1-1-2012

Essays on Executive Compensation

Qi Liu
University of Pennsylvania, qiliu@wharton.upenn.edu

Essays on Executive Compensation

Abstract
I study executive compensation in various situations, including the cases where (i) CEOs have relative wealth concerns (RWCs); (ii)inside debt can be a part of an optimal contract; (iii) there are ambiguous information about firm value.

The first chapter, "Relative Wealth Concerns and Executive compensation", studies the implications of RWCs on executive compensation. I first study the case in which a CEO's effort increases firm value without changing firm risk. In this case, RWCs will result in an increase in CEO incentives. This effect is larger if aggregate risk is higher, so RWCs can lead to a positive relation between CEO incentives and aggregate risk. CEOs with RWCs willingly risk exposure to aggregate shock to keep up with their peers. This help to reduce risk premium paid to the CEOs. As a result, RWCs can be beneficial to shareholders' payoffs. I also provide a simple explanation for the pay-for-luck puzzle. I next examine the case in which the CEO's effort affects both the mean and variance of firm value. I show that RWCs can lead to a negative relation between CEOs' risk-taking behavior and their incentives, which is consistent with some empirical evidence. Lastly, I show that RWCs render options preferable to stock where aggregate risk is much larger than idiosyncratic risk.

The second chapter, "Inside Debt", is a coauthored paper with my advisor Alex Edmans. We justify the use of debt as efficient compensation. We show that inside debt is a superior solution to the agency costs of debt than the solvency-contingent bonuses proposed by prior literature, since its payoff depends not only on the incidence of bankruptcy but also firm value in bankruptcy. Contrary to intuition, granting the manager equal proportions of debt and equity is typically inefficient. The optimal ratio depends on the trade off of the importance between project selection and effort. The model generates a number of empirical predictions consistent with recent evidence.

In the last chapter, "Incentive Contracting under Ambiguity-Aversion", I study the effect of ambiguity on the pay structure of executive compensation. I show that when there is ambiguity in firm risk and if a manager is risk-averse and ambiguity-averse, stock-based contracts always impose a high risk premium. But option-based contracts can induce the manager to perceive a low risk and thus pay a low risk premium, which makes options less costly than stock. I also show that a manager tends to perceive a higher risk when she is granted higher incentives, which makes shareholders reluctant to grant managers high incentives and take advantage of some small improvements. As a result, compensation contracts exhibit an inertia property. Lastly, if ambiguity comes from the expected market returns, tying CEO pay to the market is optimal, which provides an explanation for the pay-for-luck puzzle.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Finance

First Advisor
Alex Edmans

Keywords
Executive compensation

Subject Categories
Finance and Financial Management

ESSAYS ON EXECUTIVE COMPENSATION

Qi Liu

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2012

Supervisor of Dissertation

____

Alex Edmans, Assistant Professor of Finance

Graduate Group Chairperson

____

Eric Bradlow, Professor of Marketing, Statistics and Education

Dissertation Committee

Alex Edmans (Chair), Assistant Professor of Finance

Franklin Allen, Nippon Life Professor of Finance and Economics

Itay Goldstein, Associate Professor of Finance

Philipp Illeditsch, Assistant Professor of Finance

To my parents.

Acknowledgements

I am extremely grateful to my dissertation advisors: Franklin Allen, Alex Edmans, Itay Goldstein and Philipp Illeditsch for their ideas, guidance and continued support. I would also like to thank Andrew Abel, Philip Bond, Xavier Gabaix, Todd Gormley, Doron Levit, Nikolai Roussanov, Luke Taylor, John Zhu and seminar participants at Delaware, Hong Kong University, SAIF, Peking University and Wharton for helpful comments and suggestions. All remaining errors are my own.
Phone: 217-721-4888, E-mail: qiliu@wharton.upenn.edu, Address: 2419 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA.

ABSTRACT

ESSAYS ON EXECUTIVE COMPENSATION

Qi Liu

Alex Edmans

I study executive compensation in various situations, including the cases where (i) CEOs have relative wealth concerns (RWCs); (ii) inside debt can be a part of an optimal contract; (iii) there are ambiguous information about firm value.

The first chapter, "Relative Wealth Concerns and Executive compensation", studies the implications of RWCs on executive compensation. I first study the case in which a CEO's effort increases firm value without changing firm risk. In this case, RWCs will result in an increase in CEO incentives. This effect is larger if aggregate risk is higher, so RWCs can lead to a positive relation between CEO incentives and aggregate risk. CEOs with RWCs willingly risk exposure to aggregate shock to keep up with their peers. This help to reduce risk premium paid to the CEOs. As a result, RWCs can be beneficial to shareholders' payoffs. I also provide a simple explanation for the pay-for-luck puzzle. I next examine the case in which the CEO's effort affects both the mean and variance of firm value. I show that RWCs can lead to a negative relation between CEOs' risk-taking behavior and their incentives, which is consistent with some empirical evidence. Lastly, I show that RWCs render options preferable to stock where aggregate risk is much larger than idiosyncratic risk.

The second chapter, "Inside Debt", is a coauthored paper with my advisor Alex Edmans. We justify the use of debt as efficient compensation. We show that inside debt is a superior solution to the agency costs of debt than the solvency-contingent bonuses proposed by prior literature, since its payoff depends not only on the incidence of bankruptcy but also firm value in bankruptcy. Contrary to intuition, granting

the manager equal proportions of debt and equity is typically inefficient. The optimal ratio depends on the trade-off of the importance between project selection and effort. The model generates a number of empirical predictions consistent with recent evidence.

In the last chapter, "Incentive Contracting under Ambiguity-Aversion", I study the effect of ambiguity on the pay structure of executive compensation. I show that when there is ambiguity in firm risk and if a manager is risk-averse and ambiguity-averse, stock-based contracts always impose a high risk premium. But option-based contracts can induce the manager to perceive a low risk and thus pay a low risk premium, which makes options less costly than stock. I also show that a manager tends to perceive a higher risk when she is granted higher incentives, which makes shareholders reluctant to grant managers high incentives and take advantage of some small improvements. As a result, compensation contracts exhibit an inertia property. Lastly, if ambiguity comes from the expected market returns, tying CEO pay to the market is optimal, which provides an explanation for the pay-for-luck puzzle.

Table of Contents

List of Figures

Chapter 1

Relative Wealth Concerns and Executive Compensation

1.1 Introduction

While standard economic models usually assume that an agent's utility is derived from the absolute level of her own consumption, economists have long believed that relative consumption effects are important (i.e., one's evaluation of her consumption depends on how much the other people are consuming, Veblen (1899)). A growing body of literature incorporates such effects into asset pricing models (Abel (1990), Constantinides (1990), Galí (1994), and Campbell and Cochrane (1999)). Such relative consumption effects, however, are not extensively studied in corporate finance. In this paper, I incorporate relative wealth concerns (RWCs) into a standard principal-agent model and study how CEO compensation is affected by such relative consumption effects.

Some empirical evidence already suggests that RWCs matter in CEO compensation. Bouwman (2011), for example, finds that CEO compensation depends on the compensation of geographically-close CEOs. An examination of the relation between geography and executive compensation shows that the results are most consistent with RWCs. Shue (2011) documents the phenomenon of "pay for friend's luck": pay responds to lucky industry-level shocks to the compensation of peers in distant industries. Therefore, peer effects in compensation are not driven by similarities in underlying managerial productivity; instead, RWCs might be the driving forces. Bereskin and Cicero (2011) study developments in Delaware caselaw in the mid-1990s and find that firms not directly impacted by legal changes increased their CEOs' compensation when the legal changes directly affected a substantial number of firms in their industry. RWCs could be a potential explanation.[1] If RWCs play a role in CEO compensation, then it is not surprising to see that the level of CEO pay is affected by peer compensation, e.g., Bouwman (2011), Bizjak, Lemmon, and Naveen (2008), Faulkender and Yang (2010) find that the level of CEO pay increases in the average level of CEO pay in a reference group.[2] Although the correlation among the level of CEO pay has been examined, the interaction among CEO incentives has not been studied empirically. In this paper, I will focus on the RWCs' effects on CEO incentives. I derive the closed-form solutions in the model. The model's tractability enables me to obtain a number of clear economic effects on CEO incentives and explain their intuitions transparently.

First, I study the case in which a CEO's effort increases firm value without changing firm risk. I set up the model with a continuum of firms and CEOs. I assume that a CEO's utility function exhibits a consumption externality. In particular, I define that a CEO's RWCs are given by the product of a parameter measuring her concerns and the sum of all the CEOs' compensation (see Definition 1.2.1 in the model). A CEO's utility is not determined by the absolute level of her compensation, but depends on the difference between her compensation and her RWCs. The participation constraint will require a premium for the CEO's RWCs, i.e., the CEO desires to have higher compensation when her peers' compensation is high. Therefore, a CEO's contract must achieve two goals: one is to induce the CEO to work; the other is to make the CEO satisfied with her contract when she compares herself to her peers. The contract has to pay a premium for the CEO's RWCs, which I refer to as a "RWCs' premium".

By granting CEOs equity to induce them to work, CEO pay is positively correlated with aggregate shock (i.e., the state of the whole economy); thus, their RWCs also depend on the aggregate shock. Because CEOs are risk-averse, it is less costly to pay the RWCs' premium by using a type of payment linked to the aggregate shock rather than cash; stock, then, is preferred to cash. So RWCs lead to an increase in

[1] There are also some other explanations; in the paper, Bereskin and Cicero point to market competition as an explanation for their results.

[2] Note that the positive correlation in the level of CEO pay might be caused by many factors (e.g., competition, living costs, a common shock to managerial productivity, RWCs). Among these three papers, only Bouwman (2011) shows that her results are most consistent with RWCs.

CEO incentives. Since a CEO's RWCs also depend on how many incentives the other CEOs are granted, the model predicts that individual CEO incentives increase with an increase in aggregate CEO incentives. If aggregate risk is higher, then a CEO will face a higher risk to fall behind other CEOs; thus, stock compensation becomes more attractive than cash. Therefore, RWCs can result in a positive relation between CEO incentives and aggregate risk. On the other hand, according to the informativeness principle, firms should use fewer incentives when firm risk (which is the sum of aggregate risk and idiosyncratic risk) is higher, because the observable outcome (i.e., stock price) becomes less informative. Due to the two conflicting effects, the relation between CEO incentives and aggregate risk can be negative or positive, depending on the coefficient of a CEO's RWCs on aggregate shock. Empirical studies have found mixed evidence as to the relation between CEO incentives and firm risk. Many of them document no significant relation, or even a positive relation between incentives and firm risk (Prendergast (2002)), which is inconsistent with a standard principal-agent model. My model predicts that if the coefficient of a CEO's RWCs on aggregate shock is large, then there is a positive relation between incentives and aggregate risk; otherwise, the relation remains negative.

The optimal CEO incentives depend on individual RWCs, which depend on other CEOs' incentives. In equilibrium, each firm will choose its contract optimally given other firms' contracts. I show that because each individual CEO incentives depend on the aggregate CEO incentives linearly, this linearity leads to a unique equilibrium. In addition, notice that in the presence of RWCs, a CEO is willing to risk exposure to aggregate shock to keep up with her peers to some extent: the CEO becomes less risk-averse due to her RWCs. Thus RWCs help to reduce the risk premium paid to the CEO. In other words, the presence of RWCs allows shareholders to grant a CEO more incentives to induce a higher effort, but without paying a higher risk premium. In particular, if a CEO's productivity is high, choosing a CEO with RWCs is beneficial to shareholders' payoffs.

In reality, however, shareholders may not be able to observe a CEO's RWCs, I also study the case where the CEO's RWCs are not observable. In this case with asymmetric information, a CEO can extract an information rent, i.e., the CEO can receive a contract which makes her utility above the reservation utility. I show that for each firm, a pooling on CEO types – giving the same contract to CEOs of different types – may exist. [3] This pooling result could help to explain the practice of benchmark pay (Bizjak, Lemmon and Naveen (2008), Faulkender and Yang (2010)). For firms with certain similarities, they may fall into the same region of pooling, so they will use other firms' contracts as a benchmark. Since CEOs' incentives depend on each other, and the CEO's type at the reservation utility also depends on other CEOs' compensation, I find that multiple equilibria can arise with asymmetric information.

Lastly, I study the case where relative performance evaluation can be used. I show that if a CEO is also concerned about her friends or neighbors – entrepreneurs, investment bankers, and private equity investors for example – whose income is correlated with some lucky shock (market shock for example), then firms should compensate the CEO by a means that will correlate with that shock. Thus the model provides a simple explanation for the pay-for-luck puzzle documented by Bertrand and Mullainathan (2001).[4]

I then turn to study the case in which a CEO's effort affects not only the mean but also the variance of firm value. To improve firm value, a CEO may need to undertake risky projects; these decisions also increase the variance of firm value. By combining risk-taking efforts and RWCs, the model generates some new predictions on executive compensation. Because a CEO's effort also increases the variance of firm value and the CEO is risk-averse, higher incentives do not always lead to a higher effort. As I demonstrate that an incentive-threshold exists such that the CEO's effort decreases in her incentives when her incentives exceeds the threshold. If the CEO has no RWCs, at the optimum, the optimal CEO incentives will never exceed the threshold, because doing this will only reduce the CEO's effort and increase the risk premium paid to the CEO. Thus her effort is increasing in her incentives. However, if the CEO has RWCs, shareholders may have to grant the CEO incentives above the threshold to pay the RWCs' premium. Thus, if the CEO's RWCs are intensive, her effort can be decreasing in her incentives at the optimum. Because of this non-monotonicity between efforts and incentives, and because CEO incentives depend on peer incentives, I show that multiple equilibria could exist, such that in one equilibrium, CEOs' efforts increase in their incentives; in another equilibrium, CEOs' efforts decrease in their incentives.

So far I have confined the discussions to linear contracts for tractability.[5] In practice, however, options

[3]This pooling result seems quite unique to RWCs in the sense that in the model I also tried some other cases, e.g., a CEO's productivity or a CEO's risk-aversion is not observable, then I only get separating on CEO types.

[4]Gopalan, Milbourn and Song (2010), Noe and Rebello (2008), and Oyer (2004) provide alternative explanations.

[5]Many papers specify the form of contracts for tractability (e.g., Holmstrom and Tirole (1993), Jin (2002), Oyer (2004), and

are widely used, I extended the model to account for non-linear contracts. In particular, I compare stock to options in the presence of RWCs in a simple example. I establish that RWCs have two effects when comparing stock to options. First, the effect of RWCs on incentivizing CEO effort is stronger for options than stock. This is because if the CEO chooses to shirk, stock still remains exposed to aggregate shock; in contrast, options are more likely to expire with value zero making them completely insensitive to the shock. But, as previously explained, in the presence of RWCs, CEOs willingly risk exposure to aggregate shock because it helps them to keep up with their peers. Thus the effect of RWCs on incentivizing a CEO's effort is larger for an option-based contract. Second, because RWCs make CEOs willing to risk exposure to aggregate shock, they can help to reduce the risk premium associated with aggregate risk. Notice that stock is fully exposed to risk, but options are only partially exposed to risk; thus, RWCs reduce the risk premium more for stock than options. I show that when aggregate risk is much larger than idiosyncratic risk, the first effect dominates the second one such that options become more preferable. Where idiosyncratic risk is not too small compared with aggregate risk, however, stock will be more efficient.

