\phi(y) \equiv u(w(y))$ and denote this last constraint as $\phi''(y) \leq 0$ (substituting $w(y)$ in the objective of the principal with $u^(- 1)(\phi(y)))$. I prove the first part of the statement. Suppose by contradiction that the limited liability binds for $y' \neq \theta$. In this case

$\phi'(y') = 0$ (or else limited liability would be violated at a neighborhood of y'). Therefore, for $y < y'$, $\phi'(y) \leq 0$. But, because $\phi''(y) \leq 0$ for all y, this implies that $\phi'(y) = 0$ for all y. This in turn implies that $w(y) = \underline{w}$, a contradiction to the participation. For the second part of the statement note that a binding limited liability at $y_0 > \theta$ implies, by the first part of the statement, that $\phi''(y) \leq 0$ is binding and thus $u(w(y))$ linear in y. This in turn implies that $w(y)$ is convex. Finally suppose that u is CARA or CRRA and $f_a(y|a)/f(y|a)$ is concave. Then, if the $\phi''(y) \leq 0$ does not bind for some y, the compensation is given by the standard first-order condition: $1/u'(w(y)) = \lambda + \mu \frac{f_a(y|a)}{f(y|a)}$, where λ and μ are the Lagrange multipliers in Holmström (1979). However, the assumptions made previously imply that $\phi(y)$ is convex, a contradiction. Therefore $\phi''(y) \leq 0$ must bind for all y and thus $w(y)$ is convex for all y.□

Proof of Proposition 4.4: The first part of the argument is similar to the proof of Lemma 3.1 (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 (II.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 (II.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.

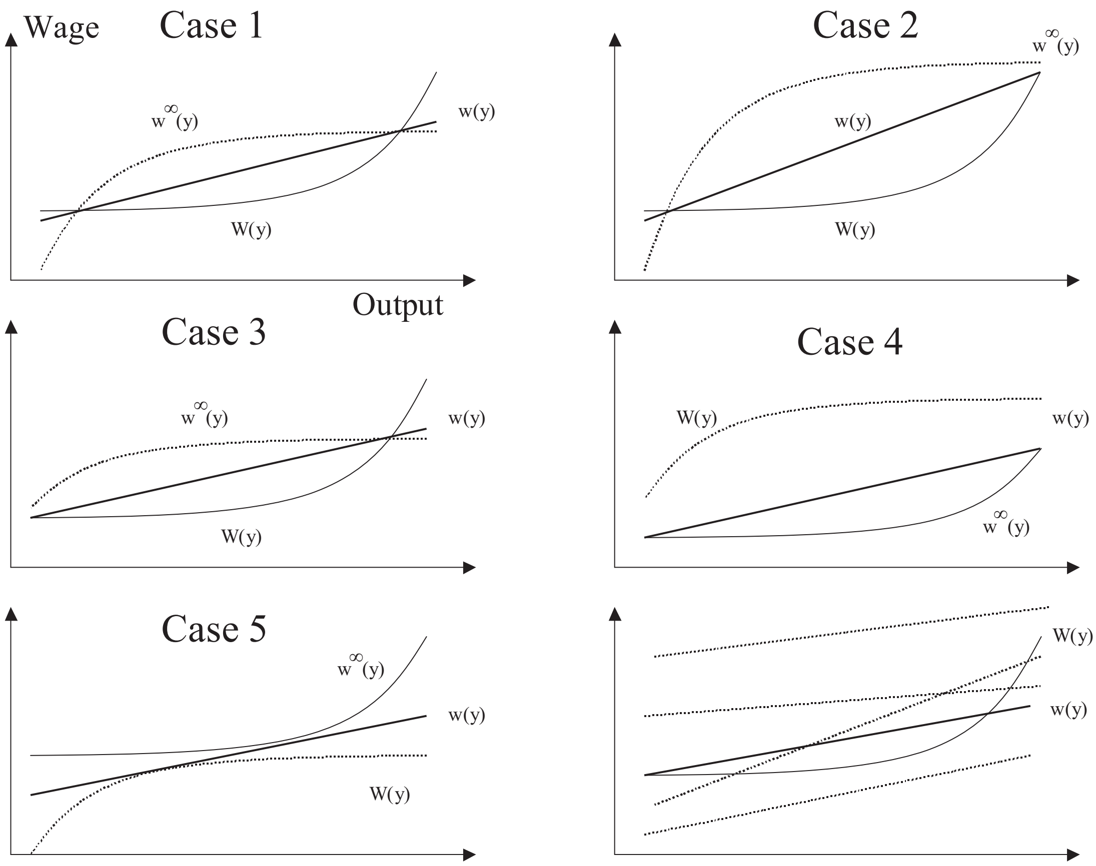


Figure II.2. Linear Contract - Cases

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 II.2. 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 4.2 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 II.2 (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 4.3, 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 4.5: 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.

$$\int g(x|a) \left(p(x)\theta + (1 - p(x))\rho(x)\right) dx \leq a \qquad (\tilde{\lambda} \tag{II.18}$$

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:

$$\frac{\partial L_1}{\partial p(x)} = g(x|a) \left(u(\underline{w}) - u(\phi(x)) - \tilde{\lambda}(\theta - \rho(x))\right) \tag{II.19}$$

From Lemma 3.1, 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:

$$-\psi'(a) + \int g_a(x|a)u(\phi(x))dx + \tilde{\lambda}\left(1 - \int g_a(x|a)\rho(x)dx\right) = 0 \tag{II.20}$$

I analyze now the problem of the principal. First, the principal can offer a contract such that Equation (II.18) 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 (II.18) 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.

$$a = \int \rho(x)g(x|a)dx \qquad (s_1) \tag{II.21}$$

$$\psi'(a) = \int g_a(x|a)u(\phi(x))dx + \tilde{\lambda}\left(1 - \int g_a(x|a)\rho(x)dx\right) \qquad (s_2) \tag{II.22}$$

$$b \geq \int g(x|a)u(\phi(x))dx - \psi(a) \qquad (s_3) \tag{II.23}$$

$$\tilde{\lambda} \leq \inf(u(\phi(x)) - u(\underline{w}))/(\rho(x) - \theta) \tag{II.24}$$

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 $x \in X'$,

$$\frac{\partial L_2}{\partial \rho(x)} = g(x|a)(-s_1 + s_2\tilde{\lambda}\frac{g_a(x|a)}{g(x|a)}) \tag{II.25}$$

This term is increasing in x and therefore $\rho(x)$ can be set equal to θ for $x \leq x_0$ and Equation (II.24) 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:

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

□

Chapter III

Optimal Incentives with Aggregate Risk

1 Introduction

According to conventional agency theory, the optimal incentive scheme should filter out risks that are not informative on the actions of the agent. This principle suggests that pay should not increase in response to aggregate shocks that are not under the control of the agent. Yet, relative performance evaluation represents a relatively small percentage of total managerial compensation. Murphy (1999) finds that some forms of compensation seem to filter out some movements in the stock price due to systematic shocks. However, this evidence concerns only cash compensation (which now represents only 50% of total pay), recent samples (not significant prior to 90), and not all industries (mostly for industrials). Analyzing total pay, Himmelberg and Hubbard (2000), Bertrand and Mullainathan (2001) and Jin (2002) find that executive wages depend on systematic risk factors (such as market indices or oil prices).[1]

The underlying reasons why incentive schemes seem to depend on these aggregate shocks are not fully understood. This problem is summarized by Garvey and Milbourn (2003): "Another important theoretical prediction comes from Holmström (1982), who suggests that the market component of a firm's returns should be removed from the compensation package since executives cannot affect the overall market by their actions and it is costly for executives to bear the related risks." Following a similar argument, Bertrand and Mullainathan (2001) explain that managers in their sample appear to be paid for luck rather than skill. Abowd and Kaplan (1999) ask: "Why should CEOs be rewarded for doing nothing more than riding the wave of a strong bull market?" They propose to improve managerial

[1] Using a different sample, Bertrand and Mullainathan find that a 1% increase in shareholder wealth due to oil prices causes a .3% increase in pay. Himmelberg and Hubbard show that a 1% increase in market return implies 0.38% increase in Cash compensation and 1.87% in Total compensation.

compensation by linking the exercise price of executive options to stock indices. Similarly, Bebchuk and Fried (2004) recommend to "encourage equity-based plans that filter out at least some of the gains in the stock price due to general market or industry movements" (p.190).[2]

This essay explores the design of efficient labor contracts in an Arrow-Debreu economy with systematic shocks. The main contribution of the essay is to describe the relationship between incentive pay and uninformative systematic shocks in a simple general equilibrium environment. I ask whether it is desirable to pay as a function of aggregate risk and/or monitor the agent's personal exposure to aggregate risk factors. In particular, I discuss the widespread idea that it is desirable to offer a wage that depends only on firm-specific risk (relative performance) and then let the agent hedge his exposure to aggregate risk through personal portfolio decisions.

The analysis shows that firms are risk-neutral to diversifiable risk but require a risk premium for wages correlated to the aggregate productivity shock. Because of these risk premia, it is cheaper to pay more during high-endowment states (expansions) than low-endowment states (recessions); this creates rational motives to offer contracts that are sensitive to systematic risk factors. In equilibrium, pay is increasing in both idiosyncratic performance and aggregate endowment. Further, in the class of decreasing absolute risk-aversion utility functions (DARA), pay-for-performance coefficients increase during expansions.

The model establishes that the optimal contract does not distort the agent's insurance against systematic shocks: in a competitive equilibrium, the agent has no incentives to trade on Arrow-Debreu securities with payoffs contingent on the aggregate shock. Indeed, given any zero-value change to the exposition of the contract to the systematic shock, the agent would trade against it and unravel it completely. In the class of hyperbolic absolute risk-aversion (HARA) utility functions, the Arrow-Debreu state prices do not depend on the informational asymmetry, even though markets are endogenously incomplete and agents are exposed to firm-specific idiosyncratic risk.

In general, the optimal contract cannot in be replicated using only relative performance pay (i.e., pay based only on the firm's idiosyncratic risk) and letting the manager trade. Because neither the systematic or idiosyncratic shock can be fully insured, it is typically desirable to condition the pay-for-performance coefficient on the realization of the systematic shock.

Related Literature

[2] One possible explanation is that the moral hazard problem depends on the business cycle. An example of this idea is developed by Oyer (2004) who argues that the outside option of the manager may vary with the business cycle. Himmelberg and Hubbard (2000) argue that CEO talent becomes more important during expansions and provide some cross-sectional support for their theory.

This essay contributes to several areas of the agency theory literature, which analyze the effect of common shocks on incentive schemes. Following Holmström (1982), a set of papers considers the optimal risk-sharing arrangement when members of a team face a common shock. In these papers, the optimal mechanism is constrained by a budget-balancing constraint and the common shock must be allocated to all agents. Holmström argues that the common shock may be filtered out, as argued in the relative performance literature, but only if there exists a risk-neutral budget-breaker. In the risk-averse case, Rasmusen (1987) shows that there are efficient contracts in which all but one agent are punished.[3]

The main difference between these papers and this one is that the problem is stated as a firm optimization in a decentralized Arrow-Debreu economy similar to Debreu (1972) with state-contingent securities, and not from the perspective of a centralized mechanism. Still, both approaches present similarities and Appendix B shows that, under certain conditions, the optimal centralized allocation can be decentralized as a competitive equilibrium. The decentralized approach has several additional advantages. First, it links the moral hazard problem to observable risk premia on financial assets (which are derived from state prices) and accommodates private trading by the agent in a straightforward manner. Second, the Arrow-Debreu economy more closely describes the problem of many firms offering compensation contracts competitively. In comparison, the more elaborate mechanisms used in the team literature are typically interpreted in the context of one firm with many workers.

A related set of papers considers moral hazard problems in which a firm offers a contract to multiple agents with a correlated performance signals. Changing how common shocks affect the informativeness of output on effort, several authors find that pay can be increasing or decreasing in the performance of the group (Nalebuff and Stiglitz (1983), Green and Stokey (1983), Magill and Quinzii (2004), Luporini (2006)). In this paper, pay will be increasing in the common shock even if common shocks are completely uninformative on effort. Other papers discuss whether relative performance evaluation should filter out industry risk (Aggarwal and Samwick (1999), Spagnolo (2000), Raith (2003), Cuat and Guadalupe (2005)). While these arguments are well-taken, this essay focuses on systematic market-wide risk, as for example movements in stock market indices. One advantage of the current approach is that it does not require aggregate shocks to be informative on individual effort and, thus, it is more parsimonious in this dimension of the problem.

Finally, several papers discuss how to contract on systematic shocks in partial-equilibrium settings. Margiotta and Miller (2000) show that, under constant absolute risk-aversion, a multi-period moral

[3]Other papers considering efficiency in partnership settings include Legros and Matthews (1993), Andolfatto and Nosal (1997), Miller (1997), Strausz (1999), Nandeibam (2002), Huddart and Liang (2003), Battaglini (2005) and Kirstein and Cooter (2006).

hazard contract can be replicated with a combination of idiosyncratic risk and private hedging. Their essay is the first to show when the hedging decision can be separated from the labor contract.[4] In the context of the Linear-Exponential Normal model (LEN), Jin (2002) and Garvey and Milbourn (2003) study the optimality of relative performance evaluation. Jin argues that the manager can trade away any unwanted market risk. However, in his model, while paying as a function of market risk is optimal, it would be optimal as well to filter out the aggregate shock and let the agent trade. In the second paper, the authors assume that adjusting exposure to the market portfolio is costly for the agent and the firm; and thus the dependence of wages on market risk is exogenously assumed rather than proved from primitive assumptions on risk-sharing. The contribution of this essay to this literature is two-fold: (i) I extend the framework to an agency setting with any concave utility functions and optimal non-linear contracts, (ii) in this extended framework, I derive formal conditions under which relative performance is suboptimal, even in the absence of capital market frictions.

2 The Model

To keep the model as transparent as possible, I initially focus on a simplified version of the model with identical agents and firms, and a single time period - these assumptions are unimportant for the main result and will be relaxed in Section 7. Assume that the economy consists of a continuum of identical firms (or principals) and agents with mass normalized to one and indexed by $k \in [0, 1]$. In the model, each firm and consumer will be small as compared to the economy; this assumption is meant to capture the idea that individual outcomes have no effect on aggregate outcomes.

□ **Consumers**

Each agent receives a utility $u_k(c) = u(c_k) - \psi(a_k)$ where c_k is final consumption and $a_k \in [\underline{a}, \overline{a}]$ is a privately observed effort decision. Assume that $u(.)$ is twice-differentiable, strictly increasing and strictly concave, and satisfies standard Inada conditions. The cost of effort $\psi(.)$ is convex, increasing and differentiable with $\psi(0) = \psi'(\underline{a}) = 0$ and $\lim_{a \to \overline{a}} \psi'(a) = +\infty$.

□ **Production**

[4] A more distantly related paper is Grossman, Hart and Maskin (1983) who study a problem in which firms are informed about productivity and workers are not. They show that this asymmetric information generates cyclical unemployment. This paper, on the other hand, is focused on a moral hazard problem and on the shape of the optimal contract, not on the cyclical properties of unemployment.

Each firm requires labor from a single agent and, without loss of generality, assume that firm k employs agent k. A firm produces an output αy_k, where $\alpha \geq 0$ is drawn from an absolutely continuous distribution with a differentiable density $h(.)$ and full support on $[\underline{\alpha}, +\infty]$ (where $\underline{\alpha} > 0$). It is convenient to interpret α as a labor productivity shock common to all firms. I think about this static equilibrium model as representing two periods, one prior to the realization of α and (y_k) and one after their realization; a high realization of α indicates in this context an expansion while a low one indicates a recession.

The distribution of y_k depends on the effort of the employed agent, a_k. Conditional on a_k, the distribution of y_k is denoted $F(.|a_k)$ and has a density $f(.|a_k)$, twice-differentiable in a_k and with compact positive-valued support $[\underline{y}, \overline{y}]$. Since the objective of the analysis is to explain the dependence of pay on systematic risk, it is assumed here that $y_k|a_k$ does not depend on α.[5] To avoid well-known situations in which relative performance pay is desirable, assume as well $(y_k)_{k' \neq k}$ is almost always independent of y_k. Following this restriction, the main focus of this essay will be exclusively *systematic* risk and not, for example, diversifiable industry or sectorial risk (since in these cases y_k may be correlated to other $y_{k'}$ in the same industry).

Assume that $f_a(y|a)/f(y|a)$ is strictly increasing in y (monotone likelihood ratio). This assumption implies in particular that for any $a > a'$, $y|a$ first-order stochastically dominates $y|a'$. To simplify notations, the mean of $y|a$ is normalized to a.

□ State-Contingent Assets

In this economy, an *aggregate state of the world* is defined as a realization of α. Let $q(\alpha)$ be the price of an asset paying one unit of good when the aggregate state is α and zero else. The state prices are normalized to $\int q(\alpha)d\alpha = 1$.

Let $\updownarrow$ be the Lebesgue measure on $\mathbb{R}$. For any possible realization of α and $(y_k)_{k\in[0,1]}$, $\int \alpha y_k d\updownarrow(k)$ exists and represents total output available in the economy. Conditional on efforts $(a_k)_{k\in[0,1]}$, the law of large numbers imply that $\int \alpha\mathbb{E}(y_k|a_k)d\updownarrow(k) = \alpha \int a_k d\updownarrow(k)$ is available in the economy. The state price $q(\alpha)$ can also be defined as an asset paying one unit of good when the aggregate endowment in the economy is $\alpha \int a_k dk$.[6]

□ Labor Contracts

[5]Magill and Quinzii (2004) relax this assumption by indexing the distribution $f(.|a_k)$ by the common shock and describe further reasons to pay as a function of the common shock.

[6]Note that from the perspective of any single agent in the economy, $\alpha \int a_k dk$ is not controlled; and thus every agent is equally informed on the realization of the aggregate endowment.

Firms can offer state-contingent contracts. Firm k's labor contract is denoted $L^k = (a_k, w(\alpha, y_k))$ and prescribes the effort to be chosen by the agent and the wage to be paid conditional on each state. Note that the wage may depend on the common shock α and the firm-specific shock y_k.[7] A contract L^k is feasible and incentive-compatible if it meets the following two constraints.

First, it must be individually rational for agent k to accept the contract. Let R denote the utility achieved by agents when contracting with other firms in the economy. Then, the wage offered by firm k must satisfy:

$$\int\int h(\alpha)u(w_k(\alpha, y_k))f(y_k|a_k)dy_k d\alpha - \psi(a_k) \geq R \tag{III.1}$$

Second, the choice of a_k is not observable by the firm. Thus, the recommended effort a_k in the contract must be incentive-compatible for the agent. Taking the first-order condition on the problem of the agent, the choice of a_k must satisfy the following Equation.[8]

$$\int\int h(\alpha)f_a(y_k|a_k)u(w_k(\alpha, y_k))dy_k d\alpha = \psi'(a) \tag{III.2}$$

□ **Firm's Problem**

The objective of firms is to maximize the value of their production plan. Formally, for a production plan with $\pi(\alpha)$ expected units of good in each state α, its value is obtained as follows: $\int q(\alpha)\pi(\alpha)d\alpha$, that is the summation of the number of units produced $\pi(\alpha)$ in each aggregate state weighted by the Arrow-Debreu price $q(\alpha)$ associated to this aggregate state.[9]

In the context of this model, taking state prices $q(.)$ as given, the value to firm k of a feasible contract L^k is:

$$V(L^k) = \int\int (\alpha y_k - w_k(\alpha, y_k))f(y_k|a_k)q(\alpha)dy_k d\alpha \tag{III.3}$$

The optimal contract L^k is then given by a solution to the following problem.

$$(P_k) \quad \max_{L^k} \quad V(L^k)$$

$$\text{s.t.} \qquad \text{(III.1) and (III.2)}$$

[7]For obvious reasons, there would be no purpose in offering a wage that depends on the firm-specific shock of other firms.

[8]The validity of the first-order approach in this setting is established in the Appendix; it holds under the same conditions as in Rogerson (1985) (monotone likelihood ratio of convexity of the distribution function).

[9]It is well-known that firms will behave as risk-neutral to any shock that does not affect the aggregate endowment (see Duffie (2001) for references).

□ **Competitive Equilibrium**

Firms are competitive and make zero profit.[10] In each aggregate state, each firm will offer a wage $w_k(\alpha, y_k)$, for a total consumed $\int\int w_k(\alpha, y_k) f(y_k|a_k) dy_k dk$. This must be equal to the total endowment in the economy $\alpha \int a_k dk$. For all α,

$$\int\int w_k(\alpha, y_k) f(y_k|a_k) dy_k dk = \alpha \int a_k dk \tag{III.4}$$

The equilibrium is formally defined below.

Definition 2.1. *A (symmetric) competitive equilibrium is* $\Gamma = (a, w(.,.), q(.), R)$ *such that:*

1. $(a, w(.))$ *maximizes* (P_k) *for all* k.
2. *Equation* (III.4) *holds for all* α.

A first-best competitive equilibrium is defined in the same manner but omitting Equation (III.2), the incentive-compatibility, in the contract design problem (P_k). In the rest of the analysis, the index k will be omitted to simplify notations. The main focus of this essay is the analysis of the optimal contract, and thus a competitive equilibrium is assumed to exist and feature a contract $w(.,.)$ that is a smooth function of both parameters.[11]

3 Analysis of the Contract

3. 1 First-Best

As a benchmark to the analysis, the solution to the first-best economy is formally stated.

Proposition 3.1. *A first-best competitive equilibrium is Pareto-efficient and satisfies (up to sets with*

[10] The assumption of zero profits can be justified by equilibrium reasoning. Suppose that firms may make positive profits and are equally owned by all agents. When a contract is signed, agents transfer their assets to the firm and are given the contractual consumption $w(\alpha, y)$ (the firm receiving the dividends from the agent's assets). In this case, equilibrium dictates that total consumption equal to total endowment in each state (Equation (III.4)). This in turn implies zero profit for firms.

[11] Existence is discussed in the general equilibrium literature (see for example Citanna and Villanacci (2002) and other referenced papers). With continuous outcomes, it is well-known that existence of an optimal contract may fail even in partial equilibrium setting (Mirrlees 1999). These issues, however, concern conceptually different questions from those discussed in this paper.

zero probability):

$$\psi'(a) = \int \alpha h(\alpha) u'(\alpha a) d\alpha \tag{III.5}$$

$$w(\alpha, y) = \alpha a \tag{III.6}$$

$$q(\alpha) = \frac{h(\alpha)u'(\alpha a)}{\int h(x)u'(xa)dx} \tag{III.7}$$

$$R = \int\int h(\alpha) f(y|a) u(\alpha a) dy d\alpha - \psi(a) \tag{III.8}$$

The first-best solution to the model is such that the disutility of effort is equal to the gains from greater consumption in all states. It is a simple implication of the first welfare theorem in economies with production (Debreu 1972) (since firms are producing state-contingent units of good). Further, the competitive equilibrium is Pareto efficient and prescribes complete insurance against idiosyncratic risk. On the other hand, the systematic shock is not insured and each agent receives a wage αa.

3. 2 Second-Best Contract

In second-best, the firm must be mindful of the need to provide incentives to the agent. The next Proposition establishes the optimality of providing some incentives regardless of the state prices or cost of effort.

Proposition 3.2. *A competitive equilibrium must be such that* $a > \underline{a}$.

While proved here in the presence of systematic risk, this property is similar to Shavell (1979) (Proposition 2, p. 59). The firm can increase effort at almost no cost by shifting from perfect insurance to a pay that depends on both idiosyncratic (to induce effort) and systematic risk (to provide insurance).

Let λ (resp. μ) denote the multiplier associated to Equation (III.1) (resp. Equation (III.2)) in Problem (P). Taking the first-order condition with respect to $w(\alpha, y)$ yields the following characterization of the optimal contract:

$$u'(w(\alpha, y)) = \underbrace{\frac{q(\alpha)}{h(\alpha)}}_{G(\alpha)} \underbrace{\frac{1}{\lambda + \mu \frac{f_a(y|a)}{f(y|a)}}}_{H(y)} \tag{III.9}$$

Because $a > \underline{a}$ is elicited (Proposition 3.2), the wage cannot be decreasing for all y and thus the multiplier μ must be strictly positive. In Equation (III.9), the marginal utility of the agent can be decomposed into an aggregate component $G(\alpha)$ and an idiosyncratic component $H(y)$. The aggregate component captures risk-aversion to aggregate risk and is driven by the common labor productivity shock α. The idiosyncratic component captures the role of incentives and its interpretation is similar to a standard agency model with no systematic risk (Holmström 1979).