This paper is related to several strands of literature. First, it is related to the literature that examines RWCs' effects. This has been extensively studied in asset pricing models. For example, relative consumption effects are used to examine equity premium puzzle (Abel (1990), Constantinides (1990), Galí (1994), Campbell and Cochrane (1999)). García and Strobl (2011) study the complementarities in information acquisition using RWCs. DeMarzo, Kaniel and Kremer (2008) show how RWCs can result in financial bubbles. Roussanov (2010) explains the under-diversification puzzle by emphasizing the desire to "get ahead of the Joneses". In corporate finance, Goel and Thakor (2010) explore merger waves using envy-based preferences. The concerns for "equity"are applied to study the dynamics of workers' wages (Cabrales, Calvó-Armengol and Pavoni (2007)). This paper introduces RWCs into a standard principal-agent model and examines how CEO incentives will be affected. Miglietta (2010) finds that RWCs make agents prefer positively correlated payoffs. Using a case involving one principal and two agents, Miglietta (2008) explains the use of firm-wide incentive contracts to employees who do not seem to need any incentive. In this paper, I focus on the interactions between shareholders and a CEO within a firm, as well as interactions among firms. I derive the optimal contract for each firm and also show that RWCs may sometimes

Bolton, Scheinkman and Xiong (2006)).

be beneficial to shareholders' payoffs. The case in which a CEO's effort affects both the mean and variance of firm value is also discussed.

Second, my paper is related to the literature on executive compensation. By incorporating RWCs into the model, I show that many features in CEO compensation can be rationalized. The standard principal-agent model predicts a negative relation between incentives and risk. However, many empirical studies find no significant relation, or even a positive relation between incentives and firm risk (Prendergast (2002)). RWCs can explain this contradiction: the relation between incentives and aggregate risk can be either negative or positive, depending on the coefficient of a CEO's RWCs on aggregate shock. Bertrand and Mullainathan (2001) document that CEOs are paid for market upswings that are beyond their control, a so-called "pay-for-luck"puzzle. Gopalan, Milbourn and Song (2010) argue that if a CEO can exert an effort to explore future market conditions and decide the firm's exposure to the market risk, then it is necessary to tie CEO pay to luck. In the model, I provide a simple explanation by showing that if a CEO compares her compensation with that of friends or neighbors, especially those whose income is correlated with lucky shock, then it is optimal to compensate the CEO by a means that is also correlated with the shock. Many empirical studies (Bizjak, Brickley and Coles (1993), Coles, Daniel and Naveen (2006), Aggarwal and Samwick (2006), Gormley, Matsa and Milbourn (2011)) have documented a negative relation between CEOs' risk-taking behavior and their incentives. I show that if a CEO's RWCs are intensive, then shareholders have to pay the RWCs' premium using more shares of stock such that a negative relation between the CEO's risk-taking behavior and her incentives arises at the optimum. Dittmann and Maug (2007) calibrate a standard principal-agent model and find that options should not be used. In contrast, I provide a rationale for the use of options by demonstrating that the incentives provided by RWCs to induce efforts are stronger for options than stock, and this effect will be dominant if aggregate risk is much larger than idiosyncratic risk.

1.2 Efforts Only Affect the Mean of Firm Value

1.2.1 Model

Suppose there is a continuum of firms and CEOs: $i \in [0,1]$. CEO i is assigned to firm i.[6] For each firm i, consider the following principal-agent model: an all-equity firm is owned by risk-neutral shareholders and managed by a risk-averse CEO. At $t = 0$, the CEO is offered a contract. At $t = 1$, she exerts an effort a_i, which is unobservable. The firm's terminal value at time $t = 2$ is given by

$$V_i = \pi_i a_i + \tilde{m} + \eta_i, \tag{1.1}$$

where $\tilde{m}$ is the aggregate shock with mean 0 and variance σ_m^2; η_i is firm i's idiosyncratic shock with mean 0 and variance $\sigma_{\eta i}^2$; $\tilde{m}$ and $(\eta_i)_{i\in[0,1]}$ are independent of each other. I assume that $(\sigma_{\eta i}^2)_{i\in[0,1]}$ are uniformly bounded, i.e., there exists $\sigma_{\max}$ such that $\sigma_{\eta i} \le \sigma_{\max}$ for each $i \in [0,1]$. $\pi_i > 0$ is the measure of CEO i's productivity.

For tractability, I confine the current discussion to linear contracts which consist of a base salary α_i[7] and $\beta_i \le 1$ shares of stock. Then CEO i's payoff at time $t = 2$ is

$$w_i = \alpha_i + \beta_i V_i = \alpha_i + \beta_i(\pi_i a_i + \tilde{m} + \eta_i).$$

At time $t = 1$, CEO i exerts an effort a_i to maximize her expected utility. I introduce RWCs into a CEO's utility function and assume that a CEO's utility is not determined by the absolute level of her wage, but depends on the difference between her wage and a portion of other CEOs' wage. Specifically, I assume that CEO i's utility is given by

$$U_i = E[u_i(a_i), H_i] = E[w_i - H_i] - \frac{1}{2}\lambda_i Var[w_i - H_i] - \frac{1}{2}a_i^2,^{8} \tag{1.2}$$

[6]Here, I do not study the problem of matching firms and CEOs. Gabaix and Landier (2008) and Edmans and Gabaix (2011) present competitive assignment models of the managerial labor market.

[7]I do not put restrictions on α_i. In Proposition 1.2.2, we will see that α_i depends on the CEO's reservation utility $\underline{u}_i$. So α_i can be negative if $\underline{u}_i$ is low.

[8]In fact, this is the certainty-equivalent of exponential utility function $E[-e^{-\lambda_i(w_i - H_i - \frac{1}{2}a_i^2)}]$ if all the shocks are normally distributed.

where λ_i is the measure of CEO i's risk-aversion; $\frac{1}{2}a_i^2$ is the cost of her effort. The novel part in the utility function is the incorporation of RWCs H_i. I define the CEO's RWCs as follows:

Definition 1.2.1. *A CEO's RWCs are the product of a parameter measuring the CEO's concerns about other CEOs' pay and the sum of all the CEO pay. Specifically, for CEO i, her RWCs are denoted by*

$$H_i = r_i \int_0^1 w_k \, dk,$$

where $r_i \geq 0$ is a parameter measuring CEO i's concerns about other CEOs' pay. A larger r_i means that CEO i cares more about other CEOs' compensation. $\int_0^1 w_k \, dk$ is the sum of all the CEO pay.

The RWCs H_i can be simplified to

$$H_i = r_i(W + M\tilde{m}),^{9} \tag{1.3}$$

where $W = \int_0^1 (\alpha_k + \beta_k \pi_k a_k) dk = r_i \int_0^1 E[w_k] dk$ is the sum of the expected CEO pay, and $M = \int_0^1 \beta_k \, dk = \int_0^1 \frac{\partial w_k}{\partial \tilde{m}} dk$ is the sum of the sensitivity of CEO pay to aggregate shock (i.e., aggregate CEO incentives in this case).

Shareholders will choose a contract to maximize the expected firm value net of CEO pay. For brevity, I drop the subscript i. For each firm, we have the following principal-agent problem:

Program 1.

$$\max_{\alpha,\beta} E[V - w] \geq 0^{10}$$

such that

$$\text{IC: } a \in \arg\max_{\hat{a}} E[u(\hat{a}), H],$$

$$\text{IR: } E[u(a), H] \geq \underline{u},$$

where $\underline{u}$ is the CEO's reservation utility.

Proposition 1.2.2.

[9] See Appendix for the derivation.

[10] If the shareholders' expected payoff is negative, I assume that shareholders will not hire the CEO and all the agents get zero payoff in expectation. See Section 1.2.2 for discussions.

1) Given a linear contract $w = \alpha + \beta V$, the optimal effort taken by the CEO is $a = \pi\beta$, thus $\frac{\partial a}{\partial \beta} > 0$, i.e., the CEO's effort is always increasing in incentives.

2) If $rM\sigma_m^2 \geq \sigma_m^2 + \sigma_\eta^2$, then $\beta^* = 1$. In this case, we have a corner solution, i.e., it is better for shareholders to sell the firm to the CEO.

If $rM\sigma_m^2 < \sigma_m^2 + \sigma_\eta^2$, then we have an interior solution. In this case, the optimal contract (α^*, β^*) is

$$\begin{cases} \alpha^* = \frac{(\pi^2 - \lambda r M\sigma_m^2)(\pi^2 + \lambda r M\sigma_m^2)}{2[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]} - \frac{\pi^2(\pi^2 + \lambda r M\sigma_m^2)^2}{[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]^2} + rW + \frac{1}{2}\lambda r^2 M^2 \sigma_m^2 + \underline{u}, \\ \beta^* = \frac{\pi^2 + \lambda r M\sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}. \end{cases}$$

2i) β^* is increasing in π, r and M; decreasing in λ and σ_η^2.

2ii) The relation between β^* and σ_m^2 is ambiguous. If $r\left[\frac{\partial M}{\partial \sigma_m^2}\sigma_m^2 + \left(1 - \frac{\lambda\sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}\right)M\right] > \frac{\pi^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$, then β^* is increasing in σ_m^2; if $r\left[\frac{\partial M}{\partial \sigma_m^2}\sigma_m^2 + \left(1 - \frac{\lambda\sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}\right)M\right] < \frac{\pi^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$, then β^* is decreasing in σ_m^2.

2iii) The expected CEO pay is

$$E[w] = \frac{\pi^4}{2[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]} + rW + \frac{1}{2}\lambda r^2 M^2 \sigma_m^2 \left[1 - \frac{\lambda\sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}\right] + \underline{u},$$

which is increasing in r, W and M.

Now we discuss the intuitions behind Proposition 1.2.2. In part 1), since a CEO's effort only increases the mean of firm value, but has no effect on the variance, giving a CEO higher incentives will always induce her to exert a higher effort. Note that a CEO's effort decision is determined by her productivity of effort and her cost of effort; at the optimum, the marginal benefit of effort for the CEO equals to the marginal cost of effort. Also note that when the target CEO's incentives are given, the other CEOs' contracts and efforts do not affect either the productivity or the cost of effort of the target CEO, so the target CEO's effort is not affected by other CEOs' contracts or efforts, even if she cares about other CEOs' compensation.[11]

In part 2i), the relations between β^* and π, λ, σ_η^2 are standard: a higher π (which is the measure of a CEO's productivity) means that the benefit of a CEO's effort is greater, thus shareholders would like to

[11]In Section 1.3.1, I will show that RWCs provide additional incentives for the CEO to work when her effort also affects the variance of firm value.

grant a CEO more incentives to induce a higher effort when π increases. Since the CEO is risk-averse, a rise in λ or σ_η^2 makes it more costly to induce a CEO's effort, so β^* is decreasing in λ and σ_η^2. The novel results of part 2i) are the relations between β^* and r and M. In the presence of RWCs, a CEO's contract plays two roles: one is to induce the CEO to work; the other is to make the CEO satisfied with her contract when she compares it to those of her peers, i.e., the contract has to pay a "RWCs' premium" to the CEO. Since CEOs are risk-averse, note that the RWCs are positively correlated with aggregate shock, thus it is less costly for shareholders to make a payment that is related to the shock, rather than pay cash to the CEO. In other words, stock compensation is better. Thus as r or M goes up, CEO incentives will rise as well. Recall that M is the sum of the sensitivity of CEO pay to aggregate shock (aggregate CEO incentives, in this case), hence the model predicts that individual CEO incentives increase when aggregate CEO incentives increase. Similarly, since participation constraint requires a premium for RWCs, part 2iii) shows that the expected CEO pay is increasing in W (aggregate expected CEO pay) and M (aggregate CEO incentives). Bouwman (2011), Bizjak, Lemmon, and Naveen (2008), and Faulkender and Yang (2010) find that the level of CEO pay increases in the average level of CEO pay in a reference group,[12] which is consistent with the model. The model also predicts a positive relation between individual CEO incentives and aggregate CEO incentives, an interesting result to be tested in the future.

Part 2ii) shows that the relation between incentives and aggregate risk depends on a CEO's RWCs. In the benchmark case with no RWCs ($r = 0$), as aggregate risk goes up, it becomes more difficult to distinguish a CEO's effort from observable outcomes. According to the informativeness principle, a firm should use fewer incentives, thus the optimal incentives are decreasing in aggregate risk. This is consistent with the standard principal-agent model. When a CEO has RWCs ($r > 0$), however, there is another effect that affects the relation between incentives and aggregate risk. As explained above, it is less costly to pay the RWCs' premium by using stock compensation than using cash. Thus, when aggregate risk goes up, stock compensation becomes more attractive than cash compensation. Shareholders are more willing to use stock to pay the RWCs' premium as aggregate risk increases. Therefore, when r is large, RWCs' effects are dominant, which leads to a positive relation between incentives and aggregate risk. This result can be used to explain the mixed empirical evidence on the relation between CEO incentives and firm risk.

[12]Bouwman (2011) use geographically-close firms as a reference group; Bizjak, Lemmon, and Naveen (2008), Faulkender and Yang (2010) use similar-sized firms in the same industry as a reference group.

The standard principal-agent model predicts that there should be a negative relationship between CEO incentives and firm risk. Yet, some empirical studies show that there is no significant relationship, or even a positive relationship between incentives and firm risk (Prendergast (2002)). By incorporating a CEO's RWCs into the model, although the relation between CEO incentives and a firm's idiosyncratic risk is still negative, the relation between CEO incentives and aggregate risk can be positive. [13] In particular, the model shows that if the aggregate CEO incentives are high, then there is a positive relation between individual CEO incentives and aggregate risk, which is a new testable prediction.

Lastly, part 2iii) predicts that the expected CEO pay increases with aggregate expected CEO pay and aggregate CEO incentives. This result is intuitive: the participation constraint requires a premium for a CEO's RWCs, which depend on other CEOs' income. If aggregate expected CEO pay goes up, individual expected CEO pay goes up as well. The increase of aggregate CEO incentives makes each individual CEO face a higher risk to fall behind other CEOs' payoffs. This has to be compensated by increasing CEO pay.

1.2.2 Equilibrium

When a CEO has RWCs, the optimal CEO incentives will depend on other CEOs' incentives by Proposition 1.2.2. Thus, I define an equilibrium as follows:

Definition 1.2.3. *A subgame perfect equilibrium is a set of functions* $(\alpha_i, \beta_i, a_i)_{i\in[0,1]}$ *that satisfy the following properties:*

1) at time 0, for each firm i*, given all the other firms' contracts* $(\alpha_j, \beta_j)_{j\neq i}$*, the contract* (α_i, β_i) *solves the principal-agent problem optimally for firm* i*.*

2) at time 1, for the CEO in firm i*, given her contract and all the other firms' contracts,* a_i *maximizes CEO* i*'s utility.*

If a CEO's RWCs are too big, i.e., if r is large, then it becomes too costly to compensate the CEO. As a result, shareholders will decide not to hire the CEO, in which case all the agents get zero payoff. In the following lemma, I study the case in which each firm will hire its CEO. I provide a necessary and sufficient condition under which a unique (hiring) equilibrium exists.

[13] Jin (2002) obtains the same result under the assumption that a CEO is risk-averse and shareholders are risk-neutral to a firm's idiosyncratic risk but risk-averse to aggregate risk.

Lemma 1.2.4. *There is a unique equilibrium in which each firm will hire its CEO (and the optimal contract for each firm is given by Proposition 1.2.2), if and only if*

$$\int_0^1 r_i\,di < 1, \text{ and} \tag{1.4}$$

$$r_i M\sigma_m^2 \leq \sigma_m^2 + \sigma_{\eta i}^2 \text{ for each } i \in [0,1], \text{ and} \tag{1.5}$$

$$\frac{\pi_i^2(\pi_i^2 + \lambda_i r_i M\sigma_m^2)}{\pi_i^2 + \lambda_i(\sigma_m^2 + \sigma_{\eta i}^2)} - E[w_i] \geq 0 \text{ for each } i \in [0,1], \tag{1.6}$$

where $M = \dfrac{\int_0^1 \frac{\pi_i^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\,di}{1-\int_0^1 \frac{\lambda_i r_i \sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\,di}$ *is aggregate CEO incentives;*

$W = \dfrac{\int_0^1 \left[\frac{\pi_i^4}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]} + \frac{1}{2}\lambda_i r_i^2 M^2 \sigma_m^2 \left[1-\frac{\lambda_i \sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\right] + \underline{u}_i\right] di}{1-\int_0^1 r_i\,di}$ *is aggregate expected CEO pay;*

$E[w_i] = \frac{\pi_i^4}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]} + r_i W + \frac{1}{2}\lambda_i r_i^2 M^2 \sigma_m^2 \left[1 - \frac{\lambda_i \sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\right] + \underline{u}_i$ *is the expected CEO pay for firm* i.

Note that (1.5) guarantees that the optimal incentives for each firm do not exceed 1. (1.4) and (1.6) guarantee that the shareholders in each firm have nonnegative payoffs. From Proposition 1.2.2, it can be seen that a CEO's effort decision is determined only by her own contract, even if she cares about other CEOs' compensation. Also, a CEO's effort is always increasing in her own incentives linearly. As a result, although RWCs make a CEO's optimal incentives depend on the other CEOs' incentives, this dependence only results in a linear relation between individual CEO incentives and aggregate CEO incentives. By aggregating all the CEOs' incentives, we will eventually obtain unique aggregate CEO incentives, which in turn determine unique optimal CEO incentives for each firm. The equilibrium is therefore unique.

1.2.3 Are RWCs Good or Bad for Shareholders?

In this section, I explore the following question: if shareholders can observe the parameter r and choose a CEO from a pool of CEOs with a different r, then will they choose a CEO with $r = 0$ or a positive r? In other words, do RWCs always hurt shareholders' value? When a CEO has RWCs, it has two effects on shareholders' payoffs. On one hand, shareholders have to pay a premium for a CEO's RWCs due to the participation constraint; this will reduce shareholders' payoffs. On the other hand, shareholders' payoffs can also benefit from a CEO's RWCs for the following reason: in the presence of RWCs, a CEO willingly risks exposure to aggregate shock to keep up with her peers. Thus RWCs help to reduce the risk premium paid to the CEO. The risk premium paid to a CEO for the aggregate risk that she bears is not proportional to her owned shares, but is proportional to the difference between her owned shares and the coefficient of her RWCs on aggregate shock (which is $\beta - rM$). Hence, when r is positive, it allows shareholders to grant a CEO more incentives to induce a higher effort, but without paying a higher risk premium. In particular, if a CEO's productivity is high, the second effect could dominate the first effect so that a positive r is optimal for shareholders.

Lemma 1.2.5. *If*

$$\frac{\lambda M \sigma_m^2 \pi^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)} > W, \tag{1.7}$$

then the optimal r^* is positive, which is given by $r^* = \frac{\lambda M \sigma_m^2 \pi^2 - W[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]}{\lambda M^2 \sigma_m^2 (\pi^2 + \lambda \sigma_\eta^2)}$. The optimal r^* is increasing in π; decreasing in W and σ_η^2.

As we argued above, the presence of RWCs has two effects on shareholders' payoffs: first, it reduces the shareholders' payoffs, as the participation constraint requires a premium for a CEO's RWCs; second, it may also increase the shareholders' payoffs, since it allows shareholders to grant a CEO more incentives to induce a higher effort without paying a higher risk premium. If π (the productivity of a CEO's effort) is large, then it strengthens the second effect; therefore, the optimal r^* is increasing in π. If W (aggregate expected CEO pay) is large, then the first effect is stronger because it now requires a higher premium for CEOs' RWCs; thus, the optimal r^* is decreasing in W. The increase of a firm's idiosyncratic risk σ_η^2 reduces the optimal CEO incentives, which reduces the second effect; so the optimal r^* is decreasing in σ_η^2.

1.2.4 Optimal CEO Incentives When CEOs' RWCs Are Unobservable

In reality, a CEO's RWCs may not be known by shareholders. If a CEO's RWCs are not observable, then how will they affect the optimal CEO incentives? Suppose that $r \in [0, r_{\max}]$ is unobservable to shareholders. The shareholders' beliefs on r are represented by the probability density function $f(r)$ with the associated cumulative distribution function $F(r)$. I assume that the distribution function satisfies the relevant monotone hazard rate conditions: $\frac{F(r)}{f(r)}$ is increasing in r and $\frac{1-F(r)}{f(r)}$ is decreasing in r.[14]

The timeline of the model is: at $t=0$, shareholders offer a menu of contracts to a CEO; at $t=1$, the CEO chooses a contract and exerts an effort; at $t=2$, all the agents' payoffs are realized. Note that when a CEO is choosing a contract and exerting an effort, both he and the shareholders do not know the other CEOs' types. As a result, the CEO and shareholders both face the uncertainty about the other CEOs' types when they are evaluating the contract and their utilities. In particular, the values of W (aggregate expected CEO pay) and M (aggregate CEO incentives) will affect a CEO's relative payoff. Given the distributions of r_i for $i \in [0,1]$, each realization of $(r_i)_{i\in[0,1]}$ determines each CEO's choice of her contract, which determines the values of W and M. Thus the distributions of W and M depend on the distributions of $(r_i)_{i\in[0,1]}$. Assume that the distributions of $(r_i)_{i\in[0,1]}$ are independent of the distributions of aggregate shock $\tilde{m}$ and idiosyncratic shocks $(\eta_i)_{i\in[0,1]}$, thus the distributions of W and M are independent of the distributions of aggregate shock $\tilde{m}$ and idiosyncratic shocks $(\eta_i)_{i\in[0,1]}$. Let $F_W(x)$ and $F_M(x)$ denote the cumulative distribution functions for W and M, respectively.[15]

To distinguish CEOs of different types, shareholders offer a menu of contracts to a CEO. Each contract consists of a base salary and some shares of stock. Specifically, in response to a CEO's claimed RWCs $\hat{r}$, the contract specifies a base salary $\alpha(\hat{r})$ and $\beta(\hat{r})$ shares of stock. Let $U(\hat{r}|r)$ denote the CEO's expected utility, given that the CEO's type is r but he claims to be of type $\hat{r}$. Then the shareholders' objective is:

Program 2.

$$\max_{\alpha(r),\beta(r)} \int_0^{r_{\max}} E[V-w|r]f(r)dr$$

such that

$$\textit{IC: } U(r|r) \geq U(\hat{r}|r), \textit{ for any } \hat{r}, r$$

$$\textit{IR: } U(r|r) \geq \underline{u}, \textit{ for any } r,.$$

Lemma 1.2.6. *Given a menu of contracts $\{(\alpha(r),\beta(r)\}_{r\in[0,r_{\max}]}$, and suppose that $\beta(r)$ is differentiable, then the contracts satisfy the IC constraint, if and only if*

[14] Many distribution functions satisfy this property, e.g., uniform, normal, logistic, exponential, and Laplace.

[15] Because there is a continuum of firms, it can be shown that by Kolmogorov's strong law of large numbers, M converges to $E[M]$ and W converges to $E[W]$ almost surely.

1)

$$\frac{\partial \beta(r)}{\partial r} \geq 0;$$

2)

$$\frac{\partial \alpha(r)}{\partial r} = -\left\{[\pi^2 - \lambda(\sigma_m^2 + \sigma_\eta^2)]\beta(r) + \lambda r E[M]\sigma_m^2\right\} \frac{\partial \beta(r)}{\partial r}.$$

The coefficient of a CEO's RWCs on aggregate shock is rM. So to catch up with the other CEOs' exposure to aggregate shock, a CEO with a higher r is more willing to choose a contract with higher incentives. Thus to distinguish CEOs of different types, it is necessary to grant higher incentives to the CEO with a higher r.

Due to the asymmetric information between shareholders and the CEO, it is possible for the CEO to receive a contract that gives her a utility above the reservation utility.[16] I refer to $U(r|r) - \underline{u}$ as *information rent*. The optimal incentives will be a trade-off between the cost of information rent and the efficiency of incentives. A CEO with a higher r is more willing to risk exposure to aggregate shock because it helps her to keep up with her peers. Thus a CEO with a higher r will gain more from a high-incentives contract. In other words, $\Delta U = U(r+\delta r|r+\delta r) - U(r|r)$, the difference in utility between two near types, is increasing in incentives. Thus, if the CEO of type r_0 is at the reservation utility with her contract, then for $r < r_0$, information rent is decreasing in incentives $\beta(r)$; for $r > r_0$, information rent is increasing in incentives $\beta(r)$. Therefore, to lower the information rent, shareholders would prefer to grant high incentives to the CEOs of types below r_0, and low incentives to the CEOs of types above r_0. Further, because the IC constraint requires that incentives must be increasing in a CEO's type, there could be a pooling on CEOs' types around r_0; that is, the contracts are the same for the CEOs of types around r_0. This result helps to explain why benchmark pay is so widely used in practice (Bizjak, Lemmon and Naveen (2008) and Faulkender and Yang (2010)). For firms with certain similarities, they may fall into the same region of pooling, so they will use other firms' contracts as a benchmark.

Proposition 1.2.7. *Suppose $r_0 = \arg\min_{r\in[0,r_{\max}]} U(r|r)$ is the type such that the CEO of type r_0 is at the reservation utility with her contract, if* $\max_{r\in[0,r_{\max}]} \sigma_m^2 E[M]\left[r + \frac{F(r)}{f(r)}\right] \leq \sigma_m^2 + \sigma_\eta^2$ *(this is a sufficient condition to ensure an interior solution), then*

[16] Note that in the case where the CEO's RWCs are observable, the base salary is always set to make her utility at the reservation utility.

1) if $r_0 = 0$*, then the optimal incentives are* $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r - \frac{1-F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$;

2) if $r_0 = r_{\max}$*, then the optimal incentives are* $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r + \frac{F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$;

3) if $0 < r_0 < r_{\max}$*, then there must exist* $r_1 \le r_0 \le r_2$ *such that* $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r + \frac{F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$ *for* $0 \le r \le r_1$*,* $\beta^*(r)$ *is constant for* $r_1 \le r \le r_2$*, and* $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r - \frac{1-F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$ *for* $r \ge r_2$.

Most of comparative statics are the same with Proposition 1.2.2: $\beta^*(r)$ *is increasing in* π*,* r*; decreasing in* λ *and* σ_η^2*; the relation between* $\beta^*(r)$ *and* σ_m^2 *is ambiguous. But the relation between* $\beta^*(r)$ *and the aggregate CEO incentives* $E[M]$ *becomes negative when* $r > r_0$ *and* $r - \frac{1-F(r)}{f(r)} < 0$.

The reason why the incentives for the types above r_0 can be decreasing in aggregate CEO incentives is due to the cost of information rent. As argued above, when aggregate CEO incentives go up, it is less costly to use stock compensation than cash to pay a CEO's RWCs' premium. But where information asymmetry is present, granting a CEO more shares of stock also increases the information rent extracted by the CEOs of types above r_0. As a result, the optimal individual incentives may be decreasing in aggregate CEO incentives, if the cost of information rent is more important.

When r is unobservable, the optimal incentives for a CEO of type r depend on the relative position between r and r_0. As argued above, when $r < r_0$, it is necessary to keep the CEO incentives *as high as possible* to reduce the information rent; when $r > r_0$, it is necessary to keep the CEO incentives *as low as possible* to reduce the information rent. Meanwhile, the location of r_0 depends on M (aggregate CEO incentives), i.e., r_0 depends on other CEOs' incentives. Hence, multiple equilibria can arise in the case with asymmetric information, such that r_0 is low in some equilibria and high in some other equilibria. In the following example, I show that there could exist multiple equilibria.

Example 1. *Suppose all the firms are identical, i.e.* $\lambda_i = \lambda$*,* $\pi_i = \pi$*,* $\sigma_{\eta i}^2 = \sigma_\eta^2$*,* r_i *are independently and uniformly distributed over* $[0, r_{\max}]$*. Let*
$M_1 = \frac{\pi^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$, $W_1 = \frac{1}{2}[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]\left[1 - \frac{1}{3}\left(\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}\right)^2\right] M_1^2 + \underline{u}$, $M_2 = \frac{\pi^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2) - \lambda\sigma_m^2 r_{\max}}$,
$W_2 = \frac{1}{1 - r_{\max}}\left[\frac{1}{2}[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]\left[1 - \frac{\lambda\sigma_m^2 r_{\max}(2 - r_{\max})}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)} - \frac{1}{3}\left(\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}\right)^2\right] M_2^2 + \underline{u}\right]$.

If

$$\lambda M_1^2\sigma_m^2\cdot\min\left(1-\frac{1}{2}r_{\max},1-\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)-W_1\geq 0,$$
$$\lambda M_2^2\sigma_m^2\cdot\max\left(1-\frac{1}{2}r_{\max},1-r_{\max}+\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)-W_2\leq 0,$$

then there exist at least two equilibria such that

1) in one equilibrium, $r_0=0$. The optimal incentives for each firm are low and given by $\beta^*(r)=\frac{\pi^2+\lambda\sigma_m^2 M_1(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$. The aggregate incentives are M_1, and the aggregate expected CEO pay is W_1.

2) in another equilibrium, $r_0=r_{\max}$. The optimal incentives for each firm are high and given by $\beta^*(r)=\frac{\pi^2+2r\lambda\sigma_m^2 M_2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$. The aggregate incentives are M_2, and the aggregate expected CEO pay is W_2.

1.2.5 A Simple Explanation of Pay-for-Luck

Empirical studies document that CEOs are often rewarded for factors that are beyond their control, i.e., paid for luck (Bertrand and Mullainathan (2001)). This phenomenon gives rise to criticism that executive compensation is inefficient and is in fact decided by executives themselves. In this section, I use RWCs to make a simple explanation of this puzzle. The intuition is as follows: suppose that firm value is affected by some lucky shock beyond a CEO's control. A CEO with RWCs will require a RWCs' premium due to the participation constraint. Because the CEO is risk-averse, if the CEO's RWCs depend on the lucky shock, then it is less costly to compensate the CEO by a means that correlates to the shock, rather than to pay cash to the CEO. In other words, shareholders should not subtract the lucky shock from the firm's terminal value when deciding CEO pay.

In (1.1), I refer to aggregate shock $\tilde{m}$ as the lucky shock, and assume that it is observable and contractible. Then the shareholders can filter out the lucky shock from CEO pay and pay the CEO for the shock separately. Hence for the CEO in firm i, her contract now has the following form:

$$w_i=\alpha_i+\beta_i(V_i-\tilde{m})+\gamma_i\tilde{m}=\alpha_i+\beta_i(\pi_i a_i+\eta_i)+\gamma_i\tilde{m}$$

A CEO may not only care about other CEOs' pay, but also be concerned about her friends or neighbors –

entrepreneurs, investment bankers, and private equity investors, etc – whose income is correlated with the lucky shock $\tilde{m}$. So I now assume that CEO i's RWCs are given by

$$H_i = r_i \int_0^1 w_k \, dk + b_i \tilde{m},$$

where the first term still reflects CEO i's concerns about other CEOs' compensation; the second term $b_i \geq 0$ reflects that the CEO will also compare her compensation to that of her friends, whose income is positively correlated with the lucky shock $\tilde{m}$.