Proposition 3.3. *$w(\alpha, y)$ is strictly increasing in α and y.*

The agent is paid more in good states of the world. This is consistent with evidence that, when controlling for firm-specific performance, contracts depend on systematic risk factors. To better understand the statement, it is helpful to reframe the result in intuitive terms starting from two perspectives. First, the Arrow-Borch conditions for perfect risk-sharing reveal that, in situations such that firm and agents are risk-averse to systematic shocks, the optimal contractual arrangement should *not* provide perfect insurance and wages should depend on common risk factors. Here, the risk-aversion of the firm is captured by $q(\alpha)/h(\alpha)$ decreasing. Second, from a general equilibrium perspective, the market clearing condition makes perfect insurance *at any price* impossible. The aggregate shock is effectively allocated to the agents in the economy as consumption risk.[12]

3. 3 Performance Pay

Since the wage is affected by aggregate shocks, the optimal performance-pay coefficients may also depend on the realization of the aggregate shocks. The next Proposition establishes the relationship between the sensitivity of the agent's pay to idiosyncratic risk and the state of the economy.

Proposition 3.4. *If absolute risk aversion is decreasing (resp. increasing), then the performance-pay coefficient (w_y) is increasing in α.*

In the class of utility functions with decreasing absolute risk aversion (DARA), performance pay should increase during expansions and decrease during recessions. Intuitively, the optimal contract should prescribe more incentive pay in states with lower risk-aversion. This implies in particular more pay for performance in states such that agents are relatively wealth which are, by market clearing, states with high productivity.

This property of the optimal contract can help rationalize a long-standing puzzle regarding the use of option compensation. Many observers have noted that, if the market decreases, the pay for performance sensitivity of executives paid with options also decreases (Murphy (1999),Abowd and Kaplan (1999)). This is because, when the option is out of the money, its delta (i.e., the derivative of its price

[12]Note that the market clearing condition implies only that the expected wage should be increasing in α, which is weaker than the claim in Proposition 3.3.

with respect to the stock price) becomes small. This view implicitly assumes that pay for performance coefficients should be disconnected from factors that are not under the control of the agent (such as aggregate market movements). In contrast, Proposition 3.4 shows that a performance sensitivity linked to aggregate risk can be desirable because of its ex-ante risk-sharing properties. Indeed, consistently with the DARA class of utility functions, the performance-pay of agents paid with options should decrease during a recession. [13]

3. 4 State Prices

The previous Propositions investigate how state prices (and aggregate risk) can affect the contract design problem. Conversely, this Section explores the effect of informational frictions on risk premia. To begin with, note that the first-best state prices (as given by Equation (III.7)) correspond to the state prices in a standard representative agent framework. In second-best, however, agents will face idiosyncratic wage shocks and thus aggregation into a representative agent may potentially fail. The next Proposition explores this question in greater details.

Proposition 3.5. *Suppose that $u(.)$ is in the HARA class (i.e., $u(x) = \frac{\gamma}{1-\gamma}(b_1 + b_2 x)^{1-\gamma}$). Then, in a competitive equilibrium with moral hazard such that a is elicited, state prices must be given by:*

$$q(\alpha) = \frac{h(\alpha)u'(\alpha a)}{\int h(x)u'(xa)dx} \tag{III.10}$$

Corollary 3.1. *In the HARA class, conditional on the elicited effort, the state prices do not depend on the informational friction (i.e., same as in first-best).*

The class of hyperbolic absolute risk-aversion (or HARA) utilities includes as a special case quadratic, constant absolute risk-aversion (CARA) and constant relative risk-aversion (CRRA) utility functions. Proposition 3.5 establishes that, for a wide class of utility functions, the informational friction does not affect the state prices and risk premia in the economy. Conditioning on effort, the state prices are the same as in an economy with a single representative agent receiving an endowment αa and hedging optimally.[14]

[13] It may also be argued that the δ of the option will also decrease following positive market shocks (since the option becomes in the money). However, it is widely documented that executives tend to exercise such options prior to maturity (Heath, Huddart and Lang 1999). This may be a reason why the presence of in-the-money options after stock market increases is viewed as less problematic in the managerial literature.

[14] If effort is assumed to take only two values, there are therefore cases such that first-best and second-best state prices exactly coincide. For the more general case with continuous effort studied here, the second-best may affect state prices only via aggregate effort.

This result contrasts with the standard intuition that aggregation should fail under incomplete markets (here, uninsurable labor income). A preliminary intuition is that the optimal contract does not distort the agent's exposure to systematic shocks. Thus, from the perspective of systematic shocks, an idiosyncratic shock is the same as an exogenous wealth shock. With HARA utility functions, two agents with different wealth should make the same choice over the composition of a portfolio of risky assets (for a general proof, see the two-fund separation theorem of Stiglitz and Cass (1970)). As a result, the idiosyncratic shock becomes irrelevant to the determination of the risky portfolio to be purchased. The result confirms prior results in the context of more specified agency problems and constant absolute risk-aversion (Margiotta and Miller (2000), Ou-Yang (2005)).[15] Note that this property would not hold for any arbitrary (not optimal) wage $w(\alpha, y)$ and thus it is a characteristic of the fully-optimal solution to the moral hazard problem.

Corollary 3.2. *In the HARA class, there exists $K > 0$, a constant that does not depend on α or y, such that:*

$$w_\alpha(\alpha, y) = KH(y)\frac{u''(\alpha a)}{u''(w(\alpha, y))} \tag{III.11}$$

In the HARA class of utility functions, the coefficient of pay for systematic risk can be captured by the ratio of the curvature of the utility function evaluated at: (i) the consumption of the representative agent (receiving αa), (ii) the consumption of the agent (receiving $w(\alpha, y)$). Here, an agent who is more successful should be given more exposure to the systematic risk factor. In other terms, a high idiosyncratic performance magnifies the dependence of consumption on the systematic shock.

3. 5 An Example

The example of logarithmic utility functions is developed in more details to illustrate how systematic risk affects the optimal contractual arrangement. Suppose here that $u(x) = x^{1-\gamma}/(1-\gamma)$ where $\gamma > 0$ (if $\gamma = 1$, $u(x)$ is assumed to be logarithmic). By Equation (III.9), the inverse of the marginal utility of the agent will be linear in the likelihood ratio and the ratio of the objective probabilities to the state prices.

$$w(\alpha, y)^\gamma = \frac{h(\alpha)}{q(\alpha)}(\lambda + \mu\frac{f_a(y|a)}{f(y|a)}) \tag{III.12}$$

Therefore:

$$w(\alpha, y) = (\frac{h(\alpha)}{q(\alpha)})^{1/\gamma}(\lambda + \mu\frac{f_a(y|a)}{f(y|a)})^{1/\gamma} \tag{III.13}$$

[15] Kahn (1990) and Kocherlakota (1994) assume that the agent may consume prior to and after the resolution of aggregate uncertainty and find effects of the informational friction on risk premia. By focusing here only on insurance against aggregate uncertainty, one is able here to attribute this dependence to intertemporal substitution effects and not to risk-aversion.

Pre-multiplying both sides of this Equation by $f(y|a)$, integrating with respect to y and using the market clearing conditions yields that:

$$\alpha a = \frac{h(\alpha)}{q(\alpha)}^{1/\gamma} \int f(y|a)(\lambda + \mu \frac{f_a(y|a)}{f(y|a)})^{1/\gamma} dy \tag{III.14}$$

One can then substitute $\frac{h(\alpha)}{q(\alpha)}^{1/\gamma}$ from Equation (III.14) in Equation (III.13).

$$w(\alpha, y) = \alpha a \frac{(\lambda + \mu \frac{f_a(y|a)}{f(y|a)})^{1/\gamma}}{\int f(y'|a)(\lambda + \mu \frac{f_a(y'|a)}{f(y'|a)})^{1/\gamma} dy'} \tag{III.15}$$

The wage can thus be written as a share of the aggregate endowment αa; this share is greater for outcomes that are more informative about effort. In addition, a greater risk-aversion γ reduces the increase in this share for a greater idiosyncratic performance, as is intuitive.

With logarithmic utilities, the wage can be simplified as: $w(\alpha, y) = \alpha a(1 + \frac{\mu}{\lambda} \frac{f_a(y|a)}{f(y|a)})$, that is one unit of the market portfolio as well as a bonus that depends on performance. Then, when the likelihood ratio is zero, the agent receives exactly the aggregate endowment.

In the logarithmic case, the agent's participation constraint and incentive-compatibility can be further simplified as follows:

$$\int \ln(\alpha a) h(\alpha) d\alpha + \int \ln(1 + \frac{\mu}{\lambda} \frac{f_a(y|a)}{f(y|a)}) f(y|a) dy - \psi(a) \quad = \quad R \tag{III.16}$$

$$\int f_a(y|a) \ln(1 + \frac{\mu}{\lambda} \frac{f_a(y|a)}{f(y|a)}) dy \quad = \quad \psi'(a) \tag{III.17}$$

To state the objective of firm k, it is now important to distinguish the average effort in the economy, a, from the effort chosen by firm k, a_k. Denoting $\phi(y) = (1 + \frac{\mu}{\lambda} \frac{f_a(y|a)}{f(y|a)})$ (a function that only depends on y), the value of firm k can be written:

$$\begin{aligned} V(L^k) &= \int \int q(\alpha)(\alpha a_k - \alpha a \phi(y_k)) f(y_k|a_k) dy_k d\alpha \\ &= \int q(\alpha) \alpha d\alpha \int (a_k - a\phi(y_k)) f(y_k|a_k) dy_k \end{aligned}$$

Let $R' = R - \int \ln(\alpha) h(a\alpha) d\alpha$ denote a normalized reserve utility. One can then rewrite a reduced contract design problem eliminating the dependence on systematic risk.

$$(P_2) \qquad \max_{\phi(.),a_k} \int (a_k - a\phi(y_k))f(y_k|a_k)dy_k$$

s.t.

$$\int f(y_k|a)u(\phi(y_k))dy_k - \psi(a_k) = R' \tag{III.18}$$

$$\int f_a(y_k|a_k)u(\phi(y_k))dy_k = \psi'(a_k) \tag{III.19}$$

For an aggregate effort a and a (net) cost of labor R', the firm does not need to know the distribution $h(.)$ of the systematic component or the state prices $q(.)$. Thus, with Logarithmic utility functions, the systematic shock has only a simple scaling effect on payments from the contract: the firm will "inflate" the wage by the total endowment.

4 Hedging Problem

4. 1 Private Hedging

In the optimal contract as defined here, the agent cannot hedge outside of the contract. As a result, the model constrains the agent to hedge *through* the labor contract. This assumption, while convenient as a solution technique, is not innocuous when interpreting the findings in a practical context. Proponents of filtering out systematic risk would generally argue that it is desirable only if the agent can privately hedge outside of the contract (or else, holding the market portfolio would generate diversification benefits). In other words, proving the dependence of consumption on aggregate risk may not be sufficient to make a case against pure relative performance evaluation. In addition, it must be established that the optimal contract cannot be replicated using a pure relative performance payment but letting the agent freely trade on capital markets (although, for obvious reasons, not on the firm's idiosyncratic risk).

The notion of private hedging is now formally defined. Let $\eta(\alpha)$ be the number of Arrow-Debreu assets purchased by the agent. To avoid unraveling of the firm's idiosyncratic risk, agent k may not trade on y_k. In the presence of hedging, the participation of an agent incorporates the return from the Arrow-Debreu assets purchased by the agent.

$$\int\int h(\alpha)u(w_k(\alpha, y_k) + \eta(\alpha))f(y_k|a_k)dy_k d\alpha - \psi(a_k) \geq R \tag{III.20}$$

Since hedging is private, the incentive-compatibility of the agent incorporates now effort and hedging. That is, for a given contract, the agent will be facing the following problem.

$$(Q_k) \qquad \max_{\tilde{\eta}(.),\tilde{a},A} \int\int h(\alpha)u(w(\alpha,y_k)+\tilde{\eta}(\alpha)-A)f(y_k|\tilde{a})dy_k d\alpha - \psi(\tilde{a})$$

$$\text{s.t.} \qquad \int q(\alpha)\eta(\alpha)d\alpha = A$$

The constraint on (Q_k) is a budget constraint and states that the agent must pay the portfolio of assets at its current state prices. A value-maximizing contract is given by $(L^k, \eta_k(.))$ and maximizes $V(L^k)$ subject to participation and incentive compatibility.

$$(S_k) \qquad \max_{L^k,\eta_k(.)} \quad V(L^k)$$

$$\text{s.t.} \qquad \text{(III.20) and } (a_k,\eta_k) \text{ maximizes } (Q_k)$$

Finally the market clearing must reflect private purchases by the agent, i.e. for all α,

$$\int (w_k(\alpha,y_k)+\eta_k(\alpha))dk = \alpha \int a_k dk \tag{III.21}$$

Definition 4.1. *A (symmetric) competitive equilibrium with hedging is* $\Gamma^s = (a, w(.), \eta(.), q(.), R)$ *such that:*

1. $(a, w(.), \eta(.))$ *maximizes* (S_k) *for all* k.
2. *Equation* (III.21) *holds for all* α.

A first-best competitive equilibrium with hedging is defined similarly by omitting a_k in the maximization (Q_k). A relative performance payment is a wage $w(.,.)$ that depends only on y but not α. Note that, because the agent may now hedge outside of the contract, a wage that does not depend on α is not necessarily incompatible with market clearing. In short-hand, a competitive equilibrium with hedging will now refer to the equilibrium concept in which the agent can privately hedge (Definition 4.1), while a competitive equilibrium will refer to the equilibrium concept in which the agent may not hedge outside of the contract (Definition 7.2).

4. 2 Incentive-Compatibility

In managerial contexts, the investment of a manager is his/her own stock is heavily controlled as the agent would prefer to unravel the idiosyncratic factor on financial markets (e.g., by short-selling company stock). A natural concern is whether the same caveats apply to the dependence of w on α: Should the firm contractually commit the agent not to offset the dependence of the wage on α? The next Proposition establishes that such constraints are unnecessary when firms offer an optimal contract.

Proposition 4.1. *Suppose* $\Gamma = (w(.,.), a, q(.), R)$ *is a competitive equilibrium (i.e., the agent cannot privately hedge), then* $\Gamma^s = (w(.,.), \eta(.), a, q(.), R)$ *with* $\eta(\alpha) = 0$ *for all* α *is a competitive equilibrium with hedging.*

The intuition at play in Proposition 4.1 is similar to that underlying Proposition III.10. In an optimal contract where the agent cannot hedge, the principal should not distort exposure to systematic risk. As a result, even if the agent were to be given the option to trade state-contingent assets, it would not be optimal to trade.[16]

To further illustrate the result, it is useful to revisit a widespread misconception in the relative performance evaluation literature. Describing previous work in the area, Bebchuk and Fried (2004) explain: "compensation for sector or market increases exposes executives to market risk. Given that agents are risk-averse and therefore value risky compensation less, it would be cheaper (...) to give agents indexed options and cash equal to the value of the market component of their options" (p.157).[17] In fact, as Proposition 4.1 shows, doing so in the manner suggested would require the *exact* same monetary expense for the firm and, eventually, the agent would unravel the transaction and hold the same consumption. Therefore, such a change to the contract would be irrelevant to the effort choice, consumption choice or total surplus created. In intuitive terms, firms and agent are "equally" risk-averse to market risk and thus reallocating market risk between the firm and the agent would not provide more efficient insurance.

4. 3 Relative Performance

As a benchmark to the analysis, Proposition 4.2 analyzes whether relative performance can be optimal in the absence of informational frictions. A relative performance payment is defined as a wage that does not depend on the realization of α.

[16]To some extent, this feature is caused by the assumption that the number of firms is large. If the number of firms was finite, the agent could unravel some of the firm's idiosyncratic risk by trading the market portfolio; this would create a motive to avoid private hedging (however, the result would carry over provided the agent were to be constrained to trade only on other firms).

[17]This argument is, for example, present in Abowd and Kaplan (1999) as a recommendation to improve executive compensation. Note that Bebchuk and Fried do not support this explanation; their observation is that CEOs use this argument to demand higher expected wages which, as shown here, is incorrect.

Proposition 4.2. *Let $\Gamma^s = (w(.,.), a, q(.), R)$ be a first-best competitive equilibrium. Then, there exists a first-best competitive equilibrium with private hedging $\Gamma = (\tilde{w}(.,.), \eta(.), a, q(.), R)$ where $\tilde{w}(.,.)$ does not depend on α (pure relative performance).*

Proof: Set $w^{**}(\alpha, y) = \int \alpha q^*(\alpha) d\alpha a^*$ (constant). It is then easily verified that $\eta^{**}(y) = \alpha a^*$ is a solution to Q^*.□

It turns out that, in first-best, the optimal contract can be rewritten in terms of a compensation that does not depend on α but letting the agent privately hedge. To summarize, while the first-best competitive equilibrium does not exclude pay as a function of systematic risk, it suggests that filtering out systematic shocks is also possible and may even be, presumably, more transparent to shareholders.

This question is now discussed when effort is unobservable. To begin with, consider the case of constant absolute risk aversion. By Proposition (3.4), the wage in an equilibrium with no hedging can be written additively as $w(\alpha, y) = w_1(\alpha) + w_2(y)$. Then, integrating with respect to y and using the market-clearing constraint yields that: $w_1(\alpha) = \alpha a - \int f(y|a) w_2(y) dy$. The optimal contract with CARA utilities is thus the sum of a unit of the market portfolio and a relative performance payment. Suppose now that, instead of offering $w(.,.)$, the principal offers a wage $\hat{w}(\alpha, y) = w_b^*(y) + w_f^*$, composed of a relative performance payment $w_b^*(y) = w_2(y) - \int f(y|a) w_2(y) dy$ and a fixed salary $w_f^* = a \int q(\alpha) \alpha d\alpha$. Then, assume that the principal lets the agent trade on the Arrow-Debreu assets. Proposition 4.1 implies that the agent will use w_f^* to purchase one unit of the market portfolio, thus consuming exactly $w(\alpha, y)$. Collecting these observation yields the next result.

Proposition 4.3. *Suppose $u(.)$ is CARA. Suppose $\Gamma = (w(.,.), a, q(.), R)$ is a competitive equilibrium with moral hazard, then there exists a competitive equilibrium with moral hazard and private hedging $\tilde{\Gamma} = (\tilde{w}(.,.), \eta(.), a, q(.), R)$ such that $\tilde{w}(.,.)$ does not depend on α and $w(\alpha, y) = \tilde{w}(\alpha, y) + \eta(\alpha)$.*

In the case of CARA utilities, the hedging and contracting decisions can be separated. The agent can be given a contract that only depends on the systematic component and use this expected surplus to trade on capital markets. This will yield a consumption that is the same as that of the optimal contract with commitment not to trade ($w(\alpha, y)$). The result corresponds to relative performance pay in the conventional sense.

In comparison, if absolute risk-aversion is not constant, the consumption of the agent cannot be written as the sum of a function of the aggregate shock and a function of the idiosyncratic shock (see also Equation (III.61)). Therefore, there will be no hope of replicating a consumption profile $w(\alpha, y)$ with a pure relative performance payment and private hedging on capital markets (since, typically, $w(\alpha, y) \neq \tilde{w}(y) + \eta(\alpha)$). As a result, in a competitive equilibrium with private hedging, firms would always offer labor contracts such that the wage depends on systematic risk factors. This observation is summarized next.

Proposition 4.4. *Suppose that $u(.)$ has non-constant absolute risk-aversion (on any interval). Then, in any competitive equilibrium with hedging, $w(\alpha, y)$ must depend on α.*

Unlike in first-best, the separation between labor wages and hedging does not hold with non-constant absolute risk-aversion. One reason for this is that the proper amount of exposure to systematic risk should depend on the realization of the idiosyncratic shock. For example, with CRRA utility functions, an agent who is more successful than his/her peers should bear more systematic risk. The problem with pure relative performance is that the agent cannot choose an exposure to systematic risk that is conditional on the realization of the idiosyncratic shock.

To make relative performance optimal, more freedom is required on the hedging choices of the agent. For the sake of the argument, it may be possible to instead allow agents to trade on capital markets and condition the *number* of the Arrow-Debreu securities purchased on the realization of the idiosyncratic shock y_k. In this case, Proposition 4.1 shows that relative performance would then yield the same consumption (and surplus) as the optimal contract without hedging. From a more practical perspective, however, such an arrangement would require an external monitor controlling whether the agent is only trading on Arrow-Debreu assets and not unraveling the firm's idiosyncratic risk. Since the final payment from hedging would depend on y_k, such monitoring may be difficult to implement.

5 Limited Information

In the benchmark model, it is assumed that the principal can contract on α, the systematic component, and y_k, the idiosyncratic component. The model, however, does not say how firms may recover α and y_k when given only information on total output $\pi_k = \alpha y_k$. To make inference on the systematic risk possible, assume that $(\pi_k)_{k \in [0,1]}$ is observable and contractible by all firms.

At first sight, this limited informational environment seems inconsequential. Letting a denote the aggregate effort (which is exogenous from the perspective of small market participants), firm k can estimate α as $\hat{\alpha} = \int \pi_k dk/a$ and y_k as $\hat{y}_k = \pi_k/\hat{\alpha}$. As a result, a competitive equilibrium when α and y are observable remains a competitive equilibrium when all $(\pi_k)_{k \in [0,1]}$ are observable. This Section addresses two additional problems that may appear in such limited informational environments.

5. 1 Private Risk Management

The model is extended as follows. As stated earlier, the principal may now only observe $(\pi_k)_{k\in[0,1]}$ and may use the cross-section of π_k to perfectly observe α. In addition, the agent can affect the firm's exposure to systematic risk by (privately) trading on Arrow-Debreu assets. This is formalized as purchasing assets $(\omega(\alpha))_\alpha$, where $\omega(\alpha)$ represents the number of assets paying one when state is α. Net of these purchases, the output of the firm will be: $\pi = \alpha y + \omega(\alpha) - \int q(\alpha')\omega(\alpha')d\alpha'$. Here, the first term represents the "true" output of the firm, the second term represents the payoff from the portfolio and the third term represents the cost of the portfolio.

The assumption that the principal may not monitor such trades is not unrealistic. In practice, managers are in charge of the daily operations of the firm and their decisions may affect to how much systematic risk the firm is exposed. It may be difficult for investors to estimate what will be the ex-post exposure of the firm to systematic risk and how much of it is unnecessary for the operations of the firm. Risk management is modeled here in a parsimonious manner by assuming that the agent can choose *any exposure with zero value*; this is meant to separate systematic risk management from (value-increasing) effort.

The portfolio of assets $(\omega(\alpha))_\alpha$ has zero value for the principal, so that one may without loss of generality assume that the optimal contract elicits $\omega(\alpha) = 0$ for all α.[18]

Proposition 5.1. *If $u(.)$ has non-decreasing absolute risk-aversion, then a contract that can be written in the form of Equation* (III.9) *will not elicit $\omega(\alpha) = 0$ for all α. If $u(x) = \ln(x)$, on the other hand, the optimal contract of Section 3. 5 is robust to risk management.*

The problem with risk management is that the agent can mimic a high idiosyncratic performance by trading on Arrow-Debreu assets. Suppose for example that, for some state α, the agent generates a true performance y_1 but is willing to report $y_2 > y_1$. When not trading on Arrow-Debreu assets, the agent will report $\pi = \alpha y_1$. Without trading, the principal would divide π by α to correctly estimate the idiosyncratic component y_1. Suppose, however, that the agent purchases a hedge $\eta(\alpha) = \alpha(y_2 - y_1)$ (*). Doing so will generate a profit equal to αy_2 and thus the principal will (incorrectly) calculate idiosyncratic performance as $\pi/\alpha = y_2$. From the perspective of the principal, the agent will appear to have produced a high idiosyncratic performance and will be rewarded for it.[19]

[18]For any optimal contract that elicits some $\omega(\alpha)$ non-zero, one may offer another equivalent contract such that $\omega(\alpha)$ is zero.