Proposition 1.2.8. *Suppose that $\int_0^1 r_k \, dk < 1$. Then the optimal contract $(\alpha_i^*, \beta_i^*, \gamma_i^*)$ for firm i is given by*

$$\begin{cases} \alpha_i^* = \frac{(\lambda\sigma_{\eta i}^2 - \pi_i^2)h_i^4}{2(\pi_i^2 + \lambda_i \sigma_{\eta i}^2)^2} + rW + \underline{u}, \\ \beta_i^* = \frac{\pi_i^2}{\pi_i^2 + \lambda_i \sigma_{\eta i}^2}, \\ \gamma_i^* = \frac{r_i \int_0^1 b_k \, dk}{1 - \int_0^1 r_k \, dk} + b_i, \end{cases}$$

where W is the sum of the expected CEO pay. So as long as $\int_0^1 b_k \, dk > 0$, $\gamma_i^* > 0$ for each $i \in [0,1]$, i.e., it is optimal to pay CEOs for luck.

1.3 Efforts Affect the Mean and Variance of Firm Value

1.3.1 Model

In the last section, a CEO can improve firm value without changing firm risk, which is exogenous at $\sigma^2 = \sigma_m^2 + \sigma_\eta^2$. In reality, however, a CEO's action may not only affect the mean of firm value, but also the variance of firm value. For example, if a CEO decides to invest in risky projects, investing in more projects will expose the firm to more risk, because the success of a project is subject to the state of the whole economy and the firm's own idiosyncratic risk. Previous literature has argued that one objective of CEO compensation is to induce a CEO's risk-taking actions that increase both the mean and variance of firm value. For example, risk-taking incentives are generally used to argue that options can be more efficient than stock (Dittmann and Yu (2008) and Feltham and Wu (2001)).

In this section, I investigate the case where firm return depends on firm risk σ^2 chosen by the CEO. As in Sung (1995) and Dittmann and Yu (2009), I represent the CEO's investment opportunity set by a frontier in (σ, a) space. A point on the frontier represents the maximum return that the CEO can achieve by accepting the risk σ. Let $a(\sigma)$ denote the maximum return associated with the target risk σ. To increase return above $a(\sigma)$, the CEO will have to take more positive-NPV projects. This can be operationalized by allowing the CEO to choose a single action a which affects both the mean and variance of the return. Thus, I assume that at $t = 1$, if a CEO exerts effort a, then the firm's terminal value at time $t = 2$ is given by

$$V = \pi a + a(\tilde{m} + \eta), \tag{1.8}$$

Compared to (1.1), in (1.8) the CEO's effort also affects the volatility of firm value, which is proportional to the effort a. Except that the form of firm value is different from Section 1.2.1, all the other assumptions and specifications remain unchanged. In particular, I still confine the discussion to linear contracts for tractability.[17]

Now since CEOs' actions also affect their firms' exposure to risk and all CEOs' payoffs are subject to aggregate shock, for any target CEO, if the other CEOs' payoffs have a high exposure to aggregate shock, it will also incentivize the target CEO to exert a higher effort for the purpose of keeping up with the other CEOs' exposure to the aggregate shock. Therefore, other CEOs' incentives and actions can also affect the target CEO's effort decision. This is different from the basic model in Section 1.2.1, in which a CEO's effort decision is only affected by her own contract – even if she is concerned about the other CEOs' payoffs.

Lemma 1.3.1. *Given a linear contract $w = \alpha + \beta V$, the optimal effort taken by a CEO is*

$$a = \frac{\beta\pi + \lambda\beta\sigma_m^2 rM}{\lambda(\sigma_m^2 + \sigma_\eta^2)\beta^2 + 1}.$$

1) The CEO's effort a is increasing in β if $\beta < \sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$ and decreasing in β if $\beta > \sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$.

2) a is increasing in π, r and M, decreasing in σ_η^2; but the relation between a and σ_m^2 is ambiguous.

[17] Edmans and Gabaix (2011) and Ou-Yang (2003) explore situations where a CEO's effort affects the volatility of firm value in a general contracting problem (i.e., they do not restrict to linear contracts). But they have some other assumptions to get the closed form solution. Edmans and Gabaix (2011) assume that a CEO can observe the noise before taking the action to ensure the optimality of log-linear contracts; Ou-Yang (2003) uses a continuous time framework as in Holmstrom and Milgrom (1987).

If $\frac{r}{\pi+\lambda\sigma_m^2 rM}\frac{\partial(\sigma_m^2 M)}{\sigma_m^2} > \frac{\beta^2}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1}$*, then a is increasing in* σ_m^2*; if* $\frac{r}{\pi+\lambda\sigma_m^2 rM}\frac{\partial(\sigma_m^2 M)}{\sigma_m^2} < \frac{\beta^2}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1}$*, then a is decreasing in* σ_m^2*.*

Since a higher effort leads to a larger variance of firm value, a risk-averse CEO will be reluctant to exert a high effort when she is granted high incentives. Thus the CEO's effort is increasing in her incentives only when her incentives are not too high. This is different from the case in Section 1.2.1 where efforts are always increasing in incentives. Given the incentives β, a CEO's effort has two effects on her payoff: first, it increases the mean of firm value. The marginal benefit of this effect is $\pi\beta$, and the marginal cost of this effect is a. Second, the effort also increases the variance of firm value. Since the CEO is risk-averse, the marginal cost of increasing firm risk is $\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2 a$. Due to the CEO's RWCs, the increase of firm risk also helps the CEO to catch up with the other CEOs' exposure to the aggregate shock. The marginal benefit of this effect is $\lambda\beta\sigma_m^2 rM$. So the total marginal benefit of the CEO's effort is $\beta\pi+\lambda\beta\sigma_m^2 rM$, and the total marginal cost of the CEO's effort is $a[\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1]$. Equating the benefit to the cost yields the optimal effort in Lemma 1.3.1. Granting the CEO a few more shares will increase the marginal benefit of effort by $\pi+\lambda\sigma_m^2 rM$, and increase the marginal cost of effort by $2a\lambda(\sigma_m^2+\sigma_\eta^2)\beta$. If the increase in the marginal benefit is larger (smaller) than the increase in the marginal cost, then granting the CEO more shares will increase (decrease) the CEO's effort. Comparing the change in the marginal benefit and marginal cost, it can be seen that there is a threshold $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$, such that when CEO incentives exceed this threshold, the CEO's effort will become decreasing in the incentives.

I now turn to the results in part 2) of Lemma 1.3.1. Since the increase of π, r, and M will increase the marginal benefit of the CEO's effort, but has no effect on the marginal cost, the CEO's effort a is increasing in π, r and M. Similarly, the increase of σ_η^2 only increases the risk that the CEO has to bear, so it only increases the marginal cost of the CEO's effort. Thus the CEO's effort a is decreasing in σ_η^2. Lastly, the increase of σ_m^2 increases both the marginal benefit and marginal cost. On one hand, the increase of σ_m^2 decreases the CEO's effort because the CEO is risk-averse and a higher CEO's effort exposes the CEO's payoff to more risk. On the other hand, because other CEOs' payoffs are correlated with aggregate shock, a higher effort can help the CEO to catch up with the other CEOs' exposure to aggregate shock. The benefit from this effect is more valuable when the variance of aggregate risk goes up; thus the increase of σ_m^2 also increases the CEO's effort due to her concerns about the other CEOs' payoffs. So the two conflicting effects make the effect of σ_m^2 on the CEO's effort ambiguous.

The shareholders will choose an optimal level of effort and a linear contract to implement the effort. Solving the model, we can obtain that

Proposition 1.3.2. *If $rM\sigma_m^2 \geq \pi(\sigma_m^2+\sigma_\eta^2)$, then $\beta^*=1$, i.e., it is better to sell the firm to the CEO. If $rM\sigma_m^2 < \pi(\sigma_m^2+\sigma_\eta^2)$, then the optimal linear contract $w=\alpha^*+\beta^* V$ is given by*

$$\beta^* = \frac{2}{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)+\sqrt{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)}},$$

where $M=\int_0^1 \frac{\partial w_k}{\partial \tilde{m}}dk$ is the sum of the sensitivity of CEO pay to aggregate shock. The base salary α^* is set to make the participation constraint binding.

1) β^* is increasing in r and M; decreasing in π, λ, and σ_η^2;

2) The relation between β^* and σ_m^2 is ambiguous. When $rM+r\sigma_m^2\frac{\partial M}{\partial \sigma_m^2}<\pi\beta^*$, β^* is decreasing in σ_m^2; when $rM+r\sigma_m^2\frac{\partial M}{\partial \sigma_m^2}>\pi\beta^*$, β^* is increasing in σ_m^2.

3) If $\lambda r M\sigma_m^2<\pi$, then $\beta^*<\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$, which implies that $\frac{\partial a}{\partial \beta}|_{\beta^*}>0$, i.e., granting the CEO more incentives will increase the CEO's effort. If $\lambda r M\sigma_m^2>\pi$, then $\beta^*>\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$, which implies that $\frac{\partial a}{\partial \beta}|_{\beta^*}<0$, i.e., granting the CEO more incentives will decrease the CEO's effort.

Compared to Proposition 1.2.2, there are two different results. In Proposition 1.3.2, the relation between β^* and π is reversed. Now β^* becomes decreasing in π. The reason is as follows: given CEO incentives, a CEO will exert an optimal effort a that is proportional to her productivity π. Thus the expected firm value will be increased by πa, which is proportional to π^2 (because a is proportional to π). But note that the CEO's effort also increases the variance of firm value, which is proportional to a^2. Because a is proportional to π, the risk premium paid to the CEO is also proportional to π^2. Hence from Proposition 1.3.2, we can see that when there are no RWCs (i.e. $r=0$), the optimal incentives do not depend on π because the benefit and the cost are both proportional to π^2. If the CEO has RWCs, they provide two benefits on shareholders' payoffs: Lemma 1.3.1 shows that the RWCs also provide incentives for the CEO to work. Such a benefit is proportional to the CEO's productivity π. Secondly, since the risk premium paid to the CEO for the aggregate risk that she bears is now equal to $\frac{1}{2}\lambda(\beta a-rM)^2\sigma_m^2$. If r is positive, it helps to reduce the risk premium. Savings on the risk premium are also proportional to π. Thus when π goes up, the benefits for the shareholders from the CEO's RWCs become relatively smaller when compared with the benefit from the CEO's effort and the cost of CEO compensation. Therefore, the optimal incentives will move toward to the ones in the case with no RWCs: the optimal incentives will decrease as π goes up.

Another different result is part 3) of Proposition 1.3.2. In Proposition 1.2.2, the CEO's effort only affects the mean, so her effort will always increase in the incentives. Here, the CEO's effort also affects the variance of firm value, so when the CEO's incentives are too high, it reduces the CEO's (risk-taking) effort. In part 3), if the CEO has no RWCs, at the optimum, β^* is less than the threshold $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$: the CEO's effort is increasing in the incentives. Intuitively, it is never optimal for shareholders to grant a CEO more than $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$ shares of stock, since doing so will only reduce the CEO's effort and increase the risk premium paid to the CEO. However, if the CEO has RWCs ($r > 0$), it is less costly to pay the RWCs' premium by using stock than cash. When r is large, it could be optimal for shareholders to grant a CEO more than $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$ shares of stock to pay the RWCs' premium. Thus in equilibrium, granting the CEO more shares can decrease her (risk-taking) effort.[18]

As r increases, Lemma 1.3.1 shows that it directly increases CEO effort. On the other hand, the increase of r also increases the optimal CEO incentives by Proposition 1.3.2. If the CEO's RWCs are large, then at the optimum the CEO's risk-taking effort may be decreasing in her incentives. Thus the increase of CEO incentives may reduce a CEO's effort. Therefore the increase of r may have an indirect effect through incentives to decrease a CEO's effort. I show in the following lemma that the total effects of a CEO's RWCs on her (risk-taking) effort are positive, i.e., the optimal CEO's effort is increasing in her RWCs. This suggests that when a CEO's RWCs go up, she will exert a higher effort. As a result, the covariance of firm value with aggregate shock goes up. Similarly, if the sensitivity of other CEOs' pay to aggregate shock goes up, the total effects on the target CEO's risk-taking action are also positive, which will push up the covariance of the target firm value with the aggregate shock. The results are summarized in the following lemma.

Lemma 1.3.3.

$$\frac{da}{dr}|_{\beta^*} = \frac{\partial a}{\partial \beta}\frac{\partial \beta}{\partial r} + \frac{\partial a}{\partial r} > 0.$$

Therefore, $\frac{d}{dr}Cov(V, \tilde{m}) > 0$. *As a CEO's RWCs increase, firm value becomes more covariant with the aggre-*

[18] Carpenter (2000) and Ross (2004) show that granting a CEO options may reduce the CEO's risk-taking incentives, but they do not endogenize the optimal CEO incentives in the model. Here I show that by endogenizing the optimal CEO incentives, granting her more incentives could reduce her risk-taking efforts only when the CEO has RWCs.

gate shock.

Similarly, $\frac{da}{dM}|_{\beta^*} > 0$ *implies that as the sum of the sensitivity of CEO pay to aggregate shock goes up, firm value becomes more covariant with the aggregate shock.*

1.3.2 Equilibrium

When a CEO cares about other CEOs' compensation, Proposition 1.3.2 shows that the optimal incentives for one firm depend on other CEOs' incentives. Lemma 1.3.1 shows that each CEO's (risk-taking) effort also depends on the other CEOs' incentives and efforts. So an equilibrium requires that each firm chooses a contract optimally given the other firms' contracts, and each CEO exerts an optimal effort given the other CEOs' incentives and efforts.

Definition 1.3.4. *A subgame perfect equilibrium is a set of functions* $(\alpha_i, \beta_i, a_i)_{i\in[0,1]}$ *that satisfy the following properties:*

1) at time 0, for each firm i*, given all the other firms' contracts* $(\alpha_j, \beta_j)_{j\neq i}$*, the contract* (α_i, β_i) *solves the principal-agent problem optimally for firm* i*.*

2) at time 1, for the CEO in firm i*, given the other firms' contracts and CEOs' efforts,* a_i *maximizes the CEO* i*'s utility.*

If r is too big, there could be two results. First, it can become too costly to compensate the CEO. As a result, shareholders will decide not to hire the CEO. Second, as CEOs' effort decisions are affected by each other, when some CEOs (a positive measure) increase their efforts, it will also incentivize the other CEOs to exert higher efforts. This, in turn, will increase the original CEO effort, and so on and so forth. Then RWCs eventually will have an amplifying effect on incentivizing a CEO's effort. Thus, if r is too big, the result could be that CEOs will exert an effort of infinite large, which is unrealistic. In the following lemma, I provide a sufficient condition under which equilibrium exists; in equilibrium, CEOs are hired and their efforts are bounded.

Lemma 1.3.5. *Define that* $\beta_i(M) = \frac{2}{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)+\sqrt{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)^2+4\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}}$ *for each* $i \in [0,1]$ *and* $W(M) =$

$\frac{\int_0^1 \left[\frac{1}{2}\frac{\beta_i(M)^2(\pi_i-\lambda_i r_i M\sigma_m^2)(\pi_i+\lambda_i r_i M\sigma_m^2)}{\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)\beta_i(M)^2+1}+\frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2\right]di}{1-\int_0^1 r_i di}$. *Suppose there exists a solution $M > 0$ to the equation*

$$\int_0^1 \frac{\beta_i(M)^2(\pi_i+\lambda_i\sigma_m^2 r_i M)}{\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)\beta_i(M)^2+1}di = M.$$

If the following conditions are satisfied, there must exist at least one equilibrium in which CEOs are hired with the optimal contracts given by Proposition 1.3.2 and their efforts are bounded.

$$\int_0^1 r_i di < 1, \text{ and} \tag{1.9}$$

$$r_i\sigma_m^2 \le \frac{2(\sigma_m^2+\sigma_{\eta i}^2)}{1+\sqrt{1+4\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}} \text{ for each } i\in[0,1], \text{ and} \tag{1.10}$$

$$r_i M\sigma_m^2 \le \pi_i(\sigma_m^2+\sigma_{\eta i}^2) \text{ for each } i\in[0,1], \text{ and} \tag{1.11}$$

$$\frac{\beta_i(M)\pi_i(\pi_i+\lambda_i r_i M\sigma_m^2)}{\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)\beta_i(M)^2+1} - E[w_i] \ge 0 \text{ for each } i\in[0,1], \tag{1.12}$$

where $E[w_i] = \frac{1}{2}\frac{\beta_i(M)^2(\pi_i-\lambda_i r_i M\sigma_m^2)(\pi_i+\lambda_i r_i M\sigma_m^2)}{\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)\beta_i(M)^2+1}+\frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2 + rW(M)$.

Note that (1.10) guarantees that each CEO's effort is bounded. (1.11) guarantees that the optimal CEO incentives do not exceed 1. (1.9) and (1.12) guarantee that the shareholders' payoffs are nonnegative for each firm. Since the CEO's effort not only affects the mean but also the variance of firm value, Lemma 1.3.1 sets forth that a CEO's effort is not monotonically increasing in the incentives. There is a threshold, $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$, such that the CEO's effort is increasing in the incentives below the threshold and decreasing in the incentives above the threshold. By Proposition 1.3.2, each CEO's optimal incentives depend on other CEOs' incentives. If the other CEOs are granted low incentives, then the target CEO's incentives can be low, such that her effort is increasing in the incentives; but if the other CEOs are granted high incentives, then the target CEO's incentives has to be high and may exceed the threshold, such that her effort becomes decreasing in the incentives. Therefore, there could exist multiple equilibria in this case. I illustrate this in the following example.

Example 2. *Suppose there are two types of firms. For $i \in [0,k]$, $\lambda_i = \lambda$, $\pi_i = \pi$, $\sigma_{\eta i}^2 = \sigma_\eta^2$, $r_i = r$. For $i \in [1-k,1]$, $\lambda_i = \lambda$, $\pi_i = \pi$, but $\sigma_{\eta i}^2 = \bar{\sigma}_\eta^2$, $r_i = \bar{r}$. Assume that the parameters satisfy the following* *conditions: 1)* $\lambda(\sigma_m^2+\sigma_\eta^2)>1$*; 2)* $\lambda r(\sigma_m^2+\bar{\sigma}_\eta^2)=\lambda\bar{r}(\sigma_m^2+\sigma_\eta^2)$*; 3)* $\frac{\bar{r}}{r}>1$*; 4)* $\frac{r+\bar{r}}{2\bar{r}}\left(1+\frac{\bar{r}-r}{\sqrt{(\bar{r}-r)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)\bar{r}r}}\right)<\frac{\sigma_m^2+\sigma_\eta^2-\sigma_m^2 rk}{\sigma_m^2 r(1-k)}<1.$ [19] *Then for firms in* $[0,k]$*, there must exist at least two equilibria: in one equilibrium, the optimal CEO incentives are less than the threshold* $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$*, thus the CEO's effort is increasing in the incentives; in another equilibrium, the optimal incentives are greater than the threshold* $\sqrt{\frac{1}{\lambda(\sigma_m^2+\sigma_\eta^2)}}$*, thus the CEO's effort is decreasing in the incentives.*

1.4 Extension: An Example of Stock Versus Options

In Sections 2 and 3, I focus on linear contracts for the purpose of tractability. In practice, though, options are widely used. Options provide insurance from the downside of returns, giving a CEO convex payoffs and inducing her to take risk (Guay (1999), Coles, Daniel and Naveen (2006)). Dittmann and Maug (2007), however, calibrate a standard principal-agent model and find that options should not be used. In this section, I will compare stock to options in the presence of RWCs in a simple example.

Suppose there are a continuum of firms and CEOs: $i\in[0,1]$. CEO i is assigned to firm i. All firms are identical in the sense that the corresponding parameters are equal for each firm. For simplicity, I assume that for each CEO, there is a binary effort decision $\{0\text{ (shirk)}, 1\text{ (work)}\}$. If CEO i chooses to shirk, then there is no cost and firm value is $V_i=e_0(\tilde{m}+\eta_i)$. If she chooses to work, then there is a cost c for her and firm value is $V_i=\pi+e_1(\tilde{m}+\eta_i)$, where $\tilde{m}$ refers to the aggregate shock that could be σ_m or $-\sigma_m$ with probability $\frac{1}{2}$ for each; η_i is firm i's idiosyncratic shock that could be σ_η or $-\sigma_\eta$ with probability $\frac{1}{2}$ for each; $\tilde{m}$ and $(\eta_i)_{i\in[0,1]}$ are independent of each other. e_0 is the firm's exposure to risk if the CEO shirks, and e_1 is the firm's exposure to risk if the CEO works. I assume that $e_1\geq e_0$. If $e_0=e_1$, then it corresponds to the situation in which the CEO's effort only increases the mean of firm value; if $e_0<e_1$, then it corresponds to the situation in which the CEO's effort increases both the mean and variance. For simplicity, I make the following assumptions:

Assumption 1: It is always beneficial for shareholders to induce CEOs to work (i.e., 1 is the optimal effort);

[19] Note that since $\lambda(\sigma_m^2+\sigma_\eta^2)>1$ and $\bar{r}>r$, $\frac{r+\bar{r}}{2\bar{r}}\left(1+\frac{\bar{r}-r}{\sqrt{(\bar{r}-r)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)\bar{r}r}}\right)<\frac{r+\bar{r}}{2\bar{r}}\left(1+\frac{\bar{r}-r}{\sqrt{(\bar{r}-r)^2+4\bar{r}r}}\right)=1$. So the condition 4) can be satisfied.

Because of the convex payoff of options, I will relax the restriction on linear contracts and make the following assumption:

Assumption 2: For shareholders in each firm, they will use one of the two strategies to induce CEOs to work: either use a stock-based contract (i.e., salary + stock), or use an option-based contract of exercise price π (i.e., salary + options).

The CEO in firm i has RWCs with her expected utility defined in (1.2). The shareholders will choose either stock or options to induce the CEO to work. For simplicity, I assume that when there is no difference between stock and options, shareholders will choose to use stock to induce the CEO's effort. Then, since firms are identical, in equilibrium, either all the firms use options or all the firms use stock.

The presence of RWCs has two effects when comparing stock with options. First, if the CEO chooses to shirk, then a stock-based contract still has an exposure e_0 to aggregate shock; but an option-based contract becomes more insensitive to shock, because options are more likely to expire with value zero. Recall that in the presence of RWCs, CEOs are willing to risk exposure to aggregate shock because it helps them to keep up with their peers. Thus, the effect of RWCs on incentivizing a CEO's effort is larger for an option-based contract than a stock-based contract. This "incentive effect" makes options more preferable.

On the other hand, to some extent, RWCs make CEOs willing to expose to aggregate shock, which reduces the risk premium associated with aggregate shock. Because stock is completely exposed to the aggregate shock, while options are only partially exposed, the reduction in risk premium is larger for stock than options. For this symmetric case, the proof in the Appendix shows that *the ratio of the risk premium of one unit of stock to the risk premium of one unit of options* is

$$\kappa(r) = \frac{4[(1-r)^2\sigma_m^2 + \sigma_\eta^2]}{(1-r)^2\sigma_m^2 + \sigma_\eta^2 + \min(\sigma_m^2, \sigma_\eta^2)},$$

which is decreasing in r. Thus this "risk premium effect" makes stock preferable. Therefore, depending on which effect is more dominant, RWCs can render either stock or options more preferable.

Lemma 1.4.1. *Suppose that $r < 1$ and $\pi \geq e_0(\sigma_m + \sigma_\eta)$. Then there exists a cutoff $\hat{\pi}$ such that shareholders in all the firms will use stock if and only if $\pi \geq \hat{\pi}$.*

1) If $\frac{\sigma_\eta}{\sigma_m} \to 0$, *then* $\frac{\partial \kappa(r)}{\partial r}$ *goes to 0, and* $\frac{\partial \hat{\pi}}{\partial r} > 0$.

2) If $\sigma_m^2 \geq \sigma_\eta^2 > \frac{4\lambda c(1-r)^2}{(1-r)(1-8\lambda c)\sqrt{1-6\lambda c}-8\lambda c}\sigma_m^2$, *then* $\frac{\partial \hat{\pi}}{\partial r} < 0$.

3) If $\frac{\sigma_m}{\sigma_\eta} \to 0$, *then* $\frac{\partial \hat{\pi}}{\partial r} = 0$.

In part 1) of Lemma 1.4.1, when aggregate risk is much more important relative to idiosyncratic risk, we can see that the ratio of risk premiums converges to a constant, and the risk premium effect disappears. As a result, the incentive effect dominates and renders options more preferable. Part 2) indicates that if idiosyncratic risk is not too small compared with aggregate risk, then the risk premium effect can dominate the incentive effect such that stock becomes more preferable. Part 3) suggests that if aggregate risk becomes much smaller than idiosyncratic risk, then both effects tend to disappear. Therefore, RWCs do not affect the comparison between stock and options.

1.5 Empirical Implications

By introducing RWCs, the model generates a number of empirical predictions on executive compensation. I outline the predictions below.

1. Individual CEO compensation and aggregate CEO compensation

Due to RWCs, one expectes to see a correlation between CEOs' compensation. The model predicts a positive correlation in the level of CEO pay and CEO incentives. Proposition 1.2.2 shows that the expected individual CEO pay is increasing in the level of aggregate expected CEO pay and aggregate CEO incentives. Bouwman (2011), Bizjak, Lemmon, and Naveen (2008), as well as Faulkender and Yang (2010) find that the level of CEO pay increases in the average level of CEO pay in a reference group. Although market competition could also lead to a positive correlation among the level of CEO pay, it is not clear how market competition will affect the relation between CEO incentives. My RWCs-based model predicts a positive relation between CEO incentives. It will be interesting to test the positive relation between individual CEO incentives and aggregate CEO incentives. To test the relation, it is critical to identify the reference group. Bouwman (2011) uses geographically-close firms as a reference group, Bizjak, Lemmon, and Naveen (2008), Faulkender and Yang (2010) use similar-sized firms in the same industry as a reference group. So geography, industry and firm size might be the potential factors to identify a reference group in an empirical test. In addition, it is also a challenge to solve the causality problem, because the increase in individual incentives and aggregate incentives may be due to other factors – for example, a common shock to aggregate risk may increase both individual incentives and aggregate incentives simultaneously.

2. CEO incentives and firm risk

Prendergast (2002) provides a survey of empirical evidence on the relation between CEO incentives and firm risk. The results are mixed: three papers find negative relation consistent with the informativeness principle, but eight other papers find no significant or even a positive relation. In Proposition 1.2.2, I show that while the relation between incentives and a firm's idiosyncratic risk is negative, the relation between incentives and aggregate risk could be negative or positive, depending on the coefficient of a CEO's RWCs on aggregate shock (i.e., rM, where r measures the CEO's concerns about other CEOs' pay, and M is aggregate CEO incentives). In particular, the model predicts that if $r\left[\frac{\partial M}{\partial \sigma_m^2}\sigma_m^2+\left(1-\frac{\lambda\sigma_m^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)M\right]>\frac{\pi^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$, then β^* is increasing in σ_m^2. So if r is big, and (or) the aggregate incentives M is big, and (or) the idiosyncratic risk σ_η^2 is big, then there is a positive relation between incentives and aggregate risk. The prediction that aggregate CEO incentives also affect the relation between individual CEO incentives and aggregate risk is novel.

3. Pay for luck

Bertrand and Mullainathan (2001) document that CEOs are rewarded for factors that are beyond their control, i.e., paid for luck. Proposition 1.2.8 provides a simple explanation of pay-for-luck puzzle. Moreover, it predicts that a CEO is paid more for luck if her concerns about other CEOs' pay are more intensive (i.e., r is bigger), or if other CEOs in the reference group are paid more for luck.

4. RWCs and risk-taking incentives

Empirical studies have documented a negative relation between a CEO's sensitivity to stock price movement and the CEO's risk-taking behavior (Bizjak, Brickley and Coles (1993), Coles, Daniel and Naveen (2006), Aggarwal and Samwick (2006)). But because CEO compensation is endogenously determined, there is a causality problem when testing the relation between incentives and risk-taking behavior. Gormley, Matsa and Milbourn (2011) address this identification problem by studying unexpected changes in the firms' business environment that increase left-tail risk. They find that managers with incentives that provide a high sensitivity to stock price movements respond to the increased tail risk with greater risk-reducing activities. My model predicts that the relation between incentives and a CEO's risk-taking behavior depends on a CEO's RWCs. Proposition 1.3.2 shows that if rM exceeds some threshold, then granting a CEO more incentives can reduce her risk-taking effort. Some other papers (Carpenter (2000), Ross (2004)) also point out that granting a CEO stock options may reduce the CEO's risk-taking incentives, but those papers do not endogenize the optimal CEO incentives. My model shows that RWCs may cause the optimal CEO incentives to be at a high level, such that a negative relation between CEO incentives and risk-taking behavior arises at the optimum.

5. Stock versus options

Lemma 1.4.1 shows that if aggregate risk is much larger than idiosyncratic risk, then the incentive effect is dominant, so options are more preferable. But if idiosyncratic risk is not dominated by aggregate risk, then the risk premium effect is dominant such that stock becomes more efficient. To my knowledge, the comparison between stock and options has not been associated with the relative importance of aggregate risk and idiosyncratic risk in the previous literature.

1.6 Conclusion

Some recent empirical studies on executive compensation (Bouwman (2011)) find that relative wealth concerns affect firm compensation policies. In this paper, I incorporate RWCs into a standard principal-agent model and study how CEO compensation will be affected. I focus on RWCs' effects on CEO incentives and find a number of interesting implications. I show that the relation between incentives and aggregate risk could be negative or positive, depending on the coefficient of the CEO's RWCs on aggregate shock. This can help to explain the mixed empirical evidence on the relation between incentives and risk. I also provide a simple explanation for pay-for-luck puzzle. If a CEO's efforts also affect the variance of firm value, I show that if the CEO's RWCs are intensive, then it can lead to a negative relation between the CEO's risk-taking behavior and her sensitivity to stock price movement. Finally, I compare stock to options in the presence of RWCs. I find that RWCs may render options or stock more preferable, depending on the relative importance of aggregate risk and idiosyncratic risk. This paper makes a number of predictions about RWCs' effects on CEO compensation, as well as some predictions concerning the relation between individual CEO compensation and aggregate CEO compensation (including the level of CEO pay and CEO incentives). The predictions, however, pose certain challenges for empirical tests: how to identify RWCs' effects; how to find an appropriate measure for a CEO's concerns about peer compensation; and how to solve the causality problem.

Chapter 2

Inside Debt

2.1 Introduction

Shareholders ultimately bear the agency costs suffered by other stakeholders (Jensen and Meckling, 1976). Therefore, it appears intuitive that they should pay the manager according to firm value, rather than equity value alone. In particular, Jensen and Meckling speculated that granting the manager equal proportions of debt and equity might attenuate the stockholder-bondholder conflicts that arise when the manager is purely equity-aligned. However, this idea of compensating the manager with "inside debt"[1] has not since been pursued further. Instead, the intervening three decades of compensation theories have focused on justifying equity-like instruments, such as stock and options. In particular, a number of models suggest that bonuses for avoiding bankruptcy, salaries or managerial reputation are adequate remedies to the agency costs of debt, leaving no role for inside debt in efficient compensation (see, e.g., Hirshleifer and Thakor, 1992; Brander and Poitevin, 1992; John and John, 1993). However, the substantial bondholder losses in the recent financial crisis suggest that the agency costs of debt are not fully solved.