[19]From an ex-ante perspective, the agent should also, in this example, sell certain Arrow-Debreu securities in other states of the world, so that the value of the portfolio of assets is zero.

Proposition 5.1 shows that private risk management can be costly to the principal and will distort the optimal contract. When absolute risk-aversion is non-decreasing, there is a *high* performance-pay in states with low aggregate endowment. Note also that the number of assets to be purchased required to mimic high performance is decreasing in α (by $(*)$). Therefore, in states with low endowment: (i) it is cheap to mimic high performance, (ii) it is valuable to do so because of high performance-pay. The combination of these two effects makes it desirable to trade on Arrow-Debreu assets.[20]

When absolute risk-aversion is decreasing, on the other hand, performance-pay is increasing in the aggregate state. Then, while it is cheap to mimic high performance in low states, this leads to a lower wage increase. In the case of a logarithmic utility function, both sides of this trade-off coincide and the risk management problem does not constrain the optimal contract. As noted in Section 3. 5, the aggregate shock simply scales contract payments and thus performance-pay is also scaled by α. This matches the number of additional assets to be purchased (which is also scaled by α).

5. 2 Joint Deviations

Even in the absence of risk management, not being able to observe α or y separately may be problematic. Recall that the firm infers the true α using a conjecture about the effort supplied by all agents in the economy ($\hat{\alpha} = \int \pi_k dk/a$). While one agent only will not affect $\int \pi_k dk$, a joint deviation by all agents may affect $\hat{\alpha}$ and lead to incorrect estimates $\hat{\alpha} \neq \alpha$.

This problem is formalized next. Suppose that the firms incorrectly believe that all agents choose a, when agents deviate to $a' \neq a$. Then, the principal will make two errors. One, when estimating α, the principal will incorrectly estimate $\hat{\alpha} = \int \pi_k dk/a = \alpha a'/a$. That is, when all agents deviate to supply less effort, the principal will understate the aggregate shock. Two, when estimating the true y_k, the principal will incorrectly use $\hat{\alpha}$ and thus will pay according to $\hat{y}_k = \pi_k/\hat{\alpha} = y_k a/a'$. That is, the principal will overstate the idiosyncratic component. Thus, jointly decreasing effort generates the following two effects: (i) a decrease in systematic pay, (ii) an increase in idiosyncratic pay.

Proposition 5.2. *If at the optimal contract:*

$$\int\int h(\alpha)f(y|a)\frac{-u'(w(\alpha,y))}{u''(w(\alpha,y))}(H(y)G'(\alpha) - H'(y)G(\alpha))dyd\alpha \neq 0 \qquad \text{(III.22)}$$

Then, there exists a profitable joint deviation.

[20] If the production technology is modified to $\alpha + y$ (instead of αy), then (i) will no longer apply and thus this conclusion will be true for any utility function but CARA. This can be proved with a straightforward modification of the proof of Proposition 5.1.

Proposition 5.2 formalizes the trade-off faced by the agents. When jointly decreasing effort, agents misrepresent the systematic as well as the idiosyncratic components of pay. Such actions are desirable whenever there is too much difference between the pay-for-performance coefficient (w_y) and the pay for systematic risk coefficient (w_α). In particular, in a broader context, Equation (III.22) suggests that pay as a function of systematic risk may play an important role in deterring joint deviations. On the other hand, settings with high-powered incentives (w_y large) may be easily prone to joint deviations. Note also that in the limiting case such that the cost of effort becomes small, $H'(y)$ will also become small and thus Equation (III.22) will be necessarily violated.

To implement the equilibrium favorable to the firms, note first that a joint deviation can always be prevented provided there is a signal observable by the firms that fully reveals that a joint deviation has taken place (in this case, the firms could offer a contract such that the wage is very low if the Arrow-Debreu prices are not equal to their no-deviation value). In this economy, a useful signal may be whether financial markets would clear at the (public) Arrow-Debreu prices corresponding to no joint deviation.[21]

The methodology will be to allow the agents to privately hedge (as in the previous Section) to endogenously tie a deviation to the state prices. Recall from Proposition 4.1 that hedging should be irrelevant to the principal provided all agents choose the recommended effort. Thus, if all agents choose the recommended effort and are given the optimal contract without hedging, they should not affect the state prices. Suppose next that the agents all deviate to $a' \neq a$. Since an effect of state prices would imply that the joint deviations is detected, it is sufficient to check that asset markets would not clear at the original state prices conditional on the joint deviation, i.e. there exists states such that the agents will demand some Arrow-Debreu assets. Stating first the hedging decision of the agents:

$$(\tilde{Q}) \qquad \max_{\tilde{\eta}(.),A} \int\int h(\alpha)u(w(\hat{\alpha},\hat{y})+\tilde{\eta}(\alpha)-A)f(y|a')dyd\alpha$$

$$\text{s.t.} \qquad \int q(\hat{\alpha})\eta(\alpha)d\alpha = A$$

Proposition 5.3. *The solution of $(\tilde{Q})$ must involve $\eta(\alpha)$ non-zero for a set of values of α with non-zero measure.*

As shown in Proposition 5.3, the agents *would* trade after a joint deviation and thus their actions would, in equilibrium, cause effects on the Arrow-Debreu prices. This result suggests that contracting on the Arrow-Debreu prices (e.g., punishing the agents when these prices are different from what they

[21] The idea is kept informal to avoid new definitions and additional notations that do not seem necessary for the intuition and results. Further, stating the extended contract formally would be fairly straightforward here.

should be) should be sufficient to deter any joint deviation.[22]

6 Constrained Contracts

6. 1 Linear Contracts

To better understand what additional constraints are imposed by linear contracts, it is now helpful to analyze the problem when firms are constrained to use linear schemes. Formally, a linear contract is defined as a payment $w(\alpha, y) = \theta_1 \alpha y + \theta_0$.[23] In a contract of this form, the agent holds a fraction θ_1 of the firm and receives a salary θ_0. Using the market clearing condition implies that for all α,

$$\alpha a = \theta_1 \alpha a + \theta_0 \tag{III.23}$$

Lemma 6.1. *If firms are constrained to linear contracts, agents fully own their firm ($\theta_1 = 1, \theta_0 = 0$).*

In contrast to partial equilibrium analysis, the performance-pay coefficient does not depend on the utility function or the production technology.[24] To make this example simpler, assume that there are now only two possible effort choices, $\underline{a}$ and $\overline{a}$. This simplification will be useful to derive a closed-form expression for aggregate risk premia. Writing the standard incentive-compatibility condition yields:

$$\int\int h(\alpha) f(y|\overline{a}) u(\alpha y) dy d\alpha - \psi(\overline{a}) \geq \int\int h(\alpha) f(y|\underline{a}) u(\alpha y) dy d\alpha \tag{III.24}$$

Because Equation (III.24) is written only in terms of exogenous parameters, a competitive equilibrium with effort may fail to exist. To proceed further, assume that this condition is strictly satisfied. Note that, as a result, a competitive equilibrium must prescribe that the incentive-compatibility is not binding.[25]

$$(P') \qquad \min_{w(.,.)} \int\int q(\alpha) \alpha d\alpha \theta_1 + \theta_0$$

[22]There are well-known techniques using voluntary reports by agents to implement a particular equilibrium (see Bassetto and Phelan (2007) for an example in which a social planner attempts to infer the aggregate endowment in the economy); thus, being able solve this problem using a more elaborate mechanism should not come as a surprise. Yet, general mechanisms have often been criticized on the grounds that they do not seem to look like actual contracts. To avoid this criticism, the approach used here will be to use only observable variables but no voluntary reports.

[23]This functional form is more restrictive than other functional forms used in the literature (e.g., $\theta_\alpha \alpha + \theta_y y + \theta_0$). The same conclusions would hold for more general functional forms but the example is analytically more straightforward in the simpler case considered here.

[24]One can for example compare this result to the LEN model used in previous papers (see Murphy (1999) for an example), in which the performance coefficient depends on the volatility of the idiosyncratic shock and the cost of effort.

[25]Unlike with fully-optimal contract, this may now be possible given the constrained contractual form.

s.t.

$$\int\int h(\alpha)f(y|\overline{a})u(\theta_1\alpha y+\theta_0)dyd\alpha-\psi(\overline{a}) \quad = R \quad (\lambda) \tag{III.25}$$

Differentiating with respect to θ_0 and setting $(\theta_0, \theta_1) = (0, 1)$,

$$\lambda = \frac{1}{\int\int h(\alpha)f(y|\overline{a})u'(\alpha y)dyd\alpha} \tag{III.26}$$

Reinjecting this expression in the first-order with respect to θ_1 yields:

$$\int q(\alpha)\alpha d\alpha = \frac{\int\int h(\alpha)\alpha y f(y|\overline{a})u'(\alpha y)dyd\alpha}{\int\int h(\alpha)f(y|\overline{a})u'(\alpha y)dyd\alpha} \tag{III.27}$$

In an economy constrained to linear contracts, then, a competitive equilibrium can prescribe risk premia that are different from those implied by a representative agent and thus there will be distortions to state prices. The distortion will exist as long as:

$$\frac{\int \alpha h(\alpha)u'(\alpha\overline{a})d\alpha}{\int h(\alpha)u'(\alpha\overline{a})d\alpha} \neq \frac{\int\int h(\alpha)\alpha y f(y|\overline{a})u'(\alpha y)dyd\alpha}{\int\int h(\alpha)f(y|\overline{a})u'(\alpha y)dyd\alpha} \tag{III.28}$$

6. 2 Limited Liability

The presence of external constraints on contract payments can disrupt the agent's hedging decision and violate Proposition 4.1. An example of such a situation can be derived in the (reasonable) situation such that agents have limited liability and cannot be paid less than $\underline{w} < 0$. To make the problem interesting, assume as well that there exists a subset of states of the world such that the limited liability binds.

While the analysis of the contract remains essentially unchanged, the presence of a limited liability may affect incentives to trade. To see this, denote $\tilde{y}(\alpha)$ as the performance threshold such that the limited liability no longer binds. It is given by:

$$u'(\underline{w}) = G(\alpha)H(\tilde{y}(\alpha)) \tag{III.29}$$

Since the previous comparative statics remain valid, the agent is now paid $w(\alpha, y) = \underline{w}$ for $y \leq \tilde{y}(\alpha)$ and $w(\alpha, y) > \underline{w}$ else. If Proposition 4.1 is still to hold, it must then hold for all α, α' that (first-order condition for (Q)):

$$\frac{u'(\underline{w})\int_{-\infty}^{\tilde{y}(\alpha)} f(y|a)dy/G(\alpha) + \int_{\tilde{y}(\alpha)}^{+\infty} H(y)dy}{u'(\underline{w})\int_{-\infty}^{\tilde{y}(\alpha')} f(y|a)dy/G(\alpha') + \int_{\tilde{y}(\alpha')}^{+\infty} H(y)dy} = 1 \qquad \text{(III.30)}$$

This implies that the term in the numerator is constant in α. But differentiating with respect to α yields that $u'(\underline{w})\int_{-\infty}^{\tilde{y}(\alpha)} f(y|a)dy(1/G(\alpha))' > 0$. As a result, with a binding limited liability constraint, the agent would trade if given the opportunity to do so, in contrast with previous results. More precisely, the agent would sell short states of the world with low α and use the proceeds to buy states with high α.[26] This is because the limited liability binds more when α is low which translates in a distortion in the hedging decision as well.

In the presence of a binding limited liability, then, the firm will always gain from controlling the hedging decision. Intuitively, the optimal contract will induce excess hedging in low-endowment states in order to loosen the limited liability constraint.

Risk-Neutrality

The assumption of risk-aversion is important for the presence of risk premia in the economy, but some elements of the approach carry over to the risk-neutral case. Assume now that $u(x) = x$. To obtain explicit solutions to the problem, the problem is drastically simplified to only two outputs $\underline{y}$ and $\overline{y}$ and two possible efforts $\underline{a}$ and $\overline{a}$. Conditional on effort a, the probability of $\overline{y}$ is denoted $p(a)$, with $p(\overline{a}) > p(\underline{a})$. Denote $\overline{\alpha} = \int h(\alpha)\alpha d\alpha$. The cost of effort is denoted $B = \psi(\overline{a}) < \overline{\alpha}(\overline{a} - \underline{a})$. The solution to this model is presented next.

Proposition 6.1. *Denote $\Delta = \underline{w} + \frac{p(\overline{a})B}{p(\overline{a}) - p(\underline{a})}$. There is a competitive equilibrium given as follows:*

(i) If $\Delta \leq \overline{\alpha a}$, (a) $h(\alpha) = q(\alpha)$ for all α, (b) $R = \overline{\alpha a} - B$, (c) $w(\alpha, \underline{y}) = \underline{w}$ (resp. $w(\alpha, \overline{y}) = \underline{w} + \frac{\alpha\overline{a} - \underline{w}}{p(\overline{a})}$).

(ii) If $\Delta > \overline{\alpha a}$, (a) $h(\alpha) = q(\alpha)$ for all α, (b) $R = \underline{\alpha}\overline{a}$, (c) $w(\alpha, \underline{y}) = \alpha\underline{a}$.

Further, for a given Δ, all competitive equilibria prescribe the same effort.

Corollary 6.1. *If $\Delta \leq \overline{\alpha a}$, there is a competitive equilibrium with $R < \alpha\overline{\alpha a} - B$ if and only if $R \in [\overline{\alpha}\underline{a}, \alpha\overline{\alpha}]$ and $\Delta = \overline{\alpha a}$.*

[26] It is open question as to what the optimal contract would look like with limited liability and private hedging; this Section only checks that Proposition 4.1 is invalidated.

Corollary 6.2. *In a competitive equilibrium with hedging, offering a relative performance payment $w(\alpha, \underline{y}) = \underline{w}$ and $w(\alpha, \overline{y}) = (\overline{\alpha}a - (1 - p(a))\underline{w})/p(a)$ (where a is the elicited effort in (i) or (ii)) achieves the same level of welfare as the contract of Proposition 6.1.*

The first part of Proposition 6.1 is similar to Tirole (2006). The term $\frac{p(\overline{a})B}{p(\overline{a})-p(\underline{a})}$ corresponds to the minimum pledgeable wealth of the agent, required to operate the firm with high effort. When the wealth of the agent lies below this bound, the firm is operated with low effort.

The general equilibrium approach shows that the payments given to the agent depends on the aggregate endowment. The contract that maximizes incentives (by inducing $\underline{w}$ conditional on $\underline{y}$) prescribes a pay-for-performance coefficient that is increasing in α.[27] This is because a high aggregate endowment includes also more resources to reward the agent.

7 Extensions

7. 1 General Production Technology and Heterogeneity

The model is extended to incorporate: non-identical agents, a general production technologies in α and y. Assume now that agent $k \in [0, 1]$ can produce output with a technology $\phi_k(\alpha, y_k)$ (instead of αy_k), increasing in the systematic shock α and the idiosyncratic shock y_k. The distribution of y_k is denoted $f^k(y|a_k)$ and may depend on k as well as effort. Each agent has a utility $u_k(c_k) - \psi_k(a_k)$, such that risk-aversion and cost of effort may depend on k as well.[28] It is now convenient to formulate the problem of agent k seeking an employment contract from firm k.

$$(P_k) \quad \max_{w(.,.),a} \quad \int\int h(\alpha) f^k(y|a) u_k(w(\alpha, y)) d\alpha dy - \psi_k(a)$$

s.t.

$$\int\int q(\alpha) f^k(y|a)(\phi_k(\alpha, y) - w(\alpha, y)) dy d\alpha \geq R^f \qquad (\lambda) \qquad \text{(III.31)}$$

$$\int\int h(\alpha) f^k_a(y|a) u_k(w(\alpha, y)) dy d\alpha = \psi'_k(a) \qquad (\mu) \qquad \text{(III.32)}$$

Equation (III.31) is the participation of the firm, where R^f denotes the expected profit expected by a firm on an employment contract. Equation (III.32) is the agent's incentive-compatibility.

[27] This contract is of course not unique when $\Delta < \overline{\alpha a}$; but it is the unique optimal contract that will be robust to all possible values of Δ compatible with high effort.

[28] The results are unchanged if firms are non-identical and owned by agents in the economy. With non-identical agents and firms, there may be employer-employee matching and bargaining, which would require a full-fledged model of job matching.

Definition 7.1. *A competitive equilibrium is* $\Gamma = ((a_k, w_k(.,.))_{k\in[0,1]}, q(.), R^f)$ *such that:*

1. *For all* k, $(a_k, w_k(.,.))$ *maximizes* (P_k).

2. *For all* α, $\int \phi_k(\alpha, y) f^k(y|a_k) dy d\updownarrow(k) = \int w_k(\alpha, y) f^k(y|a_k) dy d\updownarrow(k)$.

Lemma 7.1 verifies that firms make zero profit.

Lemma 7.1. *In a competitive equilibrium,* $R^f = 0$.

The first-order optimality condition for (P_k) is:

$$u_k'(w(\alpha, y)) \underbrace{\{1 + \mu_k \frac{f_a^k(y|a)}{f^k(y|a)}\}}_{G_k(y)} = \underbrace{\lambda_k \frac{q(\alpha)}{h(\alpha)}}_{T_k(\alpha)} \tag{III.33}$$

Lemma 7.2. *In a competitive equilibrium,* λ_k *and* μ_k *are strictly positive.*

Lemma 7.3. $G_k(y)$ *is is strictly positive and strictly increasing in* y *and* $T_k(\alpha)$ *is strictly decreasing in* α.

Proposition 7.1. *In a competitive equilibrium,* $w_k(\alpha, y)$ *is increasing in* α *and* y. *Further, if* $u_k(.)$ *has decreasing absolute risk-aversion,* w_y *is increasing in* α.

Proposition 7.1 confirms the intuition of Proposition 3.4 in the more general case of a general production technology and, possibly, heterogeneity across agents. In the DARA class of utility functions, the consumption cannot be additively separated in α and y and thus a relative performance payment cannot replicate $w_{k,\alpha,y}$ even if the agent can hedge the market portfolio.

7. 2 Multiple Periods and Serial Correlation

The model is extended first to two periods, in order to illustrate both qualitatively and quantitatively *changes* to the compensation in response to shocks to the business cycle. Assume that production takes place over two periods $t = 1, 2$. Agents have a utility function $U(c_1, c_2) = u(c_1) + \beta u(c_2)$, where c_1 represents period 1 consumption and c_2 represents period 2 consumption. Assume in addition that, in period 1, output is deterministic and equal to $a_0 > 0$. In period 2, output is equal to αy, where α is drawn from a distribution $h(.)$ and y is drawn from a distribution $f(.|a)$. Both distributions satisfy the previous restrictions.

It is now convenient to denote $q(\alpha)$ the price of an asset paying one unit of consumption in period 2, in units of consumption of period 1. Normalizing the price of a unit of consumption in period 1 to one, the firm's contracting problem can be written as follows:

$$(P) \qquad \max_{w_1, w_2(.,.), a} \int\int (\alpha y - w_2(\alpha, y)) f(y|a) q(\alpha) dy d\alpha - w_1$$

s.t.

$$u(w_1) + \beta \int\int h(\alpha) u(w_2(\alpha, y)) f(y|a) dy d\alpha - \psi(a) \geq R \tag{III.34}$$

$$\beta \int\int h(\alpha) f_a(y|a) u(w_2(\alpha, y)) dy d\alpha = \psi'(a) \tag{III.35}$$

In this problem, the firm chooses the wage in period 1, w_1, and the wage in period 2, $w_2(\alpha, y)$. As before, markets must clear. This leads to the following definition of a competitive equilibrium.

Definition 7.2. *A (symmetric) competitive equilibrium is* $\Gamma = (a, w_1, w_2(., .), q(.), R)$ *such that:*

1. $(a, w_1, w_2(., .))$ *maximizes* (P) *for all* k.

2. *For all* α, $\alpha a = \int f(y|a) w(\alpha, y) dy$ *and* $w_1 = a_0$.

To illustrate the analysis, I focus next on the case with $u(c) = \ln(c)$, $f(y|a)$ a Normally distributed with variance one and mean a truncated at $y = 0$. Since the focus of this Section is not the analysis of the optimal effort, assume here that $a = 1$ is the optimal level of effort. In this case, the likelihood ratio $f_a(y|a)/f(y|a)$ at $a = 1$ is equal to $y - \sqrt{2/\pi}$. After substituting $w_1 = a_0$, the problem of the firm is similar to the problem studied in Section 3. 5, and thus $w_2(., .)$ verifies Equation (III.15) and can be written as $w(\alpha, y) = \alpha a(1 + A(y - \sqrt{2/\pi}))$, where $A > 0$ is a constant equal to the ratio of the Lagrange multipliers μ/λ. The change in wages can then be written as follows: $w_2(\alpha, y) - w_1 = (\alpha a - a_0) + \alpha a A(y - \sqrt{2/\pi})$. The second term has mean zero under the distribution of y. The first term corresponds to the response of the wage to the aggregate shock αa. It simply says that the wage changes on average by the change in the total endowment. This is consistent with the evidence reported in Himmelberg and Hubbard (2000) and Bertrand and Mullainathan (2001), who document that changes in the wage seem to be proportional to changes in the market. In Figure III.1, the wage change $w_2 - w_1$ is plotted as a function of α and y, and clearly exhibits a greater wage for a greater aggregate shock in period 2.

Since this example is developed only in a specified two-period problem, the basic model is now extended to the general dynamic moral hazard and, possibly, serial correlation across periods. Time

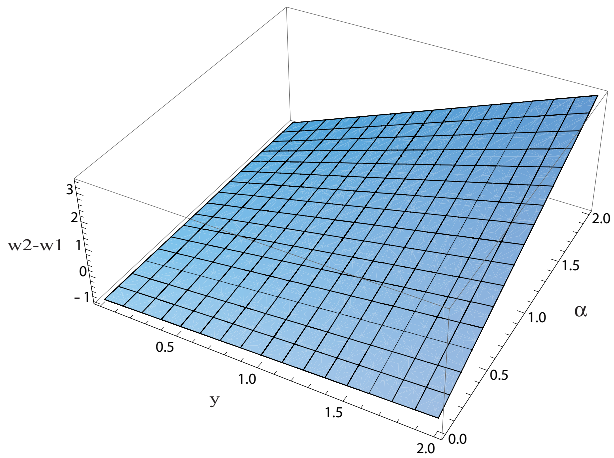


Figure III.1. Δw as a function of y and α

is discrete with $t = 0, \dots, +\infty$. At each period, firm k produces $\alpha^t y_k^t$ where α^t is current productivity shock and y_k^t is current output. To model serial correlation, let $\mathbf{y_k^t} = (\mathbf{y_k^0}, \dots, \mathbf{y_k^{t-1}})$ and $\alpha^{\mathbf{t}} = (\alpha^{\mathbf{0}}, \dots, \alpha^{\mathbf{t-1}})$ denote previous histories. The distribution of y_k^t is denoted $f(.|\alpha^{\mathbf{t}}, \mathbf{y^t}, \mathbf{a_{kt}})$ where a_{kt} is period t effort for firm k. The distribution of α^t is denoted $h(.|\alpha^{\mathbf{t}})$, and may possibly depend on all past realizations. A wage at date t is denoted $w(\mathbf{y^t}, \alpha^{\mathbf{t}}, \alpha, \mathbf{y})$, and may depend on the firm's past history $\mathbf{y^t}$, past aggregate shocks $\alpha^{\mathbf{t}}$, the current aggregate shock α and the current firm shock y. The price of receiving one unit of consumption at date t conditional on $(\alpha^{\mathbf{t}}, \alpha)$ is denoted $q(\alpha^{\mathbf{t}}, \alpha)$. In short, these functions are denoted $w_{kt}(\alpha, y)$, $q_t(\alpha)$, $f^{kt}(y|a)$ and $h_t(\alpha)$. The dependence on firm k is omitted when there is no ambiguity.[29] Finally, let $\updownarrow_\alpha^t$ and $\updownarrow_y^t$ denote the probability measure associated to outcomes $\alpha^{\mathbf{t}}$ and $\mathbf{y^t}$, respectively.

The firm and the agent sign a contract with commitment at date 0, which may dynamically insure the agent against systematic and idiosyncratic shocks.

$$\max_{w_t, a_t} \int\int \sum_{t=0}^{\infty} q_t(\alpha)(\alpha y - w_t(\alpha, y)) f^t(y|a_t) dy d\alpha d\updownarrow_\alpha^t d\updownarrow_y^t$$

$$\text{s.t.} \quad \forall \alpha^{\mathbf{t}}, \mathbf{y^t}$$

$$\int\int \sum_{t=0}^{\infty} h_t(\alpha) u(w_t(\alpha, y)) f^t(y|a_t) dy d\alpha d\updownarrow_\alpha^t d\updownarrow_y^t - \psi(a_t) \geq R \qquad \text{(III.36)}$$

(III.37)

$$\int\int \sum_{t=0}^{\infty} h_t(\alpha) u(w_t(\alpha, y)) f_a^t(y|a_t) dy d\alpha d\updownarrow_\alpha^t d\updownarrow_y^t - \psi'(a_t) = 0 \qquad \text{(III.38)}$$

[29]The previous assumptions are maintained. Note that, as a result, $f_t(y|a)$ must have mean a which is with loss of generality. However, this restriction is mostly for convenience and can be easily relaxed.

In the above program, both the wage w_t and the elicited effort can depend on $\alpha^{\mathbf{t}}$ and $\mathbf{y^t}$. The definition of a competitive equilibrium is similar to the definitions given before and not repeated here; it requires contracts to solve the above program and market clearing for any history and any date. The first-order optimality condition for this dynamic contract design is given next.

$$\frac{q_t(\alpha)}{\beta^t h_{kt}(\alpha)} = u'(w_{kt}(\alpha, y))(\lambda_{kt} + \mu_{kt}\frac{f_a^{kt}(y|a_t)}{f^{kt}(y|a_t)}) \tag{III.39}$$

This optimality condition has the same form as Equation (III.9), except that the dependence on past history is explicitly modeled. It can then be easily verified with the same steps as in Proposition 7.1 that the wage will be increasing will be increasing in both systematic and idiosyncratic risk. In addition, with DARA utility functions, the pay-for-performance coefficient will be greater when α is high.

8 Appendix

9 Appendix A: Centralized Problem

To further compare the approach to the existing literature on mechanism design, the optimal incentive scheme is now derived given a benevolent central planner. Since each agent is symmetric, one may focus on the optimal incentive scheme for a single agent. The solution to the planned problem is given as follows:

$$(P^*) \quad \max_{w(.,.),a} \quad \int\int h(\alpha)f(y|a)u(w(\alpha, y))dyd\alpha$$

s.t.

$$\psi'(a) = \int\int h(\alpha)f_a(y|a)u(w(\alpha, y))dyd\alpha \tag{III.40}$$

$$\alpha a = \int w(\alpha, y)f(y|a)dy \tag{III.41}$$

The main difference between the decentralized problem and the planned problem is that the market-clearing constraint is imposed in (P^*) as a budget-balancing constraint. Note that, as a result, any competitive equilibrium is feasible in (P^*). Conversely, a competitive equilibrium is said to be constrained-efficient if it prescribes a consumption profile $w(.,.)$ solution to (P^*). To be able to decentralize the solution to (P^*), the following stronger restriction on the problem is made: the first-order conditions in problem (P_k) are sufficient for an optimum for any normalized state prices.

Proposition A. 2. *If (P^*) has a solution, then there exists a constrained-efficient competitive equilibrium.*

Proof: Let $\nu(\alpha)$ denote the multipliers associated to Equation (III.4) and κ the multiplier associated to the incentive-compatibility. The first-order condition with respect to $w(.,.)$ for this problem is:

$$u'(w^*(\alpha,y)) = \frac{\nu(\alpha)}{h(\alpha)} \frac{1}{1+\kappa\frac{f_a(y|a^*)}{f(y|a^*)}} \tag{III.42}$$

Now consider the decentralized problem and set Γ as follows: $q(\alpha) \equiv \nu(\alpha)/\int \nu(\alpha)d\alpha$, $a \equiv a^*$, $R \equiv \int\int h(\alpha)f(y|a^*)u(w^*(\alpha,y))dyd\alpha$. One needs next to show that a^* and w^* are optima in (P_k). To show this, one may rearrange Equation (III.42) as follows:

$$u'(w^*(\alpha,y)) = \frac{q(\alpha)}{h(\alpha)} \frac{1}{1/\int \nu(\alpha)d\alpha + \kappa/\int \nu(\alpha)d\alpha\frac{f_a(y|a^*)}{f(y|a^*)}} \tag{III.43}$$

Choosing the multipliers of the decentralized problem as $\lambda \equiv 1/\int \nu(\alpha)d\alpha$ and $\mu \equiv \kappa/\int \nu(\alpha)d\alpha$, $w^*(.,.)$ is an optimum. Next, the first-order condition with respect to effort in the centralized problem is written as follows:

$$\int\int h(\alpha)f_{aa}(y|a^*)u(w^*(\alpha,y))d\alpha dy - \psi''(a^*) = \{\int \nu(\alpha)\int w^*(\alpha,y)f_a(y|a^*)dyd\alpha - \int \nu(\alpha)\alpha d\alpha\}/\kappa \tag{III.44}$$

Rewriting $\kappa = \mu\int\nu(\alpha)d\alpha$ and $\nu(\alpha) = q(\alpha)\int\nu(\alpha)d\alpha$ and simplifying yields:

$$\mu\int h(\alpha)f_{aa}(y|a^*)u(w^*(\alpha,y))dyd\alpha - \psi''(a^*) = \int q(\alpha)\alpha d\alpha - \int\int \alpha w^*(\alpha,y)f_a(y|a^*)q(\alpha)dyd\alpha \tag{III.45}$$

This is the first-order condition with respect to a in the decentralized problem.□

Proposition A2 establishes that under certain conditions, one may find a constrained efficient competitive equilibrium. In such cases, firms may sustain the same allocation as the planner even though they do not take the budget-balance constraint into account in their optimization. Note also that the statement does not exclude other competitive equilibria with inefficient effort.

10 Appendix B: Omitted Proofs:

First-Order Approach: To guarantee that the first-order approach is valid, it must be established that, when using the contract given by Equation (III.9), there exists a unique solution for a in the problem of the agent.

Writing the expected utility of the agent $U(a)$,

$$\begin{aligned} U(a) &= \int h(\alpha) \int u(w(\alpha, y)) f(y|a) dy d\alpha - \psi(a) \\ &= \int h(\alpha)([u(w(\alpha, y))F(y|a)]_{\underline{y}}^{\overline{y}} - \int u'(w(\alpha, y))F(y|a) w_y(\alpha, y) dy) d\alpha - \psi(a) \end{aligned} \quad \text{(III.46)}$$

By Proposition 3.3, $w_y(\alpha, y) > 0$. Then, it is readily verified that $U(a)$ is strictly concave whenever $F_{aa}(y|a) > 0$. This is the convexity of the distribution function condition (CDFC) given in Rogerson (1985).□

Proof of Proposition 3.1: Note first that constraint (III.1) is binding and denote λ^* its associated Lagrange multiplier. The first-order condition in $w(\alpha, y)$ can be written:

$$\lambda^* h(\alpha) u'(w(\alpha, y)) = q(\alpha) \quad \text{(III.47)}$$

Thus, w depends only on α.

The first-order condition in a yields that:

$$\int q(\alpha) f_a(y|a)(\alpha y - w(y, \alpha)) dy d\alpha - \lambda^* \psi'(a) + \lambda \int \int h(\alpha) u(w(\alpha, y)) f_a(y|a) dy d\alpha \quad \text{(III.48)}$$

Because the w does not depend on y, the above expression can be simplified as follows:

$$\int \int q(\alpha) \alpha d\alpha = \lambda^* \psi'(a) \quad \text{(III.49)}$$

Together, Equations (III.47), (III.49) and (III.1) characterize a solution to the contracting problem. Note first that $\lambda^* > 0$ or else all Arrow-Debreu prices would be equal to zero. It follows that the solution to Equation (III.49) is unique.

Consider next the restrictions imposed by the market clearing condition. From Equation (III.4),

$w(\alpha, y) = \alpha a$ for all α, y. This wage must also be a solution to Equation (III.47), i.e.:

$$\lambda^* h(\alpha) u'(\alpha a) = q(\alpha) \tag{III.50}$$

Integrating both sides of this Equation with respect to α yields:

$$\lambda^* = \frac{1}{\int h(\alpha) u'(\alpha a) d\alpha} \tag{III.51}$$

Plugging this multiplier into Equation (III.50),

$$q(\alpha) = \frac{h(\alpha) u'(\alpha a)}{\int h(x) u'(xa) dx} \tag{III.52}$$

Rewriting Equation (III.49) with the state prices:

$$\int h(\alpha) u'(\alpha a) \alpha dy d\alpha = \psi'(a) \tag{III.53}$$

To prove efficiency, consider the following planning problem.

$$(P^{**}) \qquad \max_{w(.,.),a} \int \int h(\alpha) f(y|a) u(\phi(\alpha, y)) dy d\alpha - \psi(a)$$

The solution to this planning problem corresponds to the choice of effort given by Equation (III.53).□

Proof of Proposition 3.2: Suppose $w(\alpha, y)$ does not depend on y. Then, by incentive-compatibility, all firms choose to elicit $a = \underline{a}$ and therefore, in the competitive equilibrium, pay is equal to $w(\alpha, y) = \alpha \underline{a}$. For $v \geq 0$, consider next the following alternative wage schedule:

$$u(\hat{w}(\alpha, y)) = \frac{\psi'(v)}{\int \alpha h(\alpha) d\alpha} (\alpha y - \alpha v) + \alpha \underline{a} + \psi(v) \tag{III.54}$$

Writing the incentive-compatibility condition for this wage:

$$\begin{aligned} \psi'(a) &= \int \int h(\alpha) f_a(y|a) \left(\frac{\psi'(v)}{\int \alpha' h(\alpha') d\alpha'} (\alpha y - \alpha v) + \alpha \underline{a} + \psi(v) \right) dy d\alpha && \text{(III.55)} \\ &= \psi'(v) && \text{(III.56)} \end{aligned}$$

So that this wage induces an effort equal to v. Next, note that:

$$\int\int h(\alpha)u(\hat{w}(\alpha,y))f(y|a)dyd\alpha - \psi(a) = \int \alpha\underline{a}h(\alpha)d\alpha = R \tag{III.57}$$

So that conditional on choosing effort v, the agent achieves the reserve R. It follows from Equations (III.55) and (III.57) that the new contract verifies the incentive-compatibility condition and the participation constraint.

It must now be established that this new contract makes the principal better-off. To simplify notations, let $T[y] \equiv u^{-1}(\alpha\underline{a} + \psi(v) + \frac{\psi'(v)}{\int \alpha h(\alpha)d\alpha}(\alpha y - \alpha v))$. The firm optimizes its profit $M(v)$ over $v \geq 0$:

$$M(v) = v\int \alpha q(\alpha)d\alpha - \int\int q(\alpha)f(y|v)T[y]dyd\alpha$$

The first-order condition in v yields:

$$M'(v) = \int q(\alpha)\alpha d\alpha - \int\int q(\alpha)f_v(y|v)T[y]dyd\alpha - \int\int q(\alpha)f(y|v)\frac{\frac{\psi''(v)(\alpha y - \alpha v) - \alpha\psi'(v)}{\int \alpha h(\alpha)d\alpha} + \psi'(v)}{u'(T[y])}$$

Evaluating this expression at $v = 0$, $M'(0) = \int q(\alpha)\alpha d\alpha > 0$.□

Proof of Proposition 3.3: Differentiating the market clearing condition with respect to α: $w_\alpha(\alpha,y)f_1(y)dy = A$. Thus, there exists values of y, say $\hat{y}$, such that $w_\alpha(\alpha,\hat{y}) > 0$. Differentiating both sides of Equation (III.9) with respect to α:

$$w_\alpha(\alpha,y) = H(y)\frac{G'(\alpha)}{u''(w(\alpha,y))} \tag{III.58}$$

Therefore, for all y, $Sign(w_\alpha(\alpha,y)) = Sign(-G'(\alpha))$. This is true in particular at $y = \hat{y}$ and thus $q(\alpha)/h(\alpha)$ is strictly decreasing.□

Proof of Proposition 3.4: Differentiating Equation (III.9),

$$G(\alpha)H'(y) = u''(w(\alpha,y))w_y(\alpha,y) \tag{III.59}$$

$$G'(\alpha)H'(y) = u'''(w(\alpha,y))w_\alpha(\alpha,y)w_y(\alpha,y) + u''(w(\alpha,y))w_{\alpha,y}(\alpha,y) \tag{III.60}$$

Using Equations (III.58) and (III.59) to substitute w_y and w_α in Equation (III.60) yields that:

$$w_{\alpha,y}(\alpha,y) = \frac{G'(\alpha)H'(y)}{-u''(w(\alpha,y))}\left(\frac{u'(w(\alpha,y))u'''(w(\alpha,y))}{(u''(w(\alpha,y)))^2} - 1\right) \quad \text{(III.61)}$$

Note that the term on the right-hand side is the derivative of the inverse of the Arrow-Pratt absolute risk-aversion.□

Proof of Proposition 3.5: Let $u(x) = \frac{\gamma}{1-\gamma}(b_1 + b_2 x)^{1-\gamma}$. By Equation (III.9),

$$b_2\gamma(b_1 + b_2 w(\alpha,y))^{-\gamma} = \frac{q(\alpha)}{h(\alpha)}\frac{1}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}} \quad \text{(III.62)}$$

Therefore:

$$(b_2\gamma)^{-1/\gamma}(b_1 + b_2 w(\alpha,y)) = \left(\frac{q(\alpha)}{h(\alpha)}\right)^{-1/\gamma}\left(\frac{1}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}\right)^{-1/\gamma} \quad \text{(III.63)}$$

Pre-multiplying both sides of this Equation by $f(y|a)$, integrating with respect to y and using the market clearing conditions yields that:

$$(b_2\gamma)^{-1/\gamma}(b_1 + b_2\alpha a) = \left(\frac{q(\alpha)}{h(\alpha)}\right)^{-1/\gamma}\underbrace{\int f(y|a)\left(\frac{1}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}\right)^{-1/\gamma}dy}_{K} \quad \text{(III.64)}$$

Solving for q,

$$q(\alpha) = h(\alpha)\left(\frac{(b_1 + b_2\alpha a)(b_2\gamma)^{-1/\gamma}}{K}\right)^{-\gamma} \quad \text{(III.65)}$$

Integrating both sides with respect to α and solving for K,

$$K = \left(\int h(\alpha)(b_1 + b_2\alpha a)^{-\gamma}(b_2\gamma)d\alpha\right)^{-1/\gamma} \quad \text{(III.66)}$$

Reinjecting K in Equation (III.65), implies Equation (III.7) in the case of a HARA utility function.□

Proof of Corollary 3.2: In the HARA class,

$$q'(\alpha) = \frac{h'(\alpha)u'(\alpha a) + a h(\alpha)u''(\alpha a)}{\int h(x)u'(xa)dx} \quad \text{(III.67)}$$

Then:

$$\begin{aligned} G'(\alpha) &= \frac{q'(\alpha)h(\alpha) - h'(\alpha)q(\alpha)}{h(\alpha)^2} \\ &= \frac{h(\alpha)h'(\alpha)u'(\alpha a) + ah(\alpha)^2u''(\alpha a) - h'(\alpha)h(\alpha)u'(\alpha a)}{h(\alpha^2)\int h(x)u'(xa)dx} \\ &= \frac{au''(\alpha a)}{\int h(x)u'(xa)dx} \end{aligned}$$

Reinjecting $G'(\alpha)$ in Equation (III.58),

$$w_\alpha(\alpha, y) = KH(y)\frac{u''(\alpha a)}{u''(w(\alpha, y))} \tag{III.68}$$

where $K \equiv a/(\int h(x)u'(xa)dx)$.□

Proof of Proposition 4.1: Suppose the wage is given by Equation (III.9),

$$\max_{\eta(.),a,A} \int h(\alpha) \int u(w(\alpha, y) + \eta(\alpha) - A)f(y|a)dyd\alpha - \psi(a)$$

s.t. $\int q(\alpha)\eta(\alpha)d\alpha = A(\rho)$

Holding a fixed, consider first the hedging choice $\eta(.)$. Note that the problem is concave in each $\eta(.)$ and thus it has a unique solution. The first-order condition in $\eta(.)$ implies:

$$\int h(\alpha)u'(w(\alpha, y) + \eta(\alpha))f(y|a)dy = \rho q(\alpha) \tag{III.69}$$

At $\eta(\alpha) = 0$, this condition can be written:

$$\begin{aligned} \rho q(\alpha) &= \int h(\alpha)u'(w(\alpha, y))f(y|a)dy \\ &= q(\alpha)\int \frac{f(y|a)}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}dy \end{aligned} \tag{III.70}$$

And thus the solution is obtained with $\rho = \int \frac{f(y|a)}{\lambda+\mu\frac{f_a(y|a)}{f(y|a)}}dy$. Thus $\eta(\alpha) = 0$ is an optimum to the problem regardless of a.□

Proof of Proposition 5.1: Suppose the principal anticipates no trading but the agent chooses a non-zero portfolio $(\omega(\alpha))_\alpha$. Then, for a given performance $\pi = \alpha y + \omega(\alpha) - \int q(\alpha)\omega(\alpha)d\alpha$, the principal will (incorrectly) interpret this performance as: $\hat{y} = \pi/\alpha = y + \frac{\omega(\alpha) - \int q(\alpha')\omega(\alpha')d\alpha'}{\alpha}$. The problem of the agent will then be as follows:

$$\max_{\omega(.),\tilde{a}} \int\int h(\alpha)f(y|\tilde{a})u(w(\alpha, y + \frac{\omega(\alpha) - A}{\alpha}))dyd\alpha - \psi(\tilde{a})$$

$$\text{s.t.} \qquad \int q(\alpha)\omega(\alpha) = A$$

The multiplier associated to the constraint is denoted β. Taking the first-order condition with respect to $\omega(\alpha)$ on this problem: for all α,

$$h(\alpha)\int f(y|a)w_y(\alpha, y)u'(w(\alpha, y))dy = \lambda\alpha q(\alpha) \tag{III.71}$$

Assume next that the contract is a solution to Equation (III.9). Integrating both sides with respect to α,

$$lambda = \int h(\alpha)\int f(y|a)w_y(\alpha, y)u'(w(\alpha, y))dy/\alpha d\alpha \tag{III.72}$$

By Proposition 3.3, $w_y > 0$ and thus $\lambda > 0$. Next, one may substitute u' in Equation (III.71) using Equation (III.9),

$$\int f(y|a)\frac{w_y(\alpha, y)}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}dy = \lambda\alpha \tag{III.73}$$

Differentiating both sides with respect to α,

$$\int f(y|a)\frac{w_{y,\alpha}(\alpha, y)}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}dy = \lambda \tag{III.74}$$

Under non-decreasing absolute risk-aversion, $w_{y,\alpha} \leq 0$ (by Proposition 3.4), a contradiction.

In the case of a logarithmic utility function, one may use Equation (III.15) to simplify Equation (III.73),

$$\int f(y|a)\frac{a(\frac{f_a(y|a)}{f(y|a)})'}{\lambda + \mu\frac{f_a(y|a)}{f(y|a)}}dy = \lambda \tag{III.75}$$

Thus, $\omega(\alpha) = 0$ satisfies the optimality condition for the problem of the agent.$\square$

Proof of Proposition 5.2: Suppose that the principals believe that all agents will choose a. The problem of finding the best possible joint deviation is given as follows:

$$\max_{a'} \int\int h(\alpha)f(y|a')u(w(\alpha a'/a, ya/a'))dyd\alpha - \psi(a')$$

Taking the first-order condition on this problem yields:

$$
\begin{aligned}
\psi'(a') &= \int\int f(y|a')u'(w(\alpha a'/a, ya/a'))(w_\alpha(\alpha a'/a, ya/a')/a - \frac{a}{(a')^2}w_y(\alpha a'/a, ya/a'))dyd\alpha \\
&\quad + \int\int h(\alpha)f_a(y|a')u(w(\alpha a'/a, ya/a'))dyd\alpha
\end{aligned} \tag{III.76}
$$

A sufficient condition for the existence of a profitable joint deviation is that this Equation is violated at $a = a'$. Further, note that the principal offers a contract such that incentive-compatibility is satisfied at a. This readily implies that:

$$
\int\int f(y|a')u'(w(\alpha, y))(w_\alpha(\alpha, y) - w_y(\alpha, y))dyd\alpha = 0 \tag{III.77}
$$

Substituting w_α and w_y from Equations (III.58) and (III.59) yields Equation III.22.□

Proof of Proposition 5.3: Denote $\tilde{\rho}$ the multiplier associated to the budget constraint. The first order condition in Problem $(\tilde{Q})$ is:

$$
\int h(\alpha)u'(w(\hat{\alpha}, \hat{y}) + \eta(\alpha))f(y|a')dy = \tilde{\rho}q(\hat{\alpha}) \tag{III.78}
$$

Suppose that not trading is optimal. Substituting $u'(.)$ from Equation (III.9),

$$
\int H(\hat{y})f(y|a')dy = \tilde{\rho}\frac{h(\hat{\alpha})}{h(\alpha)} \tag{III.79}
$$

As before, $\tilde{\rho} > 0$. The left-hand side does not depend on α. Thus, it must hold that $h(\alpha) = h(\alpha a'/a)$. This implies that $h(\alpha)$ is constant in α on $\mathbb{R}^+$, a contradiction.□

Proof of Proposition 6.1: There are two types of equilibria: (i) $\overline{a}$ is sustained in a competitive equilibrium, (ii) $\underline{a}$ is sustained in a competitive equilibrium.

(i) Suppose that $\overline{a}$ is chosen.

Case 1: the participation does not bind (i.e. the agent achieves a utility greater than R). Then, supposes that the firms set: $w(\alpha, \underline{y}) = \underline{w}$ and $\int w(\alpha, \overline{y})h(\alpha)d\alpha = W$. The incentive-compatibility must bind, therefore:

$$
W = \frac{B}{p(\overline{a}) - p(\underline{a})} \tag{III.80}
$$

Reinjecting this Equation in the market-clearing Equation:

$$\underline{w} + \frac{p(\overline{a})B}{p(\overline{a}) - p(\underline{a})} = \overline{\alpha a} \tag{III.81}$$

This is possible but, as explained in Corollary 6.1, more of a knife-edge situation.

Then, it must be checked that the firm does not offer a contract that elicits $\underline{a}$. Clearly, this can be achieved by setting a wage equal to R always and will lead to a surplus $\overline{a\alpha} - R$. Note that, in the equilibrium effort, the firm achieves zero profit, so that this other contract is preferred whenever $R < \overline{\alpha}\underline{a}$. To summarize case 1 is possible only when Equation (III.81) holds and $R \in [\overline{\alpha}\underline{a}, \overline{\alpha a} - B]$.

Case 2: the participation only binds (both constraints binding being obvious). Then, by market clearing, $R = \alpha\overline{a} - B$. The wage is now constructed as in Case 1 (to maximize incentives). Reinjecting this wage into the market-clearing constraint:

$$\alpha\overline{a} = p(\overline{a})(W - \underline{w}) + \underline{w} \tag{III.82}$$

Solving for W,

$$W = \underline{w} + \frac{\alpha\overline{a} - \underline{w}}{p(\overline{a})} \tag{III.83}$$

Finally, incentive-compatibility needs to be verified:

$$\overline{\alpha a} \geq \underline{w} + \frac{Bp(\overline{a})}{p(\overline{a}) - p(\underline{a})} \tag{III.84}$$

This bound implies the first part of Proposition 6.1.

(ii) Suppose all firms choose $\underline{a}$. Then, the participation must bind at $R = \overline{\alpha}\underline{a}$. One needs next to check whether there is a profitable deviation to offer a contract $\overline{a}$. There are two possible cases.