Theorists' focus on rationalizing equity pay has likely been driven by the long-standing belief that, empirically, executives do not hold debt (see, e.g., the survey of Murphy, 1999). Accordingly, Dewatripont and Tirole (1994) seek to answer the question "why are managers' monetary incentives ... traditionally correlated with the value of equity rather than the value of debt?" However, recent empirical studies

[1]This paper defines "inside" debt as debt (or any security with payoffs very similar to debt) held by the manager. It contrasts with outside debt, which is held by external investors.

(Bebchuk and Jackson, 2005; Sundaram and Yermack, 2007; Gerakos, 2007; Wei and Yermack, 2009) find that U.S. CEOs hold substantial defined benefit pensions. These are unsecured, unfunded obligations which, in nearly all cases, have equal priority with other creditors in bankruptcy and thus constitute inside debt.[2] Researchers have also noted the common use of deferred compensation, another form of inside debt, although systematic studies have so far been limited by data availability.[3]

Inside debt is therefore widespread. Such compensation contrasts with existing theories, which do not advocate debt but instead the exclusive use of equity-like compensation. Indeed, Sundaram and Yermack note the lack of a theoretical framework for their results: "the possibility of using debt instruments for management compensation has received little attention ... A top priority would appear to be the development of theory that illustrates conditions under which debt-based compensation ... represent[s] the solution to an optimal contracting problem." Does the absence of a theoretical justification mean that inside debt constitutes rent extraction, as argued by Bebchuk and Jackson?[4] Or can it be part of efficient compensation, and if so, under what conditions? Should the manager's debt-equity ratio equal the firm's, so that he is aligned with firm value as Jensen and Meckling hypothesized? What factors affect the optimal level of inside debt?

These questions are the focus of this paper. We start with a model in which the manager makes a project selection decision; the optimal project depends on a signal privately observed by the manager after contracting. We consider a set of standard securities: debt, equity and a fixed bonus that pays off only in solvency, and initially assume that the manager holds an exogenous equity stake to create risk-shifting incentives. We demonstrate that inside debt is a superior remedy to the agency costs of debt than the bonuses advocated by prior research. Bonuses are effective in encouraging the manager to avoid bankruptcy, since they are only received in solvency. However, creditors are concerned with not only the probability of default, but also recovery values in default. Optimal contracts should therefore depend on the value of assets in bankruptcy, as well as the occurrence of bankruptcy. This is the critical difference

[2]See Section 4 for further discussion of the priority of pensions in bankruptcy.

[3]Despite limited data, anecdotal evidence suggests that such compensation may be substantial. For example, Roberto Goizueta, the former CEO of Coca-Cola, had over $1 billion in deferred compensation when he died. Wei and Yermack (2009) consider total inside debt holdings (pensions plus deferred compensation).

[4]An alternative view is that pensions are tax motivated. Bebchuk and Jackson (2005) and Gerakos (2007) provide a number of arguments against this explanation, most notably that executive pensions enjoy a different tax status from employee pensions. This paper takes a neutral stance: if pensions are indeed tax-motivated, the model investigates whether the tax system encourages firms to adopt otherwise inefficient compensation schemes.

between inside debt and bonuses: inside debt yields a positive payoff in bankruptcy, proportional to the liquidation value. Thus it renders the manager sensitive to the firm's value in bankruptcy, and not just the incidence of bankruptcy – exactly as desired by creditors. By contrast, bonuses have zero bankruptcy payoffs, regardless of the liquidation value, and so represent binary options rather than debt.

This difference in payoffs is important. Even in situations where bonuses can attenuate risk-shifting, inside debt can be a cheaper solution since its sensitivity to liquidation values renders it a more powerful instrument. Moreover, in some settings, bonuses *aggravate* risk-shifting owing to their binary nature. Since they only pay off in solvency, the manager may inefficiently sacrifice liquidation value to gamble for solvency. The same issues apply to other instruments which have zero payoff in bankruptcy regardless of the liquidation value, e.g. salary (if it is junior to creditors) or reputation (under Hirshleifer and Thakor's (1992) assumption that the labor market can only assess the incidence rather than severity of bankruptcy.) For brevity, we refer to all of these instruments as "bonuses." These alternative measures were shown to be adequate under specific frameworks in which only sensitivity to the incidence of bankruptcy matters, such as where solvency can be guaranteed (John and John, 1993), or liquidation value is always zero (Hirshleifer and Thakor). In the more general setup of this paper, the manager can affect the liquidation value and so his compensation should be sensitive to it.

We then extend the model to incorporate an effort decision, which allows us to endogenize the manager's equity stake. This analysis extends previous models which focus on the agency costs of debt (project selection) and do not incorporate the agency costs of equity (effort). This is a necessary extension, since in the absence of a shirking problem, risk-shifting can be trivially solved by removing the manager's equity and giving him a flat salary; here, equity compensation is optimal to induce effort. The compensation scheme typically does not involve pure equity compensation (as advocated by some existing research) nor giving the manager debt and equity in equal proportions (as intuition might suggest, and as Jensen and Meckling hypothesized). In the most common case, equity is more effective than debt in inducing effort and so an equity bias is desired, where the manager's percentage equity stake exceeds his percentage debt holding. Even though an equity bias leads to occasional risk-shifting, this is compensated for by greater effort. The optimal debt-equity mix depends on the relative importance of these two agency problems – the ratio of debt to equity is increasing in leverage, the probability of bankruptcy and the manager's impact on liquidation value, but decreasing in growth opportunities (i.e. the effect of effort on solvency value). However, a debt bias, where the manager's percentage debt stake exceeds his percentage equity holding, may be optimal if effort has a high expected payoff in bankruptcy, either because bankruptcy is likely, or because effort is particularly productive in enhancing liquidation value. In contrast to the "agency costs of equity" nomenclature, suboptimal effort may result from insufficient inside debt, rather than equity. Indeed, a debt bias is found by Sundaram and Yermack (2007) in 13% of cases.

Finally, we relate the model's empirical implications to recent findings. Most notably, inside debt compensation is widespread, whereas solvency-contingent bonuses have not yet been documented. Also as predicted, pensions are increasing in firm leverage (Sundaram and Yermack), decreasing in growth opportunities (Gerakos, 2007) and associated with lower risk-taking, as measured by the firm's "distance to default" (Sundaram and Yermack) or credit rating (Gerakos). Wei and Yermack (2009) find that disclosures of large inside debt holdings lead to an increase in bond prices and a fall in equity prices. In addition, the model provides a theoretical framework underpinning recent normative proposals to reform executive pay by compensating the manager with debt as well as equity, to help prevent the significant bondholder losses that manifested in the recent financial crisis (see, e.g., Bebchuk and Spamann, 2009.)

Jensen and Meckling (1976) were the first to theorize the agency costs of debt. They include a brief verbal section wondering why inside debt (awarded in the same proportion as inside equity) is not used as a solution, but are "unable to incorporate this dimension formally into our analysis in a satisfactory way." They speculate that the manager's salary is a sufficient mechanism and thus have no role for inside debt. This paper shows that salaries are problematic given their insensitivity to liquidation value, and that equal proportions of debt and equity are generally suboptimal.

John and John (1993), Brander and Poitevin (1992) and Hirshleifer and Thakor (1992) also demonstrate that the agency costs of debt can be alleviated through certain compensation instruments. Since their goal is to show the effectiveness, rather than optimality, of their proposed solutions, they do not consider whether alternative mechanisms, such as inside debt, would be superior. A second distinction is that this paper incorporates the agency costs of equity as well as of debt. This provides an endogenous justification for the equity compensation that is the cause of asset substitution and allows analysis of the trade-off between effort and project selection, thus leading to empirical predictions on the optimal ratio of debt to equity.[5] Hirshleifer and Thakor (1992) show that managerial reputation (assumed to be zero in all bankruptcy states) can deter risk-shifting. In their model, liquidation value is always zero and so only sensitivity to the probability of bankruptcy matters. In the more general setup of this paper, the manager can affect the liquidation value and so his compensation should be sensitive to it, which is not achieved by a binary instrument.[6] John and John (1993) advocate two solutions to risk-shifting. The first is a solvency-contingent bonus, which has similar issues to reputation owing to its binary payoff. The second is to reduce the manager's equity. This is possible as their model has no effort decision; indeed, there is no reason to give the manager any incentive pay. Brander and Poitevin (1992) propose a more general fixed bonus, which may be triggered at levels other than solvency. They note that if the firm is sufficiently levered, no bonus can eliminate the agency costs of debt. Here, inside debt is effective even where bonuses are impotent.

Dybvig and Zender (1991) ("DZ") show that an optimal contract can alleviate the Myers and Majluf (1984) "lemons" issue, thus resurrecting the Modigliani-Miller irrelevance theorems. Although they focus on adverse selection rather than risk-shifting, the insight that incentives can achieve first-best is potentially applicable to other agency problems. However, their interest is not on what the contract *is*, but that an optimal contract (whatever form it may take) can render financing irrelevant. By contrast, this paper is focused on the *form* of pay. First, it shows that inside debt can be superior to the instruments advocated by a number of earlier papers, whereas DZ do not compare different contracts. Second, we analyze the optimal relative proportions of debt and equity, generating empirical predictions on the cross-sectional determinants of the inside debt level. While the optimal contract in the core DZ model aligns the manager with firm value, here the manager should not hold debt and equity in equal proportions if there is an effort decision.[7]

[5]Jensen and Meckling consider the agency costs of debt and equity separately, not simultaneously. Biais and Casamatta (1999) and Hellwig (2009) do consider both agency costs together. They do not analyze executive compensation (which remains pure equity), but an entrepreneur's choice of outside financing. Stoughton and Talmor (1999) also consider contracting under both an effort and investment decision. Investment is undertaken by shareholders (rather than the manager) and does not involve risk-shifting as the firm is unlevered.

[6]While Hirshleifer and Thakor consider reputation in the managerial labor markets, Diamond (1989) considers reputation in debt markets and shows that it can deter risk-shifting; however, it requires the manager to expect to continue to raise debt in the future. The solutions considered here and in earlier papers work in a one-shot game.

[7]When DZ extend their model by introducing an effort decision, they are unable to solve for the optimal contract and note that the solution may not exist. This does not matter for them, since the form of the contract and its comparative static determinants are not the focus of their paper. Their goal is to show that "*if* there is a solution, it is independent of capital structure."

2.2 Debt and Project Selection

2.2.1 The Model

The model consists of four periods. At $t=-1$, shareholders offer a contract to the manager. At $t=0$, risky debt is raised with face value of F and market value $D_0 < F$. $t=0$ total firm value (gross of expected pay) is $V_0 = E_0 + D_0$. All agents are risk-neutral and the risk-free rate and reservation wage are normalized to 0. Bondholders observe the manager's contract when calculating D_0, and so at $t=-1$, shareholders select the contract that maximizes V_0 minus expected pay. The assumption that F is exogenous is discussed at the end of Section 2.2.2.

At $t=1$, the manager chooses one of two mutually exclusive projects: R (risky) or S (safe). (S can be thought of as the firm's "status quo" state, and the manager is deciding whether to switch to the riskier project R.) R has probability p_R of "success," in which case the firm is worth V_{GR} at $t=2$. In "failure" (which occurs with probability $(1-p_R)$), firm value is V_{BR}. S pays V_{GS} with probability p_S, and V_{BS} otherwise. We assume that $p_R \leq p_S$, $V_{GR} \geq V_{GS} \geq V_{BS} \geq V_{BR}$, $V_{GR} > F$ and $V_{BS} < F$: failure of either project leads to bankruptcy, and success of R leads to solvency. We subdivide the model into two cases, depending on whether the success of S leads to solvency: $V_{GS} \geq F$ (Case 1) and $V_{GS} < F$ (Case 2). At $t=2$, all payoffs are realized and the debt matures.

If all parameters were known at $t=-1$, shareholders would know the optimal project and can implement it with certainty. In reality, unforeseen projects often appear after pay is set, and so the contract should induce the manager to accept (reject) any new project R that appears and offers a higher (lower) NPV than the status quo S. We therefore assume $V_{GR} \sim U[V_{GS}, V_{GRH}]$. The optimal project is not known in advance; either R or S may be first-best depending on the realization of V_{GR}, which is observed privately by the manager at $t=1$. All other parameters are public at $t=-1$.

While the core model focuses on project selection, it can easily be extended to involve other agency costs of debt. Hence the terms "asset substitution" and "risk-shifting" should be interpreted as any action that benefits shareholders but reduces total firm value. Other examples include debt overhang, concealing information, failing to disinvest, or paying excessive dividends. Many of these actions were believed to be important in the recent financial crisis.

Prior research on the solutions to (rather than causes of) risk-shifting typically take risk-shifting incentives as given by assuming that the manager exogenously owns a proportion α of the firm's equity, and shows that certain compensation instruments can remove these incentives (e.g. John and John, 1993). This section follows this approach by also taking α as exogenous and deriving the optimal accompanying compensation scheme; Section 3 endogenizes α through the introduction of an effort decision. Our goal is to show whether and under what conditions debt is superior to the bonuses previously advocated, and so we allow the accompanying scheme to consist of a fraction β of the firm's debt and/or a bonus of J, paid if and only if the firm is solvent. The debt is "locked up" (similar to restricted stock) to prevent its subsequent sale or renegotiation; indeed, pension claims and deferred compensation cannot be sold in practice. β and J are choice variables, whereas α is currently fixed. Note that $\alpha > 0$ can arise exogenously even without the firm granting equity, for instance if the manager's labor market reputation is linked to the equity price (e.g. Hirshleifer and Thakor, 1992; Gibbons and Murphy, 1992).