Case 1: Only the incentive-compatibility binds when trying to elicit $\overline{a}$. Then, the principal will offer a contract that is the same as earlier: $w(\alpha, \underline{y}) = \underline{w}$ and $w(\alpha, \underline{y}) = \underline{w} + B/(p(\overline{a}) - p(\underline{a}))$. For this alternative contract to be profitable for the principal, it must be that:

$$\overline{\alpha\alpha} - \underline{w} - \frac{p(\overline{a})B}{p(\overline{a}) - p(\underline{a})} \tag{III.85}$$

In addition, the participation of the agent must be satisfied at this alternative contract.

$$\underline{w} + \frac{p(\overline{a})B}{p(\overline{a}) - p(\underline{a})} \geq B + \underline{a}\overline{\alpha} \tag{III.86}$$

Case 2: Only the participation binds when trying to elicit $\overline{a}$. Then, the principal will offer the same contract as in case 1 but, from the participation:

$$\overline{\alpha}\underline{a} + B = \underline{w} + p(\overline{a})(W - \underline{w}) \tag{III.87}$$

Note that this new contract will always be preferred by the principal. Finally, checking incentive-compatibility yields:

$$\overline{\alpha}\underline{a} + B \geq \underline{w} + \frac{p(\overline{a})B}{p(\overline{a}) - p(\underline{a})} \tag{III.88}$$

Collecting each of these cases yields Proposition 6.1. Corollary 6.2 is immediate.□

Proof of Lemma 7.1: The following holds:

$$\begin{aligned} 0 &= \int\int (w_k(\alpha, y) - \phi_k(\alpha, y))f^k(y|a_k)dyd\updownarrow(k) && \text{(III.89)} \\ &= \int q(\alpha)\int\int (w_k(\alpha, y) - \phi_k(\alpha, y))f^k(y|a_k)dyd\updownarrow(k)d\alpha && \text{(III.90)} \\ &= \int\int\int q(\alpha)(w_k(\alpha, y) - \phi_k(\alpha, y))f^k(y|a_k)dyd\alpha d\updownarrow(k) && \text{(III.91)} \\ &= \int Rd\updownarrow(k) = R && \text{(III.92)} \end{aligned}$$

The first Equation is true for all α by market clearing. The second Equation follows by pre-multiplying with respect to $q(\alpha)$ and integrating with respect to α. The third Equation is true by Fubini's theorem. The fourth Equation follows from Equation (III.31).□

Proof of Lemma 7.2: Since the argument given in Proposition 3.2 is unchanged, $a > \underline{a}$ is still elicited. Suppose that $\lambda_k = 0$. Then, $G_k(y) = 0$ for all y, a contradiction. Suppose that $\mu_k \leq 0$, then $w_k(\alpha, y)$ is decreasing in y. Define y_0 such that $f_a^k(y_0|a) = 0$.

$$\begin{aligned} &\int\{\int_{-\infty}^{y_0} f_a^k(y|a)u_k(w_k(\alpha, y))dy + \int_{y_0}^{+\infty} f_a^k(y|a)u_k(w_k(\alpha, y))dy\}d\alpha && \text{(III.93)} \\ \leq &\int h(\alpha)\{\int_{-\infty}^{y_0} f_a^k(y|a)u_k(w_k(\alpha, y_0))dy + \int_{y_0}^{+\infty} f_a^k(y|a)u_k(w_k(\alpha, y_0))dy\}d\alpha = 0 && \text{(III.94)} \end{aligned}$$

This contradicts $a > \underline{a}$ and Equation (III.32).□

Proof of Lemma 7.3: The properties of $G_k(y)$ follow immediately from Lemma 7.2. Suppose next that $T_k(\alpha)$ is not strictly decreasing in α for some k. Then, there exists $\alpha' < \alpha$ such that $T_k(\alpha') \leq T_k(\alpha)$.

For all k',

$$\begin{aligned} T_{k'}(\alpha') &= \lambda_{k'} \frac{q(\alpha)}{h(\alpha)} && \text{(III.95)} \\ &= \frac{\lambda_{k'}}{\lambda_k} T_k(\alpha') && \text{(III.96)} \\ &\leq \frac{\lambda_{k'}}{\lambda_k} T_k(\alpha) && \text{(III.97)} \\ &\leq T_{k'}(\alpha) && \text{(III.98)} \end{aligned}$$

Therefore $T_k(\alpha') \leq T_k(\alpha)$ for all k.

Next, $T_k(\alpha') \leq T_k(\alpha)$ implies that for all y, $u'_k(w_k(\alpha', y))G_k(y) \leq u'_k(w_k(\alpha, y))G_k(y)$ and therefore by Lemma 7.2, $w_k(\alpha', y) \geq w_k(\alpha, y)$ for all y.

Collecting these two observations, $w_k(\alpha', y) \geq w_k(\alpha, y)$ for all y and k. But then:

$$\int \int f_k(y|a_k) w_k(\alpha', y) dy d\updownarrow(k) \geq \int \int f_k(y|a_k) w_k(\alpha, y) dy d\updownarrow(k) \qquad \text{(III.99)}$$

Therefore:

$$\int \int f_k(y|a_k)(\phi_k(\alpha', y) - \phi_k(\alpha, y)) dy d\updownarrow(k) \geq 0 \qquad \text{(III.100)}$$

But $(\phi_k(\alpha', y) - \phi_k(\alpha, y))$ is strictly negative, a contradiction.□

Proof of Proposition 7.1: Define $M_k(y) = 1/G_k(y)$. The following holds:

$$\begin{aligned} w_{k\alpha}(\alpha, y) u''_k(w_k(\alpha, y)) &= T'_k(\alpha) M_k(y) && \text{(III.101)} \\ w_{ky}(\alpha, y) u''_k(w_k(\alpha, y)) &= T_k(\alpha) M'_k(y) && \text{(III.102)} \\ w_{k\alpha y}(\alpha, y) u''_k(w_k(\alpha, y)) + T'_k(\alpha) T_k(\alpha) M'_k(y) M_k(y) \frac{u'''_k(w_k(\alpha, y))}{(u''_k(w_k(\alpha, y)))^2} &= T'_k(\alpha) M'_k(y) && \text{(III.103)} \end{aligned}$$

Finally note that:

$$\frac{w_{k,\alpha,y}(\alpha,y)u_k''(w_k(\alpha,y))}{T_k'(\alpha)M_k'(y)} = 1 - M_k(y)T_k(\alpha)\frac{u_k'''(w_k(\alpha,y))}{(u_k''(w_k(\alpha,y)))^2} \quad \text{(III.104)}$$

$$= 1 - u_k'(w_k(\alpha,y))\frac{u_k'''(w_k(\alpha,y))}{(u_k''(w_k(\alpha,y)))^2} \quad \text{(III.105)}$$

$$= \frac{\partial}{\partial y}(\frac{u_k'}{u_k''})(w_k(\alpha,y)) \quad \text{(III.106)}$$

In the DARA case, $w_{k,\alpha,y} > 0.\Box$

Chapter IV

Informational Limits of Arbitrage

1 Abstract

This essay studies the informational determinants of arbitrage capital flows. In the model, outside investors have imperfect information about the potential risk taken by arbitrageurs, which creates an informational asymmetry to be resolved. I explore the relationship between primitive characteristics of the arbitrage (such as execution cost, potential risk and arbitrage return) and capital flows. Risk and cost decrease the amount of capital received by arbitrageurs, as expected. Counter-intuitively, however, high-return arbitrages may attract less capital than low-return ones. This situation may create multiple competitive equilibria: one of them price efficient, and others with prices far away from fundamentals. Implications for asset price instability, managerial compensation (high-water marks) and hedge fund capital flows are also discussed.

2 Introduction

The law of one price in financial markets is deeply rooted in the idea of perfect arbitrage. In the canonical model, a single competitive arbitrageur can raise a large amount of capital and trade against any profitable arbitrage. Existing research points out that arbitrageurs with limited capital may be unable to exploit any profitable arbitrage, since they may become insolvent prior to its conclusion. Yet, a complete theory of the limits of arbitrage also needs to explain the economic frictions preventing arbitrageurs from initially raising a large amount of capital. Further, the theory may help better understand the relationship between characteristics of arbitrages and the magnitude of the capital flows working to correct them.

This essay shows that the limited capital inflows received by arbitrageurs may be caused by informational asymmetries between arbitrageurs and outside investors. In the model, an arbitrageur learns of a risk-free profitable trade (i.e., a *perfect* arbitrage). Investors can provide outside capital; they have, however, little information on the potential risks taken by the arbitrageur. This creates an informational asymmetry to be resolved. Specifically, two actions are neither observable nor contractible. One, the arbitrageur can pursue investment activities that are privately costless (e.g., passive investment) but do not create value. Thus, the compensation arrangement must give incentives to actively implement the arbitrage. Two, arbitrageurs have limited personal wealth and thus do not bear the full consequences of their losses. To take advantage of certain incentive schemes, they may strategically choose to engage in fair gambles.

I show that arbitrageurs may decrease the amount of capital invested in the arbitrage and speculate to mimic the returns produced by an arbitrageur investing the full amount. The main result of the essay is that informational frictions can prevent certain profitable arbitrage opportunities from being exploited and, more generally, tighten the supply of arbitrage capital. The key trade-off in the essay is that, on one hand, raising a large amount of capital increases the profit from the arbitrage while, on the other hand, it makes speculation more attractive and thus increases agency cost. Indeed, the quantity of capital raised by the arbitrageur is decreasing in the maximum potential loss in a component of the arbitrage (e.g., an uncovered position). In the limit, no capital can be profitably raised for an arbitrage involving positions with unbounded potential loss.

Counter-intuitively, the capital raised by arbitrageurs is non-monotonic in the returns from the arbitrage. When the cost of executing the arbitrage is small, profitable arbitrages with a *low* net return on initial capital may not generate any capital inflows. This is because the return from the arbitrage is not sufficiently high to pay for the cost of giving incentives not to speculate. When the cost of implementing the arbitrage is large, however, arbitrages with a *high* net return cannot be exploited. For such cases, the large returns from the arbitrage increase how much the agent can speculate and make it too costly for investors to provide incentives.

In a competitive equilibrium, I show that asset prices can be responsive to liquidity shocks unrelated to fundamentals. I explain the intuition for this result next. Suppose an asset is overpriced and can be purchased by arbitrageurs at its correct fundamental value in another market (a violation of the law of one price). When the cost of the arbitrage is large, less capital is supplied to arbitrageurs when the return from the arbitrage is high. Thus, arbitrageurs will supply *fewer* units of the asset when its price is further away from its efficient price (since the price differential causes high arbitrage returns). Thus, the total supply of the asset (arbitrageurs and regular sellers) may decrease in price.

Whether price-taking arbitrageurs are able to correct the price inefficiency, then, depends on whether investors believe that the price inefficiency will indeed be corrected. This property can cause multiple competitive equilibria. If investors believe that prices will be inefficient, they will anticipate high-return arbitrages, and thus will not provide any capital (confirming expectations that prices remain inefficient). If investors believe that price inefficiencies will be corrected, they will anticipate low-return arbitrages, and thus will provide capital (confirming expectations that the price inefficiency is corrected). I find up to three competitive equilibrium outcomes: (a) no profitable arbitrage and (near) price efficiency, (b) prices only partially move toward the efficient price, (c) large price inefficiency and arbitrageurs do not receive any capital. Given liquidity shocks or additional entry by new arbitrageurs, (b) may cease to be an equilibrium thus causing a discontinuous price correction from equilibrium (b) to equilibria (a) or (c). In this respect, the model suggests that the potential variations in the capital flows received by arbitrageurs may be one of the drivers of asset price volatility, in particular for securities in which informational asymmetries between uninformed outside investors and informed arbitrageurs are significant.

Related Literature

The main contribution of this essay is to attempt to describe the informational frictions that may be at play preventing certain profitable arbitrages opportunities from being exploited. My findings have implications on several branches of the economic literature in which such frictions are typically mentioned but not explicitly modeled.

Following Shleifer and Vishny (1997), several authors explore how limited access to capital affects the trading patterns of arbitrageurs (see Cornelli and Li (2002), Liu and Longstaff (2004), Morris and Shin (2004), Mukarram and Mellob (2006)). In these papers, limited access to capital is typically exogenously assumed and motivated as the outcome of some informational friction. This essay provides an economic rationale for this assumption. Further, the model describes how the amount of capital received by arbitrageurs may itself also depend on the characteristics of the arbitrage. I further link the limits of arbitrage to fundamental characteristics of the informational environment, and not only "real" economic fluctuations. In particular, I show that capital flows depend on the quality of the risk controls (i.e., whether investors can prevent risk-taking) and whether execution costs are easily contractible. In addition, as in the corporate finance literature (Holmström and Tirole (1997), Tirole (2006), Biais, Mariotti, Plantin and Rochet (2007)), I show that the arbitrageur's personal wealth is key to resolve the informational friction. In particular, my framework predicts that economies in which arbitrageurs have more personal wealth should feature fewer price inefficiencies.

The approach further contributes to several bodies of empirical literature. A first set of papers explores whether demand curves for financial securities slope downward in response to exogenous demand shocks or, as suggested by standard arbitrage pricing arguments, stay flat. Kaul, Mehrotra and Morck (2000) show that prices seem to respond permanently to changes in weights in the Toronto Stock Exchange 300 index, as a result of the trades of index funds. Mitchell, Pulvino and Stafford (2002) examine cases in which the price of a firm becomes different from the total value of its subsidiaries. Bollen and Whaley (2004) show that implied volatilities for options are sensitive to the order flow. Other papers confirming these findings for other investment classes include Brennan and Schwartz (1990), Griffin and Topaloglu (2003), Gagnon and Karolyi (2003), Cherry (2004), Blouin, Hail and Yetman (2005), Hau, Massa and Peress (2005) and Greenwood (2005). As a rationale to this evidence, this essay finds that demand curves would be flat without any informational frictions but may slope down (and even become non-monotonic) when the risk of the arbitrage and its execution cost are too large.

In a related area, several authors discuss the investment strategies of arbitrageurs during periods commonly described as "bubbles" and argue that, empirically, arbitrageurs seem to be unable to correct certain price inefficiencies. Several authors provide evidence that hedge funds were present during the internet "bubble" and avoided most of the downturn, although they did not trade enough to fully correct the bubble (Ofek and Richardson (2003), Brunnermeier and Nagel (2004)). In other examples, several authors suggest evidence of uncorrected "excessive" market movements (Garber (2000), Lowenstein (2004), Shiller (2005)). The model provides conditions under which bubbles cannot be arbitraged because arbitrageurs do not receive enough capital. In particular, a bubble can be self-fulfilling: if prices move too far away from their arbitrage price, then returns from the arbitrage may become too large for arbitrageurs to receive any capital.

The findings are also related to an extensive literature on the life-cycle of hedge funds which are the closest empirical match to who I describe in this essay as "arbitrageurs." An intriguing fact in the hedge fund literature is that, unlike in the case of mutual funds, well-performing funds tend to close their funds to new investors (see Jagannathan, Malakhov and Novikov (2006), Horst and Verbeek (2007), among others). Although observers typically argue that this feature may be due to the lack of additional investment opportunities, there is no simple reason why a decrease in the number of profitable arbitrages should be correlated to a high past performance. The model provides a different rationale. Under certain conditions, high-return arbitrages may be difficult to arbitrage and receive less capital. If investors learn that the arbitrage is more profitable, it may become desirable to provide *less* capital.

Finally, the model provides several additional implications in terms of the optimal contractual arrangement for hedge fund managers. Most hedge funds use compensation schemes known as *high-water marks*: fund managers are paid a proportional fee of total fund size (management fees between 1.5% to 4%) and, if they perform above a certain benchmark, an additional performance fee (usually 20% of additional performance). Although several authors assess the appropriate parametrization for such schemes (Goetzmann, Ingersoll and Ross (2003), Panageas and Westerfield (2005), Kouwenberg and Ziemba (2005), Hodder and Jackwerth (2007)), the prevalence of this particular contract is still an open question. This essay provides some preliminary intuition why such contracts may be desirable. In the model, the participation of agents with low-return arbitrages binds, and thus they are paid exactly their management cost (a proportion of total fund size). With high-return arbitrages, however, incentive-compatibility binds, and arbitrageurs are paid a proportion of realized performance.

3 The Model

An arbitrageur (hereafter agent) learns of an arbitrage opportunity that provides a certain profit $r > 0$ per dollar of initial capital invested.[1] Here, the arbitrage is defined as a trading strategy that requires a certain amount of capital immobilized for the period of the arbitrage; r is also interpreted as a measure of the price inefficiency that can be exploited. Let $a \in [0, \overline{a}]$ denote the total amount invested in the arbitrage, where $\overline{a}$ is finite but large.[2] In the rest of the paper, $\mathcal{F}_{ra}(.)$ denotes the distribution of a random variable almost always equal to ra; it corresponds to the distribution of the profit produced by an arbitrageur investing a in the perfect arbitrage.

The agent has a utility function $w - \gamma a$, where w represents monetary compensation and $\gamma > 0$ represents a private cost associated to implementing the arbitrage (its interpretation is similar to Berk and Green (2004), i.e. transaction or management/organizational cost).[3] I extend the model to non-linearities in money in Section 7; however, it is useful first to focus on a problem in which the agent's problem features constant returns to scale (and thus capital flows are always large in the absence of informational frictions).[4] In the model, investing more capital is more costly to the agent. For example, hedge funds charge management fees proportional to fund size (versus fixed fees) which are generally motivated as

[1] This is a net profit. If $\$1$ is invested, the gross proceeds will be $\$1 + r$. The risk-free rate is normalized to one.

[2] This upper bound is for technical reasons since risk-neutrality will lead to bang-bang policies $a \in \{0, \overline{a}\}$. All results are presented for $\overline{a}$ large (to save space, omitting cases that disappear as $\overline{a}$ grows large).

[3] Note that γ represents private non-contractible cost and not cost that can be fully expended to investors (in such cases, they would be included as part of r).

[4] A non-linear convex cost function $f(a)$ (instead of γa) delivers results that are qualitatively similar to those of Section 7. If the cost function is concave, as long as $\overline{a}$ is chosen to be large, the solution of the model will be nearly identical to the solution of this model but choosing $\gamma \equiv \lim_{a \to +\infty} f'(a)$ (since concavity will only reinforce the bang-bang nature of the problem and only solutions $a \in \{0, \overline{a}\}$ will be chosen). The only important assumption is that $f'(a)$ does not converge to zero when a becomes large, or else cost becomes negligible relative to the total capital invested.

operating expenses.[5] I motivate this assumption as non-contractible trading costs or the attention cost of finding mispriced assets to invest a dollars of capital.[6]

In the model, an arbitrage opportunity is profitable if $r > \gamma$, that is, if its profit is greater than its cost. Because an infinitely-wealthy agent would be able to trade away any arbitrage opportunity without using any outside capital, it is assumed that the agent has limited wealth $-\underline{w} > 0$. Investors offer $(A, w(.))$, where A is the total initial capital given to the agent and $w(y) \geq \underline{w}$ is the wage paid when a profit y is reported. Investors observe final profit but are unable to monitor how A is invested.

Specifically, given initial capital A, two decisions are unobservable to outside investors. First, the agent chooses the amount invested in the arbitrage, namely $a \in [0, A]$. Any choice of $a < A$ corresponds to the agent shifting a capital $A - a$ to passive risk-free investments that do not generate value (their return is zero) but are privately costless. Unlike in the mutual fund industry, for example, hedge funds have very few reporting obligations and monitoring their actual portfolio may be difficult. Second, when investing a, the agent may also choose to enter in fair gambles, in other words, implement a profit y with distribution $\hat{F}(.)$ and mean lesser or equal than ra. When $\hat{F} \neq \mathcal{F}_{ra}$, the agent is strategically choosing an imperfect arbitrage. Assume that $\hat{F}(.)$ must have support over $[\theta A, +\infty)$, where $\theta \in [-\infty, r]$ is an exogenous parameter representing the maximum loss limit in a position required to be able to implement one dollar of arbitrage.

The parameter θ also describes how well the risk controls implemented by investors can be used to monitor risk-taking by the agent. As is intuitive, this assumption captures that the total maximum risk required to be able to implement the arbitrage is greater in the case of a large fund. For example, in the context of this model, an agent with enough capital to implement ten (simultaneous) Put-Call arbitrages would need ten times the maximum risk required for an agent with only enough capital to implement one Put-Call arbitrage.[7]

The model is intended to capture an important feature of the relationship between arbitrageurs and outside investors. If investors do not know the actual details of the arbitrage, they will not be able to perfectly monitor whether the agent is actively implementing the arbitrage, or shifting funds to other

[5] Note that the true total operating expenses are unobservable to investors and private information to the hedge fund management.

[6] As a simple example, consider the case of an agent engaging in Put-Call arbitrages. The agent may either wait and do nothing, or watch current orders and take mispriced options until he/she runs out of capital. Presumably, except in the (rare) case of completely automated arbitrages, one would expect the total cost to be greater given larger quantities of capital to manage.

[7] The assumption that the maximum loss is per-dollar invested is not the driver of limited capital flows: there may be limited capital flows even if the maximum loss is in absolute terms and does not depend on total capital invested. As $\overline{a}$ grows large, an absolute maximum downside risk can be obtained as a special case of the framework by setting $\theta = 0$ (since in this case the level of the downside risk will be small as compared to total capital).

investments that may involve more risk or lower returns. An arbitrage that is, in theory, risk-free and profitable can potentially turn into a risky and low-value investment if the agent is not given appropriate incentives.

Several examples are given next to intuitively illustrate what is represented by θ. To implement a Put-Call arbitrage, the agent will need to short-sell either the Call or the underlying share. The arbitrage, although risk-free if implemented, involves components with unbounded possible loss (i.e., holding short the Call or the share). In the (extreme) case in which investors do not monitor the total risk of the agent's portfolio, they will have to allow for positions with unbounded possible losses, namely $\theta = -\infty$. If the agent strategically chooses not to implement the arbitrage, he may use such uncovered positions to achieve any distribution $\hat{F}(.)$ with support over $\mathbb{R}$ and ex-ante claim that these operations were part of the actual arbitrage.

On the other hand, investors may be able to establish risk controls such that any short-selling of a share is matched to a purchase of a Call. In this case, these arbitrageurs will be constrained to choose perfect arbitrages (i.e., any uncovered transaction would be detected) and thus $\theta = r$. Of course, outside the simple case of the Put-Call parity, such situations should be fairly rare since, if the exact arbitrage were public information, outside investors could directly implement the arbitrage without hiring the agent. The most common case would be in-between situations in which implementing the arbitrage requires some loss limit $\theta \in (-\infty, r)$ which is not fully monitored.[8] As an example, $\theta = 0$ corresponds to situations in which investors cannot make a loss on their capital, while $\theta = -1$ corresponds to cases in which investors with limited liability may lose at most their capital. Situations with $\theta < -1$ correspond to cases in which the arbitrageur's positions are backed with the assets from a financial institution (e.g, proprietary trading desks).

Assume the agent must receive an expected utility $b > -\underline{w} > 0$. Clearly, there would be no purpose in giving excess capital $A > a$ to the agent and thus one may restrict the attention to contracts such that actively investing $a = A$ is elicited. An optimal contract is defined as $\Gamma = (a, w(.), \hat{F}(.))$ where $a = A$ is the initial capital raised, $w(.)$ is the monetary compensation and $\hat{F}(.)$ is the distribution of the profit chosen by the arbitrageur. The optimal contract solves the following problem.

$$(P) \qquad \max_{a, w(.) \geq \underline{w}, \hat{F}(.)} \int_{a\theta}^{+\infty} (y - w(y)) d\hat{F}(y)$$

[8]Later in the paper, I show that this loss limit can also be endogenized from primitives on the volatility of the assets.

$$s.t. \quad \int_{a\theta}^{+\infty} w(y)d\hat{F}(y) - \psi(a) \quad \geq b \tag{IV.1}$$

$$(a, \hat{F}(.)) \in \arg\max_{\tilde{a} \leq a, F(.)} \int_{a\theta}^{+\infty} w(y)dF(y) - \psi(\tilde{a}) \tag{IV.2}$$

$$\text{s.t.} \quad \int_{a\theta}^{+\infty} ydF(y) \leq r\tilde{a} \tag{IV.3}$$

$$\int_{a\theta}^{+\infty} dF(y) = 1 \tag{IV.4}$$

The agent chooses a perfect arbitrage when $\hat{F}(.) = \mathcal{F}_{ra}(.)$ and speculates when $\hat{F}(.) \neq \mathcal{F}_{ra}(.)$.[9] Finally, assume that investors have to hire the agent and the agent needs investors to execute the arbitrage. For example, the agent may be involved in other activities (e.g., marketing, general management) that are not directly related to the investment problem but have value greater than b.[10]

4 Preliminaries

4. 1 First-Best

A first-best solution to the model is defined by omitting Equation (IV.2), that is, investors control both investment and risk. Clearly, investors bind the participation of the agent: $W^* = b + \gamma a^*$. The next statement follows immediately.

Proposition 4.1. *In first-best, the following capital flows are optimal:*

(i) If $r = \gamma$, $a^* \in [0, \overline{a}]$.

(ii) If $r > \gamma$, $a^* = \overline{a}$.

In first-best, all profitable arbitrage opportunities are exploited. There are also other cases such that the first-best outcome can be achieved. First, first-best can be achieved if investors can contract on a but not on the distribution chosen by the agent $\hat{F}(.)$. This is because investors could then offer a flat payment W^*. Second, first-best can also be attained if $\theta = r$ or investors observe whether a perfect or imperfect arbitrage is chosen. In such cases, investors could offer $w(a^*) = W^*$ if a perfect arbitrage is chosen and $w(y) = \underline{w}$ otherwise.

[9] I view here θ as the expertise of investors, so that arbitrageurs need outside investors to establish risk controls $\theta > -\infty$. Alternatively, one may assume that the arbitrageur could also invest his own funds without outside investors, which would imply a lower bound for b. This bound can be easily derived formally. To begin with, the end-of-period wealth of the arbitrageur must always satisfy the limited liability (even in the worst case), so that the arbitrageur must invest a such that: $(1-\theta)a \leq -\underline{w}$. This means that the arbitrageur can invest a maximum of $\underline{w}/(1-\theta)$ ($+\infty$ if $\theta \geq 1$). If $\theta \geq 0$, then, the arbitrageur invests $-\underline{w}$ and can earn a utility $-\underline{w}(r-\alpha)$, and therefore $b \geq -\underline{w}(r-\alpha)$. The case with $\theta < 0$ is similar (although the investor would invest $a < -\underline{w}$ and keep some wealth as collateral).

[10] This assumption is not important for the qualitative results, but removes a number of additional cases to be considered when the profits from the arbitrage are not sufficient to pay for the cost of the agent b.

4. 2 Perfect Arbitrage

The model is solved next when the decision to invest and speculate is under the control of the agent (second-best), that is, as stated in problem (P). Figure IV.1 describes how the agent will respond to a particular compensation and initial capital a. The bold curve $w(.)$ represents some (arbitrary) compensation given to the agent and the dotted curve $\gamma a'$ represents the personal cost to the agent.

Suppose, for example, that the agent chooses some investment $\tilde{a}$. Given $\tilde{a}$, two types of arbitrage are possible. First, the agent may choose a perfect arbitrage and receive a compensation $w(\tilde{a})$ (white square). Second, the agent may choose an imperfect arbitrage, for example, choosing a profit with support over the peaks at each side of $\tilde{a}$ (left black square). This second option will generate a higher compensation and thus should be preferable to a perfect arbitrage. Repeating this argument for all possible choices of $\tilde{a}$, the presence of imperfect arbitrages lets the agent reach up to $\tilde{w}(y)$ (small dots), the smallest concave function greater than $w(.)$.

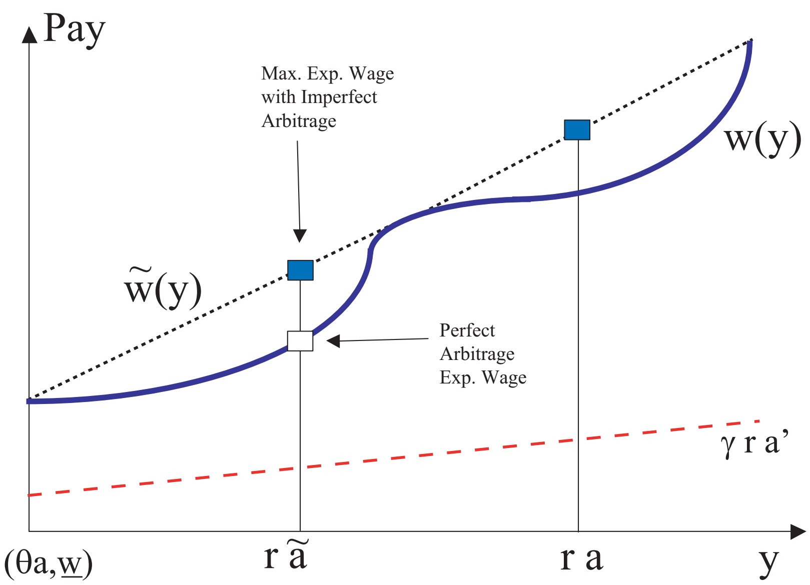


Figure IV.1. Concavification of the Wage

A utility-maximizing agent will choose the investment that maximizes the difference between $\tilde{w}(a')$ and $\gamma a'$. In the case of Figure IV.1, for example, investing the initial capital a is incentive-compatible but involves the agent choosing an imperfect arbitrage.

Note that investors are risk-neutral and thus are indifferent between: (i) offering $w(.)$ and letting the agent choose an imperfect arbitrage to reach up to $\tilde{w}(ra)$, and (ii) offering $\tilde{w}(.)$ directly and eliciting a perfect arbitrage instead. As a result, one may restrict the attention to compensation arrangements eliciting perfect arbitrages without loss of optimality.[11] This observation is formalized next and proved in the Appendix.

Lemma 4.1. *There exists an optimal contract eliciting a perfect arbitrage.*

In a perfect arbitrage, a distribution of the form $\mathcal{F}_{ra}(.)$ must be elicited. To achieve this distribution, minimizing wages for any off-equilibrium outcome $y \neq ra$ is optimal. Thus, one may restrict the attention without loss of optimality to contracts of the following form: (i) investors offer capital a, (ii) the wage prescribes $w(ra) > \underline{w}$ and $w(y) = \underline{w}$ for any $y \neq ra$, (iii) the agent chooses $\mathcal{F}_{ra}(.)$.[12] In short-hand, a contract of this form is denoted $< W, a >$, where $W = w(ra)$ is the wage paid when reporting a profit ra and a is the amount of capital to be invested.

4. 3 Incentive-Compatibility

Given that the agent chooses a perfect arbitrage, the reported profit y will be perfectly informative on the choice of the agent. This property may suggest that first-best (and thus Proposition 4.1) should also hold. This is not (always) the case here. Figure IV.2 gives an example where first-best fails. To begin with, suppose that the agent is given a contract $< W^*, a^* >$, as defined earlier. This contract seems to be geared to elicit a^* since it minimizes wages over states that should not occur when the agent invests a^*.

The agent does not choose first-best in this example. In Figure IV.2, the rightmost square corresponds to the first-best choice; it implies no rents for the agent. The agent may also speculate and choose any distribution over $y = \theta a^*$ (the worst-case outcome) and $y = ra^*$ (the only outcome with pay above $\underline{w}$): the set of feasible utilities (gross of private cost) is delineated by the straight line (M). Suppose the agent deviates to underinvest to $a < a^*$ and mimics an investment equal to a^* by speculating. In the plot, the investment choice that maximizes the utility of the agent is $a = 0$, i.e. no capital is invested.

Proposition 4.2. *If $\theta = -\infty$, $a = 0$ (no capital) must be elicited.*

In the case of an unbounded maximal loss, the agent will not be able to raise any capital, regardless of the potential return of the arbitrage. Given the ability to choose imperfect arbitrages, the agent can respond to any incentive scheme by choosing an imperfect arbitrage that exposes the firm to very large losses with a small probability. He can then use this imperfect arbitrage to almost always achieve the maximum possible compensation. Here, these imperfect arbitrages act as substitutes for investment. To

[11] There are also other optimal contracts involving imperfect arbitrages; however, they all imply the same equilibrium investment and expected compensation.

[12] A contract of this form is, of course, not the unique optimal contract (for off-equilibrium values of y); but the predicted investment (which is the focus here) is unique.

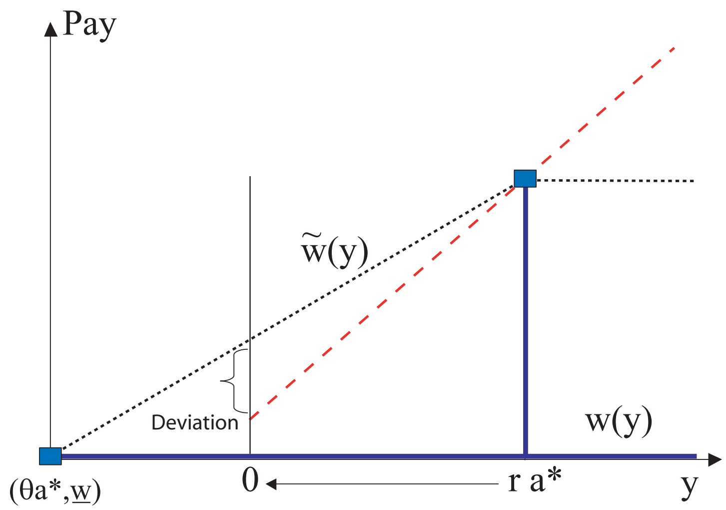


Figure IV.2. Optimal Investment Choice

give content to the contract design problem, therefore, assume in the rest of the essay that $\theta \in (-\infty, r)$ and meets the condition in Lemma 4.2.

Lemma 4.2. *$< w, a >$ is incentive-compatible if and only if*

$$\frac{w - \underline{w}}{ra - \theta a} \geq \gamma \tag{IV.5}$$

Even when the maximum loss is bounded, investors must be mindful of a joint deviation in investment and risk and, to deter unwanted risk-taking, must offer a compensation that mitigates incentives to speculate. Equation (IV.5) reveals that, to ensure that the agent will not speculate, his compensation must increase (at least) by $\gamma(r - \theta)$ for each additional dollar of capital invested. Therefore, when $r > 1 + \theta$, compensation increases faster than what is required by the participation. This causes an indirect incentive cost that is considered by investors when choosing the initial amount of capital.

5 Optimal Capital Flows

5. 1 Small Cost

Let a^{**} denote the second-best capital inflow. Taking into account the incentive cost, there are three possible equilibrium outcomes: (a) a solution is efficient (resp. distortionary) if it prescribes $a^{**} = \bar{a}$

(resp. $0 < a^{**} < \overline{a}$), (b) a solution prescribes no rents (resp. rents) when the participation of the agent binds (resp. does not bind), (c) a solution prescribes market breakdown when $a^{**} = 0$ under $r > \gamma$. Finally, to remove a number of cases that occur only when $\overline{a}$ is small (and disappear when $\overline{a}$ grows large), I make the assumption that, if $r > 1 + \theta$, $\overline{a} > \max(\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}, \frac{b-\underline{w}}{\gamma(r-\theta-1)}$ (*) and restrict the attention to parameter values that satisfy this restriction.

Proposition 5.1. *Suppose $\gamma < 1$, then for all parameter values satisfying (*),*

(i) If $\gamma \leq 1 + \theta$ (low-risk arbitrage),

1. *For $r = \gamma$, $a^{**} \in [0, \overline{a}]$ [no rents].*
2. *For $r \in (\gamma, 1 + \theta]$, $a^{**} = \overline{a}$ [efficient, no rents].*
3. *For $r > 1 + \theta$, $a^{**} = \overline{a}$ [efficient, rents].*

(ii) If $\gamma > 1 + \theta$ (high-risk arbitrage),

1. *For $r \in [\gamma, \underline{r})$, $a^{**} = 0$ [breakdown].*
2. *For $r = \underline{r}$, $a^{**} \in [\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}, \overline{a}]$.*
3. *For $r > \underline{r}$, $a^{**} = \overline{a}$ [efficient, rents].*

where: $\underline{r} = -\gamma\theta/(1-\gamma)$.

Consider first the case in which $\gamma < 1$, that is, the cost of the arbitrage is not too large. Case (i) corresponds to situations in which the maximal loss is small. It includes the situation in which investors cannot make a loss ($\theta = 0$). In this case, any profitable arbitrage will attract large capital flows $\overline{a}$. When r is sufficiently large, the agent is given a rent (the participation does not bind). This additional compensation cost is, however, more than fully paid back with the proceeds from the arbitrage.

Case (ii) corresponds to situations in which the maximal loss is large. It includes, for example, the common case in which investors can lose at most their initial capital ($\theta = -1$). In this case, there exists an informational lower bound $\underline{r}$ corresponding to the minimum return required to be able to exploit the arbitrage. Here, incentive-compatibility makes it necessary to pay the agent more than γ per unit of capital invested. As a result, a return strictly greater than γ is required for the arbitrage to be profitable. This bound becomes more demanding in response to factors that aggravate the agency problem: it decreases in the risk of the arbitrage θ and increases in the cost of the arbitrage γ.

5. 2 Large Cost

The analysis is extended to cases such that $\gamma > 1$, i.e. arbitrages that are very profitable but involve greater implementation cost.[13].

Proposition 5.2. *Suppose $\gamma > 1$, then for all parameter values satisfying (*),*

(i) If $\gamma \leq 1 + \theta$ (low-risk arbitrage),

1. *For $r = \gamma$, $a^{**} \in [0, \overline{a}]$ [no rents].*
2. *For $r \in (\gamma, 1 + \theta]$, $a^{**} = \overline{a}$ is chosen [efficient, no rents].*
3. *For $r \in (1 + \theta, \underline{r})$, $a^{**} = \overline{a}$ is chosen [efficient, rents].*
4. *For $r = \underline{r}$, $a^{**} \in [\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}, \overline{a}]$.*
5. *For $r > \underline{r}$, $a^{**} = 0$ [breakdown].*

(ii) If $\gamma > 1 + \theta$ (high-risk arbitrage), $a^{**} = 0$ [breakdown].

When the cost is large, surprisingly, high-return arbitrages are more difficult to exploit than low-return ones. In the model, a large cost means that incentives to deviate are much greater per unit of capital. By Equation (IV.5), the wage must then increase by $r\gamma$ for each additional unit of capital under management when the incentive-compatibility binds. This wage paid to the agent, therefore, can increase faster than r. In case (i), any high-return arbitrage receives zero capital inflow. That is, arbitrageurs are unable to correct very important price inefficiencies. In case (ii), no arbitrage (even the ones with lower return) can be exploited. In this case, arbitrageurs are unable to correct *any* price inefficiency.

5. 3 Discussion

The model describes which arbitrages can be exploited. In the model, arbitrages have three characteristics: execution cost, potential risk and return. The table below summarizes the main findings:

	Small Cost	Large Cost
Low Risk	All arbitrages are exploited	Only low-return arbitrages
High Risk	Only high-return arbitrages	No arbitrage can be exploited

[13]To save space, the knife-edge case $\gamma = 1$ is not stated. It can be obtained by setting $\underline{r} = +\infty$ and taking (i) in Proposition 5.1 and (ii) in Proposition 5.2

Second-best capital flows and profits can be different from first-best. First, in first-best, the profit of investors grows large as returns increase (at rate $r - \alpha$ per unit of capital invested). In contrast, given high risk and large cost, no arbitrage can be exploited in second-best, even one with a very large return. In fact, given low-risk and large-cost, high-return arbitrages are those receiving no capital. This finding is broadly consistent with the intuition that arbitrageurs are unable to exploit large price inefficiencies due to the (potential) risks than arbitrages involve.

The analysis can shed some light on the causes of the rapid growth in the total capital invested in hedge funds over the last two decades. Between 1992 and 2007, the size of the hedge fund industry grew from approximately $\$100$ billion to more than $\$2.4$ trillion (HFN and Van Hedge Advisors). The standard theory of the limits of arbitrage developed in Shleifer and Vishny (1997) relates arbitrages to asset volatility, yet asset volatility does not seem to have decreased over this period. The theory presented here suggests that a number of factors reducing the informational frictions between arbitrageurs and outside investors may have been one of the causes of this evolution. Following the collapse of the Barings bank and the move of several major banks away from proprietary trading desks, financial investors have moved toward hedge funds with limited liability. This legal innovation (as well as a wider acceptance of margins for limited liability partnerships) may have worked to reduce the maximum risk θ necessary to implement an arbitrage. Simultaneously, the development of greater computing speed and more efficient risk management models may also have contributed to the ability to instantaneously monitor the positions of arbitrageurs. Other factors may have made the actions of arbitrageurs more transparent to investors. Investors have now easy access to databases and performance benchmarks; by classifying funds by asset class and strategy, databases should make it easier to detect inappropriate risk-taking. Finally, more integrated global markets have reduced overall execution cost and may have expanded the expertise of local arbitrageurs from domestic to global markets. In the context of this model, these factors may have expanded the class of arbitrage opportunities that can receive capital.

A second implication of the model concerns the payments received by hedge fund managers. One interesting aspect of hedge fund compensation schemes is the presence of high-water mark contracts. In these contracts, the manager is paid a proportion of total fund size and a performance bonus proportional to returns above a certain target. The model provides some preliminary intuition for such contracts. When the return r is low, the participation of the agent binds, and thus the agent receives $b + \alpha a^{**}$, a proportion of total fund size a^{**}. When the return r is high, the incentive-compatibility constraint binds, and thus the agent receives $\gamma r a^{**} - \gamma \theta a^{**} + \underline{w}$ (Equation (IV.5)). In other words, an incentive fee proportional to performance is paid only for large returns. Offering a high-water mark contract, in this model, would yield optimal wages for a cross-section of different managers with different returns r. Note also that the performance threshold at which the performance payment activates is given by $1+\theta$ if cost is low and $\underline{r}$ when cost is large; both are decreasing in the potential risk of the arbitrage.

5. 4 Solvency Risk

I endogenize next the choice of θ by incorporating into the model solvency risk in the spirit of Shleifer and Vishny (1997). Suppose that, for each dollar invested, the value of the agent's portfolio may temporarily become equal to ρ (prior to the conclusion of the arbitrage), where ρ is a random variable drawn from a continuous distribution with density $h(.)$ and full support over $(-\infty, r]$.

After this temporary price movement, the value of the portfolio will be equal to ρa. Assume that the agent must respect the maximum loss of constraint and liquidates whenever $\rho a \leq \theta a$, that is the value of the portfolio falls below the maximum acceptable loss.[14] Although $h(.)$ is exogenous, the maximum risk θ is now optimally chosen by investors. Assume also that: (i) $h'(\rho)/h(\rho) > 1/(r-\rho)$, (ii) $\lim_{\rho\to-\infty} h(\rho)(r-\rho) > \gamma$, and (iii) there exists θ such that investors make positive profits.[15]

Investors resolve the following trade-off: on the one hand, reducing θ reduces the chances that the positions will be liquidated prior to the conclusion of the arbitrage but, on the other hand, increasing θ makes it easier to give incentives to the agent.

Proposition 5.3. *Let $\tilde{\theta}$ be the (unique) solution to:*

$$\gamma = h(\tilde{\theta})(r-\tilde{\theta}) \tag{IV.6}$$

Then, the optimal choice of downside risk is given by $\theta^* = \min(r-1, \tilde{\theta})$.

Equation (IV.6) shows that θ decreases when γ increases (by condition (ii)). Thus, investors are willing to tolerate more downside risk for arbitrages with greater execution cost. If, in addition, $h(.)$ is flat close to $\tilde{\theta}$, then an increase in return r will also lead to: 1. more downside risk whenever θ^* is given by Equation (IV.6), 2. less downside risk when $\theta^* = r-1$. Finally, the shape of the density $h(.)$ may affect the optimal choice of downside risk as well as the resulting contract. Specifically, if some probability mass is shifted from high values of ρ toward low values of ρ (representing more volatile environments with greater extreme risks), a choice of $\tilde{\theta}$ that is originally low risk may shift to a high potential risk, as is intuitive.

[14] For simplicity, it is assumed also that when liquidating, the signal y is still produced and thus the optimal incentive contract is unchanged.

[15] Condition (iii) can be easily recovered from the solution to the first-order conditions in Proposition 5.3.

6 Capital Markets

6. 1 Competitive Equilibrium

The previous Section shows how capital inflows will respond to the presence of an arbitrage opportunity r. To better understand how the presence of arbitrageurs will affect asset prices, r is now derived endogenously from equilibrium restrictions. I follow here the initial model by taking θ as an exogenous parameter.

Consider an underlying financial asset, over which the arbitrage is implemented. This asset may be a treasury bond, a stock traded in multiple markets or a security that is a bundle of other securities also traded separately. Assume that there exists a continuously decreasing (resp. increasing) demand (resp. supply) for the asset denoted $D(P)$ (resp. $S(P)$); $D(P)$ and $S(P)$ are assumed to intersect once at P^{na}. To establish the arbitrage, suppose this security is available in another market at a price P_0.[16] Traders buying and selling the security need the security for different motives (e.g., diversification or liquidity) and do not know how to obtain the security at price P_0.[17] To simplify the exposition, $P > P_0$ (the asset is overpriced) since the other case $P < P_0$ is symmetric to this one.

There exists N arbitrageurs who can trade the asset at price P_0. To capture how much initial capital is needed to implement the arbitrage, suppose arbitrageurs need to provide complete initial funding for any security that is held short. Specifically, arbitrageurs will sell a security at price P and buy a security at price P_0. The transaction will generate a net surplus $P - P_0$ for a margin requirement P (the price of the security held short). Thus, the initial required capital will be $P - (P - P_0) = P_0$.[18] Denote $\tilde{S}(P) = S(P) + aN/P_0$, the total supply of the asset including the supply from arbitrageurs, where a represents the capital inflow received by arbitrageurs.

Definition 6.1. *A capital market equilibrium is defined as (P, a) such that:*

(i) Markets clear: $D(P) = \tilde{S}(P)$.

(ii) a is a solution to (P) with $r = (P - P_0)/P_0$.

A capital market equilibrium is defined by incorporating the additional capital provided by arbitrageurs. Next, define a first-best equilibrium by setting $a = a^*$ instead of condition (ii) in the defi-

[16]For simplicity, P_0 is assumed to be determined in a large market which is unaffected by the actions of arbitrageurs.

[17]For the purpose of this model, whether regular traders are also trying to learn about the fundamental value of the asset is irrelevant, since this is assumed to be already captured in the demand and supply curves; the only required assumption for this model is that some gains to trade exists among traders.

[18]Most of the results are unchanged for a smaller margin requirement $mP < P$; however, if P becomes too large and $m < 1$, there may be, in theory, arbitrages that do not require any initial capital since $P - P_0 \geq mP$. This would put an additional upper-bound on $|P - P_0|$ but would not alter the results below this bound.

nition. Denote P^* (resp. P^{**}) the equilibrium price in a first-best (second-best) competitive equilibrium. Assume also that the arbitrage is potentially profitable: $(P^{na} - P_0)/P_0 > \gamma$.

6. 2 Asset Prices

First-Best

Consider first situations in which the arbitrage opportunity can be fully corrected.

Claim 1: In a first-best equilibrium, $P^* = P_0(1 + \gamma)$.

Claim 2: If $\gamma < 1$ and $\gamma \leq 1 + \theta$, in a capital market equilibrium, $P^{**} = P^*$.

In first-best, as shown in Proposition 4.1, any price $P > P^*$ will generate a very large inflow of arbitrage capital. Similarly, when cost and risk are not too large, the second-best solution to the model will produce the same pattern. In these cases, the price inefficiency is corrected (net of cost) and neither arbitrageurs nor investors make a profit.

Small Cost, High Risk ($1 + \theta < \gamma < 1$)

In this case (Proposition 5.1- (ii)), a price P^* would generate no capital inflow and thus cannot be an equilibrium price.

Claim 3: Suppose $1 + \theta < \gamma < 1$. A capital market equilibrium exists if and only if either holds:

(a) $P^{na} \leq P_0(1 + \underline{r})$, then $P^{**} = P^{na}$.

(b) $P^{na} > P_0(1 + \underline{r})$ and:

$$D(P_0(1 + \underline{r})) \geq S(P_0(1 + \underline{r})) + \frac{(1-\gamma)(b - \underline{w})}{\gamma(\gamma - 1 - \theta)} N/P_0 \qquad \text{(IV.7)}$$

then $P^{**} = P_0(1 + \underline{r})$.

In (b), arbitrageurs may realize a rent whenever the equilibrium capital inflow is strictly greater $\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}$. There are also cases such that a capital market equilibrium fails to exist, in particular when the number of arbitrageurs is too large or the cost of hiring the agent is too important. A plausible conjecture is that, in these situations, the market price would shift between $P_0(1 + \underline{r})$ and P^{na} as arbitrageurs enter and exit the market.