Case 1: $V_{GS} \geq F$

Incentive compatibility requires the manager to choose S, i.e., inequality (2.2) below is satisfied, if and only if it has a higher NPV, i.e., inequality (2.1) is satisfied:

$$p_R V_{GR} + (1-p_R)V_{BR} \leq p_S V_{GS} + (1-p_S)V_{BS}, \tag{2.1}$$

$$\begin{aligned} \text{iff} \qquad & p_R[\alpha(V_{GR}-F)+\beta F+J]+(1-p_R)\beta V_{BR} \\ \leq \quad & p_S[\alpha(V_{GS}-F)+\beta F+J]+(1-p_S)\beta V_{BS}. \end{aligned} \tag{2.2}$$

The "and only if" requirement ensures that the contract does not lead to excessive conservatism. If and only if R has a higher NPV, the contract must ensure that the manager chooses R over S. The Appendix shows that the above is satisfied if and only if:

$$\beta = \alpha - \frac{J(p_S - p_R)}{F(p_S-p_R)+(1-p_S)V_{BS}-(1-p_R)V_{BR}}. \tag{2.3}$$

Under first-best project selection, R is chosen if and only if V_{GR} exceeds a cutoff V_{GR}^{**}, where

$$V_{GR}^{**} = \frac{p_S V_{GS} + (1-p_S)V_{BS} - (1-p_R)V_{BR}}{p_R}.$$

Let $q = \Pr(V_{GR} > V_{GR}^{**})$, i.e. the probability that R is first-best. The optimal compensation scheme is the cheapest incentive compatible contract, i.e. solves

$$\min_{\beta,J} q\left[p_R\left(\beta F + J\right) + (1-p_R)\beta V_{BR}\right] + (1-q)\left[p_S\left(\beta F + J\right) + (1-p_S)\beta V_{BR}\right] \text{ s.t. (2.3).} \tag{2.4}$$

Proposition 2.2.1. *If $V_{GS} \geq F$, the optimal compensation scheme (β^*, J^*) is given by*

$$\begin{cases} (\alpha, 0) \ \textit{if}\ p_R(1-p_S)V_{BS} > p_S(1-p_R)V_{BR}, \\ \left(\alpha - \frac{J(p_S-p_R)}{F(p_S-p_R)+(1-p_S)V_{BS}-(1-p_R)V_{BR}}, J\right) \ \textit{if}\ p_R(1-p_S)V_{BS} = p_S(1-p_R)V_{BR}, \\ \quad \textit{where}\ J \in \left[0, \alpha\left[F + \frac{(1-p_S)V_{BS}-(1-p_R)V_{BR}}{p_S-p_R}\right]\right], \\ \left(0, \alpha\left[F + \frac{(1-p_S)V_{BS}-(1-p_R)V_{BR}}{p_S-p_R}\right]\right) \ \textit{if}\ p_R(1-p_S)V_{BS} < p_S(1-p_R)V_{BR}. \end{cases} \tag{2.5}$$

If $p_S >> p_R$ and V_{BS} is close to V_{BR}, the main advantage of S over R is its greater probability of solvency. Hence the bonus should be used exclusively: its zero bankruptcy payoff makes it particularly sensitive to solvency. Indeed, John and John (1993) assume $p_S = 1$ and find that a bonus can be effective. However, if p_S is close to p_R and $V_{BS} >> V_{BR}$, the main advantage of S is its greater liquidation value. Inside debt should be used exclusively as, unlike the bonus, its payoff is sensitive to liquidation value. Indeed, if $p_S = p_R$, the bonus is completely ineffective.

Case 2: $V_{GS} < F$

If $V_{GS} < F$, the firm is definitely liquidated if S is undertaken. Even though it leads to certain liquidation, S can still be preferred, if (2.1) is satisfied. The incentive constraint (2.3) becomes:

$$\beta = \alpha + \frac{p_R J}{p_S V_{GS} + (1-p_S)V_{BS} - (1-p_R)V_{BR} - p_R F}. \tag{2.6}$$

Proposition 2.2.2. *If $V_{GS} < F$, the optimal compensation scheme (β^*, J^*) is given by $\beta^* = \alpha$, $J^* = 0$.*

In Case 1, (2.3) shows that bonuses (partially) alleviate asset substitution and reduce the amount of inside debt required. In Case 2, bonuses *exacerbate* asset substitution owing to their binary nature. Since $V_{GS} < F$, the bonus is only received if the risky project is chosen and is successful. It thus induces the manager to choose R, even if (2.1) is satisfied and so voluntary liquidation through the choice of S is efficient. Introducing a bonus increases the level of debt required to achieve optimal project selection (see (2.6)) and so is counter-productive. The optimal compensation scheme therefore involves zero bonus and only inside debt. This may explain why the solvency-contingent bonuses advocated by prior literature are rarely used. Since inside debt is even more favored in Case 2, we consider Case 1 for the remainder of the paper.

2.2.2 Discussion

We now discuss whether alternative mechanisms can attenuate asset substitution and thus render inside debt unnecessary. Jensen and Meckling (1976) speculated that salaries might constitute inside debt. Most theory papers (e.g. Innes, 1990) assume that salary is junior to creditors and thus not received in bankruptcy; this is indeed the case in countries with pure liquidation bankruptcy codes and little room for renegotiation in debt workouts (Calcagno and Renneboog, 2007). In such a case, salary functions like a bonus and is thus different from inside debt. By contrast, Calcagno and Renneboog cite bankruptcy regulations in certain countries (e.g. US, UK and Germany) that management can use to ensure that salaries are senior to creditors in a bankruptcy, and give a number of examples where this occurred. If salaries are received in all states of nature, they have no effect on incentive constraints and thus the manager's decisions. (They do not mention any cases in which salaries have equal priority to other creditors, nor are we aware of any.)

Private benefits, such as firm-specific human capital, prestige, perks, and the present value of future wages are principally determined by whether the firm is solvent: if the manager is fired upon bankruptcy, he can no longer derive benefits from incumbency, regardless of liquidation value. They therefore have a very similar effect to bonuses J. Moreover, private benefits plausibly increase with shareholder value and thus may be incorporated into α, increasing the need for inside debt.

In the model, the face value of debt is set at the first-best level F, which is optimal in the absence of agency costs of debt. These costs are foreseen by rational creditors and thus shareholders suffer a discount when raising debt. Alternatively, the trade-off theory would advocate lowering debt to a second-best level $F^{SB} < V_{BR}$, so that the firm is never bankrupt and no discount is suffered. This loses some of the benefits of debt, such as tax shields. Either way, shareholders have an incentive to reduce the agency costs of debt, to augment tax shields or to reduce the discount. Results would be unchanged (at the cost of complicating the model) if F was endogenized through the introduction of taxes, as in Hirshleifer and Thakor (1992): inside debt allows the firm to increase F and thus create additional tax shields.

Covenants are an imperfect solution due to the incompleteness of contracts: see, for example, the discussion in Myers (1977). Covenants may increase asset substitution, as the manager risk-shifts even when the firm is some distance from bankruptcy to avoid breaching the covenant. Also, covenants may not be breached until after the key decision has been made (e.g., R was irreversibly chosen and failed, leading to the covenant violation).[8]

2.3 Debt, Equity, Project Selection and Effort

Section 2 followed prior literature by taking α as exogenous and deriving the optimal accompanying compensation scheme to achieve efficient project selection. However, a complete analysis must provide an endogenous justification for equity compensation, else the optimal contract would be $\alpha = \beta = 0$. This section endogenizes $\alpha > 0$ as the solution to an effort problem. In addition to project selection, the manager now also makes an effort decision. He chooses a pair $\left(e_g, e_b\right)$ where $e_g, e_b \in [0, e_H]$ and $e_H < 1$. e_g increases the firm's solvency value by g with probability e_g and costs the manager $\frac{1}{2}e_g^2$; e_b increases the firm's bankruptcy value by b with probability e_b, where $V_{BS} + b < F$, and costs the manager $\frac{1}{2}e_b^2$. If the firm has abundant intangible growth opportunities, such as employee training and building customer relationships, g will be high; if there is scope to scrap investment projects or liquidate assets, b will be high.

To keep the model tractable in the presence of an effort decision, it is necessary to specialize it to $p_R = p_S = p$. This rules out complex feedback effects between effort and project selection.[9] The two de-

[8]Several studies demonstrate that executive compensation has a significant effect on risk, despite covenants. Examples include DeFusco et al. (1990), Guay (1999), and Coles et al. (2006).

[9]Tractability concerns also motivate the assumption that effort only affects firm values, not the probability of success. If effort affects both parameters, it is impossible to solve analytically for the manager's effort choice, and to compute the effect of compensation on effort. In reality, it is plausible that effort increases the probability of success. This has a very similar effect to

cisions can thus be analyzed separately, allowing the effects of compensation on each to be seen cleanly. For example, an increase in β directly raises effort, since the manager benefits more from enhancing liquidation value. Raising β also directly leads to S being chosen with greater frequency. If $p_R = p_S = p$, this change in project selection does not affect the probability of solvency (which always equals p), and thus does not feed back into the effort decision (which depends on the probability of solvency). If instead $p_S > p_R$, the greater frequency with which S is chosen leads to a higher probability of solvency. If effort is particularly productive in solvency, this creates a positive feedback effect on effort: in (2.8) below, e_g is increasing in the probability of solvency. While such feedback makes it easier to justify inside debt, the model can no longer be solved analytically since project selection and effort cannot be considered separately. A consequence of $p_R = p_S$ is that the bonus is ineffective (see Proposition 2.2.1) and so efficient compensation involves debt and equity alone. Note that even with $p_S > p_R$, the bonus provides no incentives to exert effort: a rise in bankruptcy value has no effect on the bonus but increases the value of inside debt. Thus, if inside debt is preferred to bonuses for solving asset substitution, it remains superior when an effort decision is introduced.

Shareholders maximize firm value net of pay to the manager, i.e. solve

$$\max_{\alpha,\beta} V_0 - \alpha E_0 - \beta D_0 = \max_{\alpha,\beta} (1-\alpha) E_0 + \left(1-\beta\right) D_0, \tag{2.7}$$

subject to $\alpha \geq 0$ and $\beta \geq 0$.[10] We first consider the manager's effort decisions.

Lemma 2.3.1. *The manager chooses effort levels*

$$\begin{aligned} e_g^* &= p\alpha g, \\ e_b^* &= (1-p)\beta b. \end{aligned} \tag{2.8}$$

We now consider project selection. Firm value is maximized if R is selected if and only if $V_{GR} > V_{GR}^{**}$,

raising g, the value created by effort in success, in the analysis that follows (i.e. it increases the optimal ratio of α to β), and thus has the same comparative static effect.

[10] This is because of limited liability. In addition, we rule out $\alpha < 0$ since it is illegal for CEOs to short their own firm. Allowing $\beta < 0$ would, under some parameters, lead to the uninteresting result that the CEO should borrow to buy the entire firm. It is standard that a moral hazard problem under risk neutrality requires limited liability, otherwise the first-best can always be achieved. In addition, $\beta < 0$ would correspond to the CEO being given a loan by his firm. In the U.S., executive loans were prohibited by the 2002 Sarbanes-Oxley Act.

where V_{GR}^{**} is defined by

$$V_{GR}^{**} = V_{GS} + \frac{(1-p)}{p}(V_{BS} - V_{BR}). \tag{2.9}$$

However, the manager will choose R if and only if $V_{GR} > V_{GR}^{*}$, where

$$V_{GR}^{*} = V_{GS} + \frac{\beta(1-p)}{\alpha p}(V_{BS} - V_{BR}). \tag{2.10}$$

The cutoff V_{GR}^{*} is undefined for $\alpha = \beta = 0$, since the manager is indifferent between all projects. We assume that he chooses the efficient project as this is Pareto optimal.[11] The optimal levels of debt and equity are determined by a trade-off between their differential effects on effort and project selection. We are interested not only in the absolute levels of α and β but also their relative magnitudes – in particular, Jensen and Meckling (1976) speculated the optimal compensation scheme would involve $\beta = \alpha$. We thus also study the ratio β/α which we define as k. In particular, we are interested in whether $k = 1$ (i.e. $\beta = \alpha$) or whether the contract involves an equity bias ($k > 1$) or a debt bias ($k < 1$). Our main results are summarized in the following Proposition.

Proposition 2.3.2. *The optimal compensation scheme* (β^*, α^*) *satisfies the following:*

(i) $\beta^* < \frac{1}{2}$ *and* $\alpha^* < \frac{1}{2}$;

(ii) If $p^2 g^2 + (1-p)^2 b^2 > \frac{V_{GRH}+V_{GS}}{2} + \frac{1}{2}\frac{(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}} + (1-p)V_{BR}$ *and* $(1-p)^2 b^2 - [pF + (1-p)(V_{BS} - V_{BR})] \leq 0$, *then* $\alpha^* > 0$;

(iii) If $p^2 g^2 + (1-p)^2 b^2 > \frac{V_{GRH}+V_{GS}}{2} + \frac{1}{2}\frac{(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}} + (1-p)V_{BR}$ *and* $\frac{2(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}} + (1-p)^2 b^2 > pF + (1-p)V_{BR}$, *then* $\beta^* > 0$.

If the conditions in (ii) and (iii) are satisfied, we also have the following comparative statics:

(iv) β^* *is increasing in* b; α^* *is increasing in* g;

(v) If $pF + (1-p)V_{BR} > 2(1-p)b^2$, *then* $k^* < 1$. k^* *is increasing in* $V_{BS} - V_{BR}$ *and decreasing in* g *and* p;

(vi) If $p\left[\frac{V_{GRH}+V_{GS}}{2} - F\right] > p^2 g^2 + \frac{1}{2}\frac{(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}}$, *then* $k^* > 1$. k^* *is increasing in* b, *and decreasing in* $V_{BS} - V_{BR}$ *and* F.

(vii) D_0 *is increasing in* β, E_0 *is decreasing in* β.

[11] If we instead assume that he always takes project R if $\alpha = \beta = 0$, this allows us to drop the first condition in parts (ii) and (iii) of Proposition 2.3.2.

We now discuss the intuition behind each component of Proposition 2.3.2. Starting with part (i), increasing α augments effort, but is costly to the firm. Given the convex cost function, equity has a diminishing marginal effect on effort; when $\alpha \geq \frac{1}{2}$, this benefit is insufficient to outweigh the costs and so the optimal α is less than $\frac{1}{2}$. A similar argument applies for β.

The common condition in parts (ii) and (iii), $p^2 g^2 + (1-p)^2 b^2 > \frac{V_{GRH}+V_{GS}}{2} + \frac{1}{2}\frac{(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}} + (1-p)V_{BR}$, is sufficient to rule out the optimum involving $\alpha^* = \beta^* = 0$. Intuitively, if effort is very unproductive (g and b are small), then it is not worth giving the manager incentive compensation to induce effort; instead, the firm should set $\alpha = \beta = 0$ to guarantee optimal project selection. The condition guarantees that effort is sufficiently important for at least one of α or β to be strictly positive.

Combining this with the second condition in part (ii) is sufficient to guarantee $\alpha^* > 0$. The latter is a technical condition to rule out a boundary case. If b is high, then effort is effective in bankruptcy; if p is low, bankruptcy is likely. Both factors mean that effort has a high expected productivity in bankruptcy; since debt is sensitive to the bankruptcy payoff, the optimal β is high. The disadvantage of high β is that it leads to excessive conservatism in project selection, and so $\alpha > 0$ is typically optimal to counterbalance this. However, if β is sufficiently high that $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} > 1$, then we are at the boundary case where S is always selected. The cutoff V_{GR}^* exceeds V_{GRH}, and so R is never chosen. Even if α increases by a small amount, the cutoff remains above V_{GRH} and so project selection is unchanged. The second condition in part (ii) places an upper bound on b and $(1-p)$ to rule this out.

Part (iii) gives a sufficient condition for $\beta^* > 0$, and thus pure equity compensation to be suboptimal. The condition is more likely to be satisfied if $V_{BS} - V_{BR}$ and b are high, and p is low. $V_{BS} - V_{BR}$ measures the manager's ability to destroy liquidation value by inefficiently choosing R, and thus the magnitude of the asset substitution effect. If the firm has few tangible assets, creditors recover very little regardless of the severity of liquidation: both V_{BS} and V_{BR} are close to zero. If the firm has illiquid, tangible assets (such as buildings) that cannot be eroded by risk-taking, both V_{BS} and V_{BR} are high. In both cases, $V_{BS} - V_{BR}$ is low: there are few gains from making the manager sensitive to the liquidation value, since he has little effect on it. On the other hand, if R reduces liquidation value (such as an advertising campaign, which transforms tangible cash into an intangible asset), $V_{BS} - V_{BR}$ is high and so the optimal level of inside debt is strictly positive. Similarly, if effort is effective at improving liquidation value (b is high), then debt is effective at increasing effort and so again inside debt is justified. If p falls, bankruptcy becomes likelier. This increases the severity of the asset substitution issue; it also augments the effectiveness of debt in inducing effort, because debt is sensitive to the firm's value in bankruptcy which is enhanced by effort. Both factors lead to $\beta^* > 0$.