Large Cost, Low Risk $(1 + \theta > \gamma > 1)$

Consider now the case of a large cost but a low risk (Proposition 5.2-(i)).

Claim 4: Suppose $1 + \theta > \gamma > 1$.

(a) $P^{**} = P^*$ (no arbitrage opportunities) is a capital market equilibrium price.

(b) If $P^{na} > (1 + \underline{r})P_0$, $P^{**} = P^{na}$ is a capital market equilibrium price.

(c) If, in addition to (b), Equation (IV.7) holds, $P^{**} = (1 + \underline{r})P_0$ is a capital market equilibrium price.

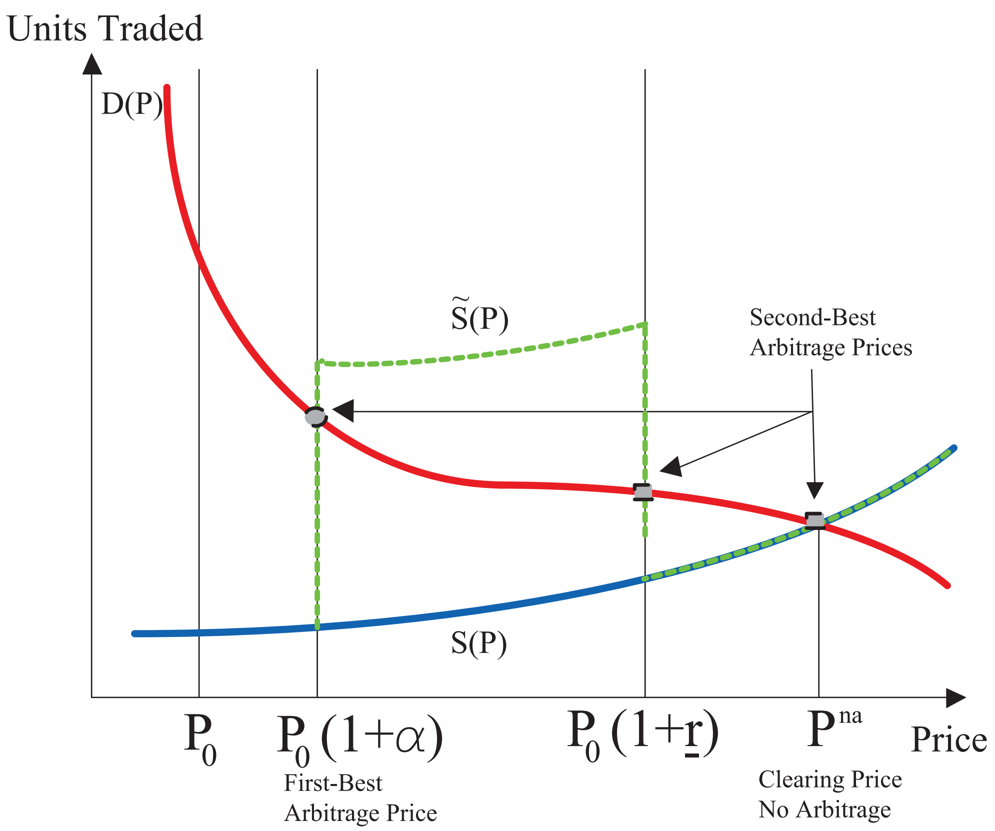


Figure IV.3. Case: $1 + \theta > \gamma > 1$

As shown in Figure IV.3, the market has between one and three equilibria. There is always an equilibrium with no profitable arbitrages ($P^{**} = P^*$). In this equilibrium, second-best and first-best coincide. When the market inefficiency is too large, another equilibrium arises such that arbitrageurs do not trade and the inefficiency is not corrected ($P^{**} = P^{na}$). This result occurs because the price inefficiency is too large for price-taking arbitrageurs to receive any capital. In fact, this equilibrium may seem more "focal" than the previous one since it prescribes using the price prior to the intervention of the arbitrageurs P^{na} as one way to coordinate expectations. Finally, when the number of arbitrageurs is not too large, there will be a third equilibrium with a smaller arbitrage return ($P^{**} = (1+\underline{r})P_0$). This equilibrium will generally be the most preferred by arbitrageurs since they can achieve a rent (the participation does not bind). To summarize, each of these equilibria has plausible features: (a) $P^{**} = P^*$ is (almost) price efficient, (b) $P^{**} = P^{na}$ presupposes using "pre-arbitrage" prices as a coordination device, (c) $P^{**} = (1+\underline{r})P_0$ provides rents to arbitrageurs.

This multiplicity of equilibria also suggests that informational frictions may be a source of price volatility. For example, consider a situation in which a small number of arbitrageurs are able to implement an arbitrage and coordinate their expectations on their most preferred price (c). Now, suppose (say, in the next period) that the demand for the asset by regular traders decreases for exogenous reasons (a common example is a change in index weights or trades caused by repurchases of mutual fund investors). In this case, Equation (IV.7) may fail and thus the price may adjust to one of the two other equilibria (a) or (c). Similarly, a sudden increase in the number of arbitrageurs (see, for example, the recent increase in the number of hedge funds) may shift asset prices away from (c). The price correction may be large despite no change to the fundamentals.

Large Cost, High Risk ($\gamma > 1$ and $1+\theta < 1$)

Finally, if cost and risk are large, arbitrageurs will not be able to receive any capital inflow, regardless of the return of the arbitrage.

Claim 5: If $\gamma > 1$ and $1+\theta < 1$, $P^{**} = P^{na}$ is a capital market equilibrium price.

7 Extension to Decreasing Marginal Utility

7. 1 Capital Flows

The analysis is extended to the case of an arbitrageur with decreasing marginal utility (or, equivalently, risk-aversion). Decreasing marginal utility captures economic reasons why very large capital flows, such as $\overline{a}$, could be expensive to implement and yields a smoother characterization of optimal capital flows.

Assume the arbitrageur receives $u(w) - \gamma a$ ($u' > 0$, $u'' < 0$, $u'(\underline{w}) < \gamma$, $\lim_{w \to +\infty} u'(w) = 0$) to capture decreasing marginal utility to monetary transfers. Suppose in addition that $u^{-1}(b) = 0$ (i.e., the labor cost of hiring an agent with no capital is zero) and, to simplify notations, $u(\underline{w}) = \underline{w}$.[19] In first-best, the participation of the agent binds and thus $W^* = u^{-1}(b + \psi(a^*))$. The optimal capital raised by the investors solves the following problem: $\max_a ra - u^{-1}(b + \psi(a))$. Taking the first-order condition yields the next Proposition.

Proposition 7.1. *In first-best, the optimal capital raised a^* is given as follows:*

$$\gamma = ru'(u^{-1}(\gamma a^* + b)) \tag{IV.8}$$

Let (a^*, W^*) be the first-best outcome; it equates the marginal disutility of investment with the marginal cost of compensating the agent to invest more. Unlike with a risk-neutral agent, the total capital raised is now a smooth function of r. Let $\phi(x) \equiv u((u')^{-1}(x))$; solving for a^*, one may rewrite Equation (IV.8) as follows:

$$a^* = \frac{\phi(\frac{\gamma}{r}) - b}{\gamma} \tag{IV.9}$$

It follows that a^* is increasing in r (i.e., more profitable arbitrages induce greater capital flows).[20] In the rest of this Section, assume that $\phi(1) - b > 0$ so that a positive capital inflow is always chosen in first-best. This benchmark is used to interpret any $a < a^*$ as a compression in capital flows due to informational frictions.

The analogue to Lemma 4.2 is stated in the case of decreasing marginal utility.

Lemma 7.1. $< w, a >$ *is incentive-compatible if and only if:*

$$\frac{u(w) - \underline{w}}{ra - \theta a} \geq \gamma \tag{IV.10}$$

To avoid unnecessary cases to consider, suppose that the outside option of the agent is sufficiently large, i.e. $(r-\theta)/r\gamma \geq u'(u^{-1}(b))$ for all $r > \max(\gamma, 1+\theta)$.[21] Proposition 7.2 characterizes the optimal capital flows in the presence of risk-aversion.

[19] The assumption that $u^{-1}(b) = 0$ is useful to ensure that in first-best some capital is always raised for any $r > \gamma$.

[20] Unlike with risk-aversion, the presence of limited capital flows is no longer surprising with decreasing marginal utility (since it occurs even in first-best). The model, however, yields additional predictions on the relationship between a and the characteristics of the arbitrage.

[21] When this assumption is violated, there will be situations such that only the incentive-compatibility is binding for r large. Except for this additional case to be considered (in case (iii) of Proposition 7.2), the results are unchanged.

Proposition 7.2. *Let $\tilde{r}$ and $\hat{a}$ be defined as follows:*

$$\phi(\gamma/\tilde{r}) = \frac{b-\underline{w}}{\tilde{r}-1-\theta}+b \quad \text{(IV.11)}$$
$$\hat{a} = \frac{b-\underline{w}}{\gamma(r-1-\theta)} \quad \text{(IV.12)}$$

Then:

(i) For $r \in (\gamma, \tilde{r})$, $a = a^*$ is chosen [efficient, no rents].

(ii) For $r \geq \max(\gamma, \tilde{r})$, $a = \hat{a}$ is chosen [distortionary, no rents]

There is a first region of arbitrages such that the first-best capital a^* is chosen (case (i)). For low-profitability arbitrages, the capital raised in first-best is low and thus the wage is also low. Thus, even in second-best, the marginal utility of the agent is high. This makes it "cheap" to prevent speculation. There are of course cases such that region (i) is empty; however, $\tilde{r}$ will always be greater than γ whenever $1+\theta \geq \gamma$, i.e. the risk of the arbitrage is not too important. In case (ii), the incentive-compatibility condition becomes binding, and thus the capital raised by the agent becomes decreasing in r. The intuition for this finding is similar to the previous results.

As the arbitrage becomes very profitable, it will attract almost no capital. Indeed, in the limit, the gross proceeds for investors of very profitable arbitrages ($r\hat{a}$) converges to $\frac{b-\underline{w}}{\gamma}$, finite, as r becomes large. That is, investors are unable to fully benefit from the arbitrage even though the potential first-best surplus (ra^*) becomes unbounded.

7. 2 Asset Prices

To incorporate financial markets, it is now useful to consider separately the cases $P > P_0$ or $P < P_0$ which may have different properties. Formally, using the same notations and margin requirements as in Section 6, an arbitrage will produce a return $r = |P - P_0|/\min(P_0, P)$. To make the analysis as tractable as possible, assume that $1+\theta < \tilde{r} < \gamma$ (i.e, $\tilde{a}$ is always chosen in second-best).

If $P < P_0$ (resp. $P > P_0$), an arbitrageur receiving capital a will be buying (resp. selling) a/P (resp. a/P_0) units of the asset to implement the arbitrage. To capture the total demand and supply of the asset, define next $\tilde{D}(P) = D(P) + 1_{P<P_0} Na(r)/P$ and $\tilde{S}(P) = D(P) + 1_{P>P_0} Na(r)/P_0$, where $a(r)$ is the capital received by arbitrageurs. Replacing $a(r)$ with $\hat{a}$ with $r = |P - P_0|/\min(P_0, P)$ the following expression of the total demand and supply can be obtained.

$$\tilde{D}(P) = D(P) + 1_{P<P_0} \frac{b - \underline{w}}{\gamma(P_0 - (2+\theta)P} \tag{IV.13}$$

$$\tilde{S}(P) = D(P) + 1_{P>P_0} N \frac{b - \underline{w}}{\gamma(P - (2+\theta)P_0} \tag{IV.14}$$

Definition 7.1. *A capital market equilibrium is defined as* (P, a) *such that:*

(i) Markets clear: $\tilde{D}(P) = \tilde{S}(P)$.

(ii) a *is a solution to* (P) *with* $r = |P - P_0| / \min(P, P_0)$.

As before, assume that there exists a profitable arbitrage. Denoting P^{na}, the equilibrium without arbitrage ($D(P^{na}) = S(P^{na})$), $|P^{na} - P_0| / \min(P_0, P^{na}) > \gamma$.

Arbitrage against overpricing ($P > P_0$)

Suppose first that $P - P_0 > \gamma P_0$, i.e. the asset is overpriced. Further, to guarantee the existence of capital market equilibria, assume here the following analogue to Equation (IV.7) (i.e., the number of arbitrageurs is not too large):

$$N \frac{\phi(1) - b}{\gamma} / P_0 + S((1+\gamma)P_0) \leq D((1+\gamma)P_0) \tag{IV.15}$$

Figure IV.4 represents: the original demand and supply curves (bold), the total supply curve (bold-dotted), the total supply curve (long dots). A (first-best) capital market equilibrium is a point were total (first-best) demand and supply curves intersect. Condition (IV.15) ensures that the first-best total supply curve is lower than the original demand at $P = (1+\gamma)P_0$, and thus there must exist at least one capital market equilibrium.

The presence of the arbitrageurs shifts the supply curve upwards, as arbitrageurs now supply the asset. However, because the supply of arbitrageurs is decreasing as one moves further away from P_0 (by Equation (IV.14)), the total supply may become decreasing in prices. As shown earlier, while regular traders supply the asset more with high prices, it is more difficult for arbitrageurs to raise capital and thus they will supply less of the asset.

In this respect, as shown in Figure $IV.14$ there can be multiple capital markets equilibria (square). While all of them prescribe prices lower than the no-arbitrage equilibrium, they can prescribe different deviations from price efficiency. These situations with multiple equilibria are more likely to occur when the supply of assets by regular traders is flat, the demand changes slope fast or the number of arbitrageurs is sufficiently large.[22] In comparison, there is only one first-best capital market equilibrium (black circle) and it prescribes always a greater price efficiency ($|P - P_0|$ small) than second-best.

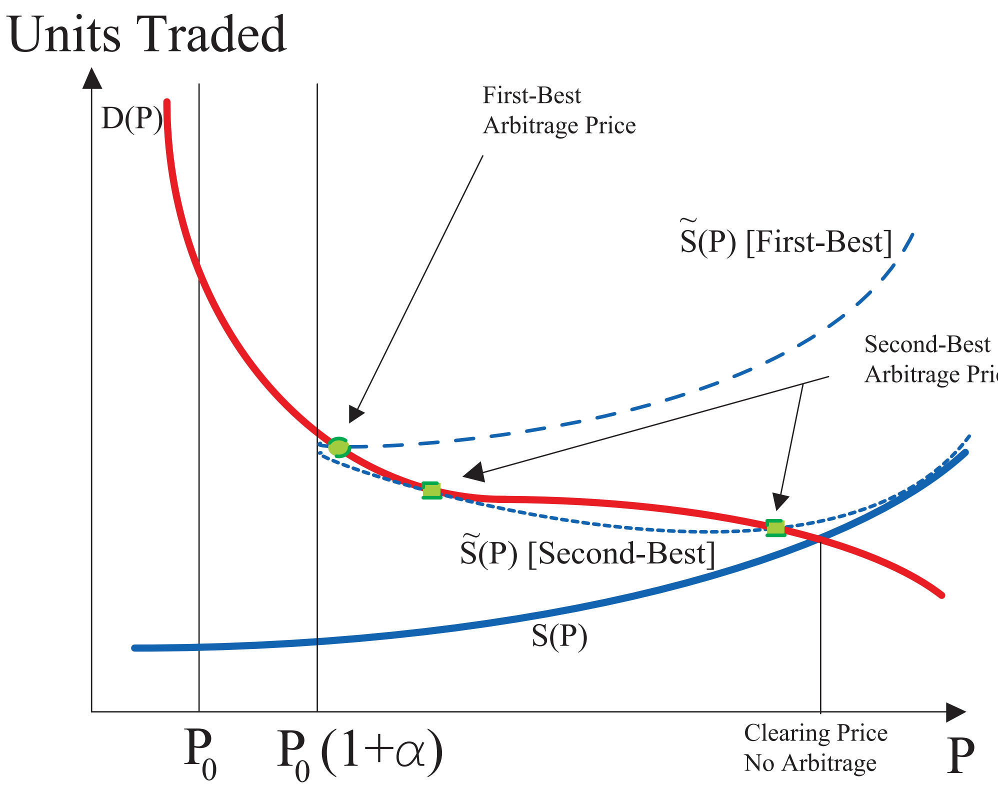


Figure IV.4. Asset Market Equilibrium

An additional implication of the presence of multiple equilibria can be seen by analyzing the response of prices to shocks to the demand of assets. Suppose in Figure IV.4, the demand curve shifts upwards, for example in response to a sudden increase in the attractiveness of the asset. Here, the second-best equilibrium with P closest to P_0 will no longer be an equilibrium. As in the case of risk-neutrality, this equilibrium with higher price efficiency is also the most fragile given a shock to the demand curve.

Arbitrage against underpricing ($P < P_0$)

Consider next the case in which the asset is undervalued, i.e. $P_0 - P < \gamma P$. In this case, arbitrageurs will demand the asset. Further, from Equation (IV.13), this demand will increase as P increase as long as $\theta > -2$, i.e. the risk of the arbitrage is not too large. This will generate features that are qualitatively

[22]To some extent, this may be a plausible description for the internet bubble, since most of the assets held by internet entrepreneurs were locked (flat supply), many banks were actively trading on these stocks and hedging internet stocks using more conventional assets whose business was de facto fairly close (large number of arbitrageurs), demand had been extremely high at lower prices (Ofek and Richardson 2003).

the same as with overpricing.

More interestingly, assume next that the risk of the arbitrage is large, i.e. $\theta \leq -2$. In this case, as P increases toward P_0, the demand of assets by arbitrageurs will decrease. Here, as P becomes small, the need for capital decreases fast which can be extremely helpful when risk is large. As with overpricing, assume that:

$$N \frac{\phi(1) - b}{\gamma} / P + D((1+\gamma)P_0) \leq S((1+\gamma)P_0) \tag{IV.16}$$

Unlike with overpricing, the total demand curve, even including arbitrageurs, will remain decreasing and thus the capital market equilibrium will be unique. Thus, for high-risk arbitrages, multiplicity of equilibria does not occur over underpriced assets. An additional implication of the model is that large market corrections due to multiple equilibria should not occur when the asset is underpriced (although they might when the asset is *over*-priced).

8 Concluding Remarks

This essay presents a theory why some seemingly profitable arbitrages receive either no or limited capital inflows. In particular, capital flows are recovered as a function of primitive characteristics of arbitrages and arbitrageurs. Further research would be helpful to explain how this endogenously derived initial capital may interact with the arbitrageur's ability to stay solvent in a multi-period setting. A plausible conjecture is that it should be desirable for investors to elicit imperfect arbitrages (with a risk of becoming insolvent) as one way to reduce the need for initial capital and/or the maximum potential loss which are both sources of agency cost. Doing so may provide an economic foundation for the observed risk-taking and failures of arbitrageurs.

Another interesting puzzle relates to the presence of (large) positive alphas on hedge funds. In this paper, investors make zero profit in a competitive equilibrium and, thus, the presence of positive alphas is unexplained. There are two potential explanations for this puzzle. First, exposure to systematic risk by arbitrageurs may be hard to measure. Second, and perhaps more interestingly, outside investors in these funds are often large institutions or wealthy individuals; they may be providing active monitoring services whose cost are not captured by these alphas (for example, θ in the model may be the endogenous outcome of this monitoring). The exact practical importance of this monitoring, and its impact on arbitrage fund flows, is still not fully understood.

There are also several open questions with respect to the compensation of hedge fund managers.

In the current setting, predictions are only in terms of the cross-section of observed wages, and not the wage contract given to an arbitrageur (the latter is indeterminate given that only one outcome is observed in equilibrium). Some preliminary attempts in this direction using extensions of this framework (with an additional noisy signal or if the return of the arbitrage is unobservable to investors) suggest that linear or option-like contracts are optimal. There seems to be, however, more structure needed to fully rationalize the high-water mark contracts observed empirically.

Finally, in the last Section, P_0 (the arbitrage price) was presented as an exogenous parameter, although it may itself be sensitive to supply and demand shocks and may be affected by the actions of arbitrageurs. By making P_0 endogenous and driven by market clearing conditions, three problems could be potentially answered. First, are arbitrageurs always beneficial to market efficiency or can their actions create equilibria with profitable arbitrages, even in cases such that an equilibrium with efficient prices exist? Second, how much does capital flow from one market to another in the presence of violations of the law of one price and how do these capital flows affect macroeconomic variables? Third, in a model with shocks to demand and supply, do arbitrageurs increase or reduce asset price volatility and correlations across markets?

9 Appendix: Omitted Proofs

Proof of Lemma 4.1: The method for this proof is to construct a new contract that yields weakly more utility to both contracting parties and elicits a perfect arbitrage. To do so, one should verify that this new contract is desirable to the agent and does not generate deviations from the previous investment.[23]

Let $\hat{F}$ (resp. a) be the distribution (resp. investment) chosen by the agent in response to a contract $w(.)$. Construct the compensation $\hat{w}(ra) = \int w(y)d\hat{F}(y)$ and $\hat{w}(y) = \underline{w}$ for $y \neq ra$. Let $\tilde{F}$ (resp. a) denote the distribution (resp. investment) 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 investment $\tilde{a} = a$ and set $y = ra$. This generates an expected utility $\hat{w}(ra) - \gamma a$.

Claim 2: Under $\hat{w}$, the distribution $\tilde{F}$ must have its support included in $\{\theta a, ra\}$.

[23]The result is robust to risk-aversion by the agent (i.e., utility of the agent is $u(w) - \gamma a$). A more general proof is available on demand.

Suppose not. Define an alternative hedging strategy G as follows: $\int_{y=ra} dG(y) = \int_{\theta a}^{+\infty}(y-\theta a)/(ra-\theta a)d\hat{F}(y)$ and $\int_{y=\theta a} dG(y) = \int_{\theta a}^{+\infty}(ra-y)/(ra-\theta a)d\hat{F}(y)$. It follows that:

$$\begin{aligned}
\int_{\theta a}^{+\infty} ydG(y) &= ra\int_{\theta a}^{+\infty}\frac{y-\theta a}{ra-\theta a}d\tilde{F}(y) + \theta a\int_{\theta a}^{+\infty}\frac{ra-y}{ra-\theta a}d\tilde{F}(y) \\
&= \int_{\theta a}^{+\infty} yd\tilde{F}(y)
\end{aligned}$$

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

$$\begin{aligned}
\int_{\theta a}^{+\infty} \hat{w}(y)dy &= \hat{w}(ra)\int_{y=ra} d\tilde{F}(y) + \underline{w}(1-\int_{y=ra} d\tilde{F}(y)) \\
&< \hat{w}(ra)\int_{y=ra} dG(y) + \underline{w}(1-\int_{y=ra} dG(y)) \qquad \text{(IV.17)}
\end{aligned}$$

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 $\{\theta a, ra\}$. Define the distribution G' as follows: for any $X' \subset X\backslash\{\theta a\}$, $\int_{X'} dG'(y) = \int_{y=ra} d\tilde{F}(y)\int_{X'} d\hat{F}(y)$ and $\int_{y=\theta a} dG'(y) = \int_{y=\theta a} d\tilde{F}(y) + \int_{y=ra} d\tilde{F}(y)\int_{y=\theta a} d\hat{F}(y)$. Suppose that the agent follows G' and $\tilde{a}$ instead of $\hat{F}$ and a with $w(.)$.

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

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

$$\begin{aligned}
\int_{\theta a}^{+\infty} w(y)dG'(y) - \gamma\tilde{a} &= (\int_{y=\theta a} d\tilde{F}(y) + \int_{y=ra} d\tilde{F}(y) \int_{y=\theta a} d\hat{F}(y))w(\theta a) \\
&\quad + \int_{y=ra} d\tilde{F}(y) \int_{y\neq\theta a} w(y)d\hat{F}(y) \\
&\geq \int_{\theta a}^{+\infty} w(y)d\hat{F}(y) - \gamma a + \int_{\theta a}^{+\infty} w(y)d\hat{F}(y) \int_{y=ra} d\tilde{F}(y) - w(ra)\int_{y=ra} d\tilde{F}(y) \\
&\geq \int_{\theta a}^{+\infty} w(y)d\hat{F}(y) - \gamma a + \int_{y=ra} d\tilde{F}(y) \int_{\theta a}^{+\infty} (w(y) \\
&\quad -w(ra))d\hat{F}(y) \\
&\geq \int_{\theta a}^{+\infty} w(y)d\hat{F}(y) - \gamma a
\end{aligned}$$

This is a contradiction to $(a, \hat{F})$ chosen by the agent given wage $w(.)$.

To conclude, note that the cost of compensating the agent is weakly less under $\hat{w}(.)$ than $w(.)$. By Claim 1, the if $w(.)$ satisfies the participation so does $\hat{w}(.)$. By Claim 3, the agent chooses a in response to $\hat{w}(.)$. As a result, investors also achieve weakly more under $\hat{w}(.)$.$\square$

Proof of Proposition 4.2: Assume here that $w(.)$ is an arbitrary compensation which yields a finite utility to the agent. Let $(\hat{F}^n, a^n)_{n=1}^{\infty}$ be a sequence of actions for the agent such that $\int_{-\infty}^{+\infty} yd\hat{F}^n(y) \leq ra^n$ and $\int_{-\infty}^{+\infty} w(y)d\hat{F}^n(y) - \gamma a^n$ converges to $\sup_{a,\hat{F}(.)s.t.\int_{-\infty}^{+\infty} yd\hat{F}(y)\leq ra} \int_{-\infty}^{+\infty} w(y)d\hat{F}(y) - \gamma a < +\infty$. One needs to show that necessarily a_n must converge to 0.

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

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

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

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

In intuitive terms, when choosing G^n, the agent samples between $-n$ and $F^n(.)$.

First, $a = 0$ and G^n is feasible by the agent. To see this,

$$\begin{aligned}\int_{-\infty}^{+\infty} ydG^n(y) &= \frac{ra^n}{n+ra^n}(-n) + (1 - \frac{ra^n}{n+ra^n})\int_{-\infty}^{+\infty} ydF^n(y) \\ &\leq \frac{ra^n}{n+ra^n}(-n) + (1 - \frac{ra^n}{n+ra^n})ra^n \\ &\leq 0\end{aligned}$$

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

$$\begin{aligned}\int_{-\infty}^{+\infty} w(y)dG^n(y) &= \frac{ra^n}{n+ra^n}w(-n-y) + (1 - \frac{ra^n}{n+ra^n})\int_{-\infty}^{+\infty} w(y)d\hat{F}^n(y) \\ &\geq \frac{ra^n}{n+ra^n}\underline{w} + (1 - \frac{ra^n}{n+ra^n})\int_{-\infty}^{+\infty} w(y)d\hat{F}^n(y) \qquad \text{(IV.18)}\end{aligned}$$

Taking the limit on n,

$$\lim_{n\to+\infty}\int_{-\infty}^{+\infty} w(y)dG^n(y) = \lim_{n\to+\infty}\int_{-\infty}^{+\infty} w(y)d\hat{F}^n(y)$$

And therefore, since F^n yields a utility that converges to the supremum of the problem, it must that γa^n converges to zero, i.e. a^n converges to 0.□

Proof of Lemma 4.2: Under the prescribed actions, the agent obtains $w - \gamma a$. 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 ra and θa. Let p denote the probability that $y = \theta a$ and $1 - p$, the probability that $y = ra$. For $\hat{F}(.)$ to be feasible, $p\theta a + (1-p)ra = r(a - \epsilon)$. Therefore: $p = \epsilon r/(ra - \theta a)$. Reinjecting in the problem of the manager: $\frac{r\epsilon}{ra-\theta a}\underline{w} + (1 - \frac{r\epsilon}{ra-\theta a})w - \gamma(ra - \epsilon r)$. This problem is concave. For $\epsilon = 0$ optimal, the first-order condition is given by Equation (IV.5).□

Proof of Proposition 5.1: There are three possible cases: (1) the participation binds, (2) the incentive-compatibility binds, (3) both constraints bind.

(1) Suppose the participation binds, i.e. $w(ra) = b + \gamma a$. The problem of investors will now be the same as in first-best and prescribes: (a) if $r < \gamma$, $a = 0$, (b) if $r = \gamma$, $a \in [0, \overline{a}]$, (c) if $r > \gamma$, $a = \overline{a}$. In (b) and (c), it must be verified that Equation (IV.5) holds at $\overline{a}$. Rewriting Equation (IV.5) with $a = \overline{a}$ yields:

$$b - \underline{w} \geq \gamma\overline{a}(r - \theta - 1) \qquad \text{(IV.19)}$$

This Equation will be satisfied when $r \leq \theta + 1$ but will never be satisfied, by (*), when $r > \theta + 1$.

(2) Suppose the incentive-compatibility binds, i.e. $w(ra) = \gamma a(r - \theta) + \underline{w}$. Investors maximize the following objective: $a(r - \gamma r + \gamma\theta) - \underline{w}$. Then: (a) if $r < \underline{r}$, $a = 0$, (b) If $r = \underline{r}$, $a \in [0, \overline{a}]$, (c) if $r > \underline{r}$, $a = \overline{a}$.

The participation of the agent must now be verified. The utility achieved by the agent will be: $\gamma a(r - \theta - 1) + \underline{w}$. Thus, $r > \theta + 1$ must hold for this inequality to hold. In case (c), this inequality will hold (under (*)). For case (b), $a \in [\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}, \overline{a}]$.

Thus, $\overline{a}$ is chosen for $r > V \equiv \max(\theta + 1, \underline{r})$. Further, one may verify that when $\theta + 1 > \gamma$, $V = \theta + 1$ and when $\theta + 1 \leq \gamma$, $V = \underline{r}$.

(3) Note that if $r < V$, either (1) can be chosen (when $r < 1 + \theta$) or binding the incentive-compatibility constraint cannot yield positive to investors (when $r < \underline{r}$). Thus, binding both constraints cannot be optimal here.

There are two cases to consider: (A) $\gamma \leq 1 + \theta$, (B) $\gamma > 1 + \theta$.

(A) In this case, for all $r \leq 1 + \theta$, there exists a first-best solution (1) and thus it should be chosen. For all $r > 1 + \theta$, there exists a second-best solution such that the participation does not bind and thus it should be chosen also.

(B) In this case, (1) is never possible. For $r \in [\gamma, \underline{r})$, $a = 0$ is optimal. For $r = \underline{r}$, $a \in [\frac{(1-\gamma)(b-\underline{w})}{\gamma(\gamma-1-\theta)}, \overline{a}]$ is optimal. For $r > \underline{r}$, $a = \overline{a}$ is optimal.□

Proof of Proposition 5.2: The steps are similar to Proposition 5.1 so this proof will refer to the same steps when they are unchanged. As before, there are three cases: (1) the participation binds, (2) the incentive-compatibility binds, (3) both constraints bind.

Note first that (1) is not affected by $\gamma > 1$. In (2), investors now maximize: $a(r - \gamma r + \gamma\theta)) - \underline{w}$. This implies that (2) with $a = \overline{a}$ is possible only when $r \in (1 + \theta, \underline{r})$. For $r = \underline{r}$ the result in the previous Section still holds. (3) is unchanged.

To conclude, note that if $\gamma \leq 1 + \theta$, then $1 + \theta < \underline{r}$ and thus (2) never occurs. If $\gamma > 1 + \theta$, then $\gamma > \underline{r}$ and thus (2) never occurs either.□

Proof of Proposition 5.3 Suppose first that $\alpha < 1$. As long as the participation binds, the wage paid to the agent does not depend on θ and thus it is optimal to choose θ as small as possible. Thus, θ

should be chosen at least smaller than $r-1$ (at which point the incentive-compatibility binds). When the incentive-compatibility binds, the principal now solves the following problem: $a(\int_{-\infty}^{\theta}\rho h(\rho)d\rho + r\int_{\theta}^{+\infty}\rho h(\rho)d\rho - \gamma(r-\theta)$. The last term in this objective is the optimal wage given by Equation (IV.5). Taking the first-order condition yields $\tilde{\theta}$. The case with $\alpha > 1$ is the same.□

Proof of Proposition 7.2: There are three possible cases: (1) the participation binds, (2) the incentive-compatibility binds, (3) both constraints bind.

(1) Suppose the participation binds. It is therefore optimal to choose a^*. One may verify next that incentive-compatibility holds, i.e. $\gamma a(r-\theta-1) \leq b - \underline{w}$. This Equation is always met for $r \leq \theta + 1$. For $r > \theta + 1$, $a^* \leq \hat{a}$ where:

$$\hat{a} = \frac{b - \underline{w}}{\gamma(r-1-\theta)} \tag{IV.20}$$

Let $\tilde{r}$ be defined as the unique solution to $a^* = \hat{a}$, that is:

$$\phi(\gamma/\tilde{r}) - b = \frac{b - \underline{w}}{\tilde{r} - 1 - \theta} = \tag{IV.21}$$

(2) Suppose Equation (IV.10) binds. One should verify first that the participation holds, i.e. $a(r-\theta)\gamma + \underline{w} - \gamma a \geq b$. Rewriting this expression, it must be that $r > \theta + 1$ and $a \geq \hat{a}$. Next, one may state the problem of investors after having substituted the wage of the agent: $\max ra - u^{-1}(a(r-\theta)\gamma + \underline{w})$. Taking the first-order condition on this problem and solving for a,

$$a^{**} = \frac{\phi(\frac{r-\theta}{r}\gamma) - \underline{w}}{\gamma(r-\theta)} \tag{IV.22}$$

However, this is always smaller than $\hat{a}$. This is a contradiction. So that, the incentive-compatibility cannot be the only constraint to bind.

(3) Suppose both the incentive-compatibility and the participation bind. Then, $a = \hat{a}$ must be chosen. It must be checked whether $\hat{a}$ yields a strictly positive profit to investors. Calculating the profit of investors with $\hat{a}$ yields: $r\hat{a} - u^{-1}(b + \gamma\hat{a})$.

Several preliminary remarks are helpful in the analysis.

(a) Differentiating $r\hat{a} - u^{-1}(b + \gamma\hat{a})$ with respect to r yields that it is increasing from $r = 1 + \theta$ to $r = r_2$ and decreasing from $r = r_2$ to $r = +\infty$, where $r_2 > 1 + \theta$.

(b) $\lim_{r \downarrow 1+\theta} r\hat{a} - u^{-1}(b + \gamma\hat{a}) = -\infty$.

(c) $\lim_{r \to +\infty} r\hat{a} - u^{-1}(b + \gamma\hat{a}) = \frac{b-w}{\gamma} > 0$.

It is thus established that there exists a unique $\tilde{r}'$ such that $\tilde{r}'\hat{a} - u^{-1}(b + \gamma\hat{a}) = 0$. For $r < \tilde{r}'$, the profit of investors is negative and for $r \geq \tilde{r}'$, the profit of investors is positive. Finally, it is easily verified that $\tilde{r}' < \tilde{r}$ where $\tilde{r}$ is defined as $\hat{a} = a^*$ (if this were not the case, first-best would yield negative profits to investors).

Collecting these results, the following holds:

1. For $r \leq \tilde{r}$, $a = a^*$ is chosen.

2. For $r > \tilde{r}$, $a = \hat{a}$ is chosen.

This concludes the proof.□

Bibliography

Abowd, John M. and David S. Kaplan, "Executive Compensation: Six Questions that Need Answering," *Journal of Economic Perspective*, 1999, *13* (4), 145–168.

Aggarwal, Rajesh K. and Andrew A. Samwick, "Executive Compensation, Strategic Competition, and Relative Performance Evaluation: Theory and Evidence," *Journal of Finance*, 1999, *54* (6), 1999–2043.

Andolfatto, David and Ed Nosal, "Optimal Team Contracts," *Canadian Journal of Economics*, 1997, *30* (2), 385–396.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, "Coherent measures of risk," *Mathematical Finance*, 1999, *9* (3), 203–228.

Arya, Anil, Jonathan Glover, and Shyam Sunder, "Earnings Management and the Revelation Principle," *Review of Accounting Studies*, 1998, *14* (2), 7–34.

— , — , and — , "Are Unmanaged Earnings Always Better for Shareholders?," *Accounting Horizons*, 2003, *17* (1), 111–116.

Bassetto, Marco and Christopher Phelan, "Tax Riots," 2007. Review of Economic Studies (forthcoming).

Battaglini, Marco, "Sequential Voting with Abstention," *Games and Economic Behavior*, 2005, *51* (2), 445–463.

Bebchuk, Lucian and Jesse Fried, *Pay Without Performance: The Unfulfilled Promise of Executive Compensation*, Harvard University Press, 2004.

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

Bertrand, Marianne and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones without Principles are," *Quarterly Journal of Economics*, 2001, *116* (3), 901–932.

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

__ **, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet**, "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, 2007, *74*, 345–390.

Blouin, Jennifer, Luzi Hail, and Michelle Yetman, "Capital Gains Taxes, Pricing Spreads and Arbitrage: Evidence from U.S. Cross-Listed Firms," 2005.

Bollen, Nicholas P. B. and Robert E. Whaley, "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?," *Journal of Finance*, 2004, *59* (2), 711–753.

Brennan, Michael J. and Eduardo S. Schwartz, "Arbitrage in Stock Index Futures," *Journal of Business*, 1990, *63* (1), 7–31.

Brunnermeier, Markus K. and Stefan Nagel, "Hedge Funds and the Technology Bubble," *Journal of Finance*, 2004, *54* (5), 2013–2040.

Burgstahler, David and Ilia Dichev, "Earnings Management to Avoid Earnings Decreases and Losses," *Journal of Accounting and Economics*, 1997, *24* (1), 99–126.

Cherry, Josh, "The Limits of Arbitrage: Evidence from Exchange Traded Funds," 2004.

Citanna, Alessandro and Antonio Villanacci, "Competitive Equilibrium with Moral Hazard in Economies with Multiple Commodities," *Journal of Mathematical Economics*, 2002, *38* (1-2), 117–147.

Coles, Jeffrey L., Daniel D. Naveen, and Lalitha Naveen, "Managerial Incentives and Risk-taking," *Journal of Financial Economics*, 2006, *79* (2), 431–468.

Cornelli, Francesca and David D. Li, "Risk Arbitrage in Takeovers," *Review of Financial Studies*, 2002, *15* (3), 837–868.

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

Cuat, Vicente and Maria Guadalupe, "How Does Product Market Competition Shape Incentive Contracts?," *Journal of the European Economic Association*, 2005, *3* (5), 1058–1082.

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

Dechow, Patricia M. and Douglas J. Skinner, "Earnings Management: Reconciling the Views of Accounting Academics, Practitioners, and Regulators," *Review of Accounting Studies*, 2000, *14* (2), 235–250.

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

Demarzo, Peter M. and Darrell Duffie, "Corporate incentives for hedging and hedge accounting," *Review of Financial Studies*, 1995, *8* (3), 743–771.

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

Duffie, Darrell, *Dynamic Asset Pricing Theory*, Princeton University Press, 2001.

Freeman, Robert N. and Senyo Y. Tse, "A Nonlinear Model of Security Price Responses to Unexpected Earnings," *Journal of Accounting Research*, 1992, *30* (2), 185–209.

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.

Gagnon, Louis and G. Andrew Karolyi, "Multi-Market Trading and Arbitrage," 2003.

Garber, Peter, *Famous First Bubbles: The Fundamentals of Early Manias*, MIT Press, 2000.

Garvey, Gerald and Todd Milbourn, "Incentive Compensation When Executives Can Hedge the Market: Evidence of Relative Performance Evaluation in the Cross Section," *Journal of Finance*, 2003, *58* (4), 1557–1581.

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.

Goetzmann, William N., Jonathan Ingersoll, and Stephen R. Ross, "High-Water Marks and Hedge Fund Management Contracts," *Journal of Finance*, 2003, *58* (4), 1685–1718.

Green, Jerry R. and Nancy L. Stokey, "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 1983, *91* (3), 349–364.

Greenwood, Robin, "Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage," *Journal of Financial Economics*, 2005, *75* (3), 607–649.

Grossman, Sanford J., Oliver D. Hart, and Eric S. Maskin, "Unemployment with Observable Aggregate Shocks," *Journal of Political Economy*, 1983, *91* (6), 907–928.

Guttman, Ilan, Ohad Kadan, and Eugene Kandel, "A Rational Expectations Theory of Kinks in Accounting Earnings," *Accounting Review*, 2006, *81* (4), 811848.

Hart, Oliver and John Moore, "Incomplete Contracts and Renegotiation," *Econometrica*, 1988, *56* (4), 755–785.

Hau, Harald, Massimo Massa, and Joel Peress, "Do Demand Curves for Currencies Slope Down? Evidence from the MSCI Global Index Change," 2005.

Haubrich, Joseph G. and Ivilina Popova, "Executive compensation: a calibration approach," *Economic Theory*, 1998, *12* (3), 561 – 581.

Hayn, C., "The Information Content of Losses," *Journal of Accounting and Economics*, 1995, *20* (2), 125–153.

Healy, Paul M., "The Effect of Bonus Schemes on Accounting Decisions," *Journal of Accounting and Economics*, 1985, *7* (1-3), 85–107.

Heath, Chip, Steven Huddart, and Mark Lang, "Psychological Factors and Stock Option Exercise," *Quarterly Journal of Economics*, 1999, *114* (2), 601–627.

Himmelberg, Charles P. and R. Glenn Hubbard, "Incentive Pay and the Market for CEOs: An Analysis of Pay-For-Performance Sensitivity," 2000. Working Paper.

Hirschleifer, David and Yoon Suh, "Risk, Managerial Effort, and Project Choice" *Journal of Financial Intermediation*, 1992, *3* (2), 308–345.

Hodder, James E. and Jens Carsten Jackwerth, "Incentive Contracts and Hedge Fund Management," 2007. forthcoming Journal of Financial and Quantitative Analysis.

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

— , "Moral Hazard in Teams," *Bell Journal of Economics*, 1982, *13* (2), 324–340.

— **and Jean Tirole**, "Financial Intermediation, Loanable Funds, and Real Sector," *Quaterly Journal of Economics*, 1997, *112*, 663–692.

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

Horst, Jenke Ter and Marno Verbeek, "Fund Liquidation, Self-selection, and Look-ahead Bias in the Hedge Fund Industry," 2007. Review of Finance, forthcoming.

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

Huddart, Steven and Pierre Jinghong Liang, "Accounting in Partnerships," *American Economic Review*, 2003, *93* (2), 410–414.

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

Jagannathan, Ravi, Alexey Malakhov, and Dmitri Novikov, "Do Hot Hands Persist Among Hedge Fund Managers? An Empirical Evaluation," 2006. Working Paper.

Jin, Li, "CEO Compensation, Diversification, and Incentives," *Journal of Financial Economics*, 2002, *66* (1), 29–63.

Jorion, Philippe, *Value at Risk*, McGraw-Hill, 2006.

Kahn, James A., "Moral hazard, imperfect risk-sharing, and the behavior of asset returns," *Journal of Monetary Economics*, 1990, *26* (1), 27–44.

Kaul, Aditya, Vikas Mehrotra, and Randall Morck, "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment," *Journal of Finance*, 2000, *55* (2), 893–912.

Kirstein, Roland and Robert D. Cooter, "Sharing and Anti-Sharing in Teams," 2006.

Kocherlakota, Narayana R., "The effects of moral hazard on asset prices when financial markets are complete," *Journal of Monetary Economics*, 1994, *41* (1), 39–56.

Kouwenberg, Roy and William T. Ziemba, "Incentives and Risk-Taking in Hedge Funds," 2005. Working Paper.

Legros, Patrick and Steven A. Matthews, "Efficient and Nearly-Efficient Partnerships," *Review of Economic Studies*, 1993, *60* (3), 599–611.

Li, Wei, "Heterogeneity of the Discontinuity in Earnings Distributions," 2008. PhD Dissertation.

Liu, Jun and Francis A. Longstaff, "Losing Money on Arbitrage: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities," *Review of Financial Studies*, 2004, *17* (3), 611–641.

Lowenstein, Roger, *Origins of the Crash: The Great Bubble and Its Undoing*, Penguin USA, 2004.

Luporini, Annalisa, "Relative performance evaluation in a multi-plant firm," *Economic Theory*, 2006, *28* (1), 235–243.

M., Harris Jeffrey H. Griffin John and Selim Topaloglu, "The Dynamics of Institutional and Individual Trading," *Journal of Finance*, 2003, *58* (6), 2285–2320.

Magill, Michael and Martine Quinzii, "Common Shocks and Relative Compensation Schemes," 2004.

Margiotta, Mary M. and Robert A. Miller, "Managerial Compensation and the Cost of Moral Hazard," *International Economic Review*, 2000, *41* (3), 669–719.

Matsumoto, Dawn A., "Management's Incentives to Avoid Negative Earnings Surprises," *Accounting Review*, 2002, *77* (3), 483–514.

Meth, Bracha, "Reduction of Outcome Variance: Optimality and Incentives," *Contemporary Accounting Research*, 1996, *13* (1), 309–328.

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

Miller, Nolan H., "Efficiency in Partnerships with Joint Monitoring," *Journal of Economic Theory*, 1997, *77* (2), 285–299.

Mirrlees, James, "The Theory of Moral Hazard with Unobservable Behaviour: Part I," *Review of Economic Studies*, 1999, *66*, 3–22.

Mitchell, Mark, Todd Pulvino, and Erik Stafford, "Limited Arbitrage in Equity Markets," *Journal of Finance*, 2002, *57* (2), 551584.

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

Morris, Steven and Hyun Song Shin, "Liquidity Black Holes," *Review of Finance*, 2004, *8* (1), 1–18.

Mukarram, Attaria and Antonio S. Mellob, "Financially Constrained Arbitrage in Illiquid Markets," *Journal of Economic Dynamics and Control*, 2006, *30* (12), 2793–2822.

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

Nalebuff, Barry J. and Joseph E. Stiglitz, "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, 1983, *14* (1), 21–43.

Nandeibam, Shasikanta, "Sharing Rules in Teams," *Journal of Economic Theory*, 2002, *107* (2), 407–420.

Ofek, Eli and Matthew Richardson, "The Information Content of Losses," *Journal of Finance*, 2003, *58* (3), 1113–1138.

Ou-Yang, Hui, "An Equilibrium Model of Asset Pricing and Moral Hazard," *Review of Financial Studies*, 2005, *18* (4), 1253–1303.

Oyer, Paul, "Why Do Firms Use Incentives That Have No Incentive Effects?," *Journal of Finance*, 2004, *59* (4), 1619–1650.

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.

Panageas, Stavros and Mark M. Westerfield, "High-Water Marks: High Risk Appetites? Convex Compensation, Long Horizons, and Portfolio Choice," 2005. Working Paper.

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

Raith, Michael, "Competition, Risk, and Managerial Incentives," *American Economic Review*, 2003, *93* (4), 1425–1436.

Rajgopal, Shivaram and Terry Shevlin, "Empirical Evidence on the Relation between Stock Option Compensation and Risk Taking," *Journal of Accounting and Economics*, 2002, *33* (2), 145–171.

Rasmusen, Eric, "Moral Hazard in Risk-Averse Teams," *Rand Journal of Economics*, 1987, *18* (3), 428–435.

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

Shavell, Steven, "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics*, 1979, *10* (1), 55–73.

Shiller, Robert J., *Irrational Exuberance*, Princeton University Press, 2005.

Shleifer, Andrei and Robert W. Vishny, "The Limits of Arbitrage," *Journal of Finance*, 1997, *52* (1), 35–55.

Sloan, Richard G. and Douglas J. Skinner, "Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio," *Review of Accounting Studies*, 2004, *7* (2-3), 289–312.

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

Spagnolo, Giancarlo, "Stock-Related Compensation and Product-Market Competition," *Rand Journal of Economics*, 2000, *31* (1), 22–42.

Spear, Stephen E. and Cheng Wang, "When to Fire a CEO: Optimal Termination in Dynamic Contracts," *Journal of Economic Theory*, 2005, *120* (2), 239–256.

Stiglitz, Joseph E. and David Cass, "The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds," *Journal of Economic Theory*, 1970, *2* (2), 122–160.

Strausz, Roland, "Efficiency in Sequential Partnerships," *Journal of Economic Theory*, 1999, *85* (1), 140–156.

Tirole, Jean, *The Theory of Corporate Finance*, Princeton, New Jersey: Princeton University Press, 2006.

Zhou, Xianming and Peter L. Swan, "Performance thresholds in Managerial Incentive Contracts," *Journal of Business*, 2003, *76*, 665–696.