In sum, the previous literature's justification of exclusively equity-linked incentives is warranted for firms where the agency costs of debt are low and effort considerations are first-order, such as start-ups with high growth opportunities. However, inside debt is desirable in companies with a significant risk of bankruptcy (low p), where the investment decision affects liquidation value (high $V_{BS} - V_{BR}$), and where effort can improve liquidation value (high b). One example is LBOs, which are frequently undertaken in mature firms where the main agency problem is excessive investment. Indeed, the private equity firm can be considered the "manager" in LBOs, given its close involvement in operations, and typically holds strips of debt and equity to minimize conflicts.

If the conditions in parts (ii) and (iii) are satisfied, then we have interior solutions for α^* and β^*. This permits comparative statics, which are given in parts (iv)-(vi). The intuition for part (iv) is standard. Parts (v) and (vi) consider the optimal ratio of debt to equity, $k = \beta/\alpha$. The optimal k is a trade-off between its differential effects on project selection and effort. For project selection, $k = 1$ (i.e. $\beta = \alpha$) is optimal as then $V_{GR}^* = V_{GR}^{**}$. However, effort considerations may cause the optimal k to deviate from 1. To illustrate this, define "output" as

$$\eta = p g e_g + (1-p) b e_b - \alpha E_0 - \beta D_0,$$

i.e. the contribution to firm value provided by effort, minus the manager's pay. We prove in the Appendix that

$$\begin{aligned} \frac{\partial \eta}{\partial k} &= 2p^2 g^2 \alpha \frac{\partial \alpha}{\partial k} + 2(1-p)^2 b^2 k \alpha \frac{\partial (k\alpha)}{\partial k}, \\ \frac{\partial \alpha}{\partial k} &< \frac{\partial (k\alpha)}{\partial k}. \end{aligned} \tag{2.11}$$

If effort has a high expected productivity in solvency (either because solvency is likely (p is high) or effort is particularly effective in solvency states (g is high relative to b)), then equity is more effective than debt in inducing effort, since it is sensitive to the payoff in solvency. Thus, reducing debt and increasing equity augments effort; indeed, inspecting (2.11) shows that $\frac{\partial \eta}{\partial k} < 0$ if p is sufficiently high and g is sufficiently larger than b.[12] The Appendix proves that, if and only if $\frac{\partial \eta}{\partial k}|_{k=k^*} < 0$, $k^* < 1$ and so the contract involves an equity bias ($\beta^* < \alpha^*$). The condition in part (v), $pF > 2(1-p)b^2$, is a sufficient condition for $\frac{\partial \eta}{\partial k}|_{k=k^*} < 0$ and is indeed satisfied if p is high and b is low. Even though increasing k towards 1 would improve project selection, it would also reduce output, and so the optimal k is less than 1. The actual value of k^* is a trade-off between the positive effect on project selection and the negative effect on effort, and thus depends on the magnitude of the two agency problems. If $V_{BS} - V_{BR}$ is high, asset substitution is relatively important and so k^* is closer to the level of 1 that optimizes project selection. If g rises, the benefits from effort are more concentrated in solvency and so equity is more effective at inducing effort, reducing k^*. An increase in p augments the potency of equity in inducing effort and reduces the severity of asset substitution; both factors reduce k^*.

Conversely, if p is low or b is high relative to g, then effort has a high expected productivity in liquidation and so debt is more effective than equity in inducing effort, since it is sensitive to the liquidation payoff. If and only if $\frac{\partial \eta}{\partial k}|_{k=k^*} > 0$, we have $k^* > 1$ and the contract involves a debt bias ($\beta^* > \alpha^*$). The condition in part (vi) is a sufficient condition for $\frac{\partial \eta}{\partial k}|_{k=k^*} > 0$, and is indeed satisfied if g and p are low. Even though increasing k above 1 leads to excessive conservatism in project selection, it also improves output and so the optimal k exceeds 1. This debt bias contrasts the traditional view that insufficient effort results from the "agency costs of equity", i.e. raising too much outside equity leaves the manager with too few shares. For a firm close to bankruptcy, effort may be more efficiently induced by giving the manager more debt. For the comparative statics, the effect of $V_{BS} - V_{BR}$ is reversed: if it increases, project selection becomes more important and so k should be closer to 1; since $k^* > 1$, this involves a *reduction* in k^*. A rise in F increases the importance of asset substitution and has the same effect. By contrast, an increase in b augments the importance of effort and so raises k^*.

While parts (i) to (vi) of Proposition 2.3.2 consider the optimal compensation scheme, part (vii) addresses the effect of the compensation scheme on security prices. An increase in β leads to greater conservatism in project selection and thus increases (decreases) the value of debt (equity).

[12]Habib and Johnsen (2000) also feature the idea that equity induces an agent to improve firm value in solvency, and debt induces an agent to improve firm value in bankruptcy. They consider a different setting where equity (debt) is given to outside investors rather than the manager, to induce them to credibly assess the firm's value in its primary (secondary) use.

2.4 Empirical Implications

The model generates a number of empirical predictions involving inside debt as both an independent and dependent variable, i.e. implications for the effects of inside debt holdings and the firm characteristics that affect the optimal debt level. We outline the predictions below. Existing evidence on pensions appears to be consistent with the first five; the other predictions are unexplored and thus may be fruitful topics for future research.

The "big picture" prediction is that inside debt should be used in executive compensation, whereas prior theories do not advocate debt. Indeed, Bebchuk and Jackson (2005), Sundaram and Yermack (2007), and Gerakos (2007) document the extensive use of pensions. While pensions for rank-and-file employees are typically insured by the Pension Benefit Guaranty Corporation and thus insensitive to bankruptcy, executive pensions typically substantially exceed the maximum insured amount. In a bankruptcy, they represent unsecured, unfunded debt claims with equal priority to other unsecured creditors.[13] Thus, pensions constitute inside debt. Moreover, Sundaram and Yermack find the percentage debt stake exceeds the percentage equity stake for 13% of CEOs. This finding represents an even sharper disparity with theories that advocate only equity, whereas this paper shows that a debt bias is sometimes optimal. Moreover, current evidence understates the extent of inside debt as it typically focuses on executive pensions and ignores deferred compensation, owing to data limitations thus far. Section 409a of the Internal Revenue Code has recently increased the reporting requirements for deferred compensation, and it would be useful to test the predictions below using the CEO's total inside debt holdings; Wei and Yermack (2009) is one such paper that does this. By contrast, there appears to be very little evidence for the fixed bonuses advocated by previous theories of stockholder-bondholder conflicts. Murphy's (1999) survey documents that bonuses in practice are instead typically increasing in equity value (up to an upper limit), and thus augment risk-shifting tendencies.

In addition to this "high-level" prediction on the existence of debt, there are a number of "detail-level" comparative statics predictions. In reality, g is likely to be significantly higher than b and so $k^* < 1$; indeed, empirically, $\alpha > \beta$ for 87% of CEOs, so we use part (v) of Proposition 2.3.2 to form the empirical predictions. From (2.10), an increase in k augments the "cutoff" V_{GR}^* for the manager to choose the risky

[13] See Sundaram and Yermack (2007) and Gerakos (2007) for further detail on the priority of pensions in bankruptcy.

project. Therefore, firm risk should decrease with the manager's personal leverage (i.e. his ratio of inside debt to equity). This is indeed found by Sundaram and Yermack, using distance to default as a measure of firm risk. Similarly, Gerakos finds that CEO pension holdings are associated with higher credit ratings.

Empiricists sometimes measure the manager's personal leverage not using percentages of debt and equity ($\frac{\beta}{\alpha}$) but using dollar values ($\frac{\beta}{\alpha}\frac{D_0}{E_0}$). Using this measure, the model predicts a positive relationship between dollar personal leverage and firm leverage ($\frac{D_0}{D_0+E_0}$) for two reasons. First, an increase in firm leverage $\left(\frac{D_0}{D_0+E_0}\right)$ mechanically increases personal leverage through augmenting the second term. Second, changes in the underlying parameters can jointly increase both firm and personal leverage. From Proposition 2.3.2, an increase in b and a decrease in g both augment k. From Equations (B.1) and (B.2) in the Appendix, these changes also increase D_0 and reduce E_0. Indeed, Sundaram and Yermack find a strong correlation between firm leverage and personal leverage.

In firms with growth options, the effort decision is first-order (g is high). Consequently, pension entitlements should fall, as found by Gerakos. As with the leverage association, this relationship is difficult to reconcile with tax or stealth compensation justifications for pensions, but consistent with efficient contracting. Indeed, Gerakos concludes that contracting variables explain a greater proportion of pension levels than do measures of CEO power.

Turning from the determinants of the compensation scheme to its effects, part (vii) of Proposition 2.3.2 predicts that the value of debt (equity) should rise (fall) with inside debt holdings. SEC regulations mandated disclosure of CEO's inside debt holdings (both pensions and deferred compensation) in Spring 2007. Wei and Yermack (2009) find that firms that disclosed large inside debt holdings indeed experienced increases in bond prices and decreases in equity prices, and that post-disclosure bond yields are significantly positively related to k.

We now move to untested predictions. In Section 2, Case 1 (2) predicts that inside debt is decreasing (increasing) in private benefits. Case 2 depicts a highly levered firm for which liquidation is very likely, and so Case 1 likely applies to the majority of firms. Sundaram and Yermack find that personal leverage is significantly increasing in firm age, which is consistent since the present value of future salary is lower for older managers closer to retirement. Note that this link is not automatically mechanical – while pension benefits naturally increase over time with the CEO's tenure, the same is true for equity compensation (e.g.

Gibbons and Murphy, 1992.) However, sharper tests of this prediction may be possible with measures of private benefits that do not depend on CEO age.

The last implication comes from Proposition 2.3.2, which predicts that inside debt is most valuable where the manager has greatest effect on liquidation values. Note that this is different from raw asset tangibility: if assets are highly intangible (tangible), liquidation values are low (high) regardless of the manager's actions. A possible proxy could be the intensity of covenants.

It is important to note some caveats with interpreting recent pensions findings as being fully consistent with the model and thus evidence that real-life practices are optimal. In many firms, pensions are sufficiently large, and have sufficiently similar payoffs to debt, that they fulfill the role of inside debt advocated by the model and explain why executives do not need to hold actual debt securities in addition. However, these conditions may not be fulfilled in certain circumstances, in which case there may be an argument for supplementing pensions with actual debt, as advocated by Bebchuk and Spamann (2009) in a normative proposal for compensation reform.

First, existing studies are focused on large firms in the U.S. It is not clear whether these findings are representative of all firms, or of firms overseas. While Sundaram and Yermack find that the CEO's debt-equity mix is independent of firm size, potentially implying their results may also apply to smaller firms outside their sample, this has yet to be directly shown. Further research is necessary to investigate the generality of recent results.

Second, the model illustrates that the payoff of a pension has to be very similar to debt for it to be effective: small departures may lead to pensions either not affecting or exacerbating the issue. If debt is secured, pensions are junior and thus similar to the bonus J; they may therefore encourage risk-shifting (Proposition 3). In other cases, the payoff may be close to risk-free and thus pensions do not affect managerial incentives. Executives can put pension fund assets into a "secular" trust fund, ring-fenced from the reach of creditors, or a "springing" trust which converts into a secular trust upon trigger events, such as a credit downgrade. Note, however, that these trusts are very rare. The CEOs of Delta Airlines and AMR (the parent of American Airlines) lost their jobs after such trusts were disclosed (Sundaram and Yermack, 2007.)

Finally, debt securities may have a role even in companies where pensions are currently sufficient to mitigate risk-shifting. Such firms may have low agency costs of debt because they are addressing them by reducing leverage or inside equity.[14] Since these measures are costly, the importance of the agency costs of debt in practice cannot be ascertained solely by looking at actual cases of risk-shifting. Debt grants may allow leverage or equity to increase, and thus be a less costly solution.

2.5 Conclusion

The simplest theory of executive compensation would advocate aligning the manager with firm value. Since empiricists have long believed that managers are compensated exclusively with cash and equity in practice, a number of theory papers rationalize such a scheme. However, recent research has shown that debt-like instruments such as pensions are in fact substantial components of executive compensation. Debt is critically different from other instruments as it is sensitive to the value of assets in bankruptcy. These findings suggest the need for new theories to explain why and when inside debt has a role in efficient compensation, and how much debt should be used.

This paper is a first step in this direction. Inside debt can be a more effective solution to creditor expropriation than salaries, bonuses, reputation and private benefits, owing to its sensitivity to liquidation value. When equity compensation is endogenized via an effort decision, the optimal level of inside debt is typically strictly positive and so pure equity compensation is inefficient. However, contrary to intuition, it typically does not equal the fraction of inside equity due to a trade-off between effort and project selection. An equity bias is usually optimal; the manager's debt-to-equity ratio is increasing in his effect on the liquidation value and the probability of bankruptcy, and decreasing in growth opportunities.

The model generates a number of empirical predictions, many of which appear to be supported with existing findings. However, since data on debt compensation has only recently become available, there are a number of untested predictions that may be interesting topics for future empirical research. In terms of future theoretical directions, this paper has derived conditions under which inside debt is superior to bonuses, and analyzed the optimal relative proportions of debt and equity when compensation

[14] There is a widely documented negative relationship between α and leverage (Friend and Lang, 1988; Agrawal and Nagarajan, 1990; Ortiz-Molina, 2007). One interpretation is that firms for which high debt is optimal are reducing the manager's equity stake to attenuate risk-shifting (as predicted by John and John, 1993), or firms that require high α to induce effort are under-leveraging for the same reason.

comprises of these instruments. It would be fruitful to study the conditions under which debt is optimal in a general contract design setting.

Chapter 3

Incentive Contracting under Ambiguity-Aversion

3.1 Introduction

Standard models on executive compensation usually assume that the distribution on future firm value is known and study the optimal CEO pay under this assumption. In this paper, I study executive compensation in the case where there are multiple possible distributions on firm value. For example, due to the lack of information, this is quite possible for new firms and new industries to have some ambiguous information on their prospects.

Following the literature on ambiguity-aversion, I assume that managers are risk-averse and ambiguity-averse. When facing ambiguity, they will derive their utility from the worst case. The behavior of ambiguity-aversion is consistent with some experimental evidence, e.g., the most notable example Ellsberg Paradox (Ellsberg (1961)) and the portfolio choice experiments (Ahn, Choi, Gale, and Kariv (2009)). The applications of ambiguity-aversion have been well studied in asset pricing. Gilboa and Schmeidler (1989) provide an axiomatic basis for a "max-min" preference that is used to model ambiguity-aversion. Epstein and Wang (1994) show how such an approach could be incorporated into a dynamic asset pricing model. Moreover, people have successfully applied ambiguity-aversion to explain many empirical puzzles in finance. For example, it can be applied to explain the puzzles on excess volatility, negative skewness and excess kurtosis of stock market returns (Epstein and Schneider (2008), Illeditsch (2010)). Easley and O'hara (2009) use ambiguity-aversion to study the participation puzzle in stock markets and the implications of regulation on asset prices. Easley and O'hara (2010) show that ambiguity-aversion can explain sudden market freezes. In this paper, I apply the "max-min" preference to an ambiguity-averse manager and study how this will affect the structure of an optimal contract.

I first consider the case where firm risk is ambiguous. I assume that firm risk could be either high or low. Since a rise of firm risk only increases the variance of stock, but has no effect on the mean value of stock, a risk-averse and ambiguity-averse manager will always perceive a high risk when she owns stock. This requires shareholders to pay a high risk premium to managers. On the other hand, a rise of firm risk increases both the mean and variance of options. While an increase in the variance of options decreases the manager's utility, an increase in the mean of options also has a positive effect on her utility. If the manager is not too risk-averse, she could perceive a low risk in the worst case. As a result, shareholders only need to pay a low risk premium. If the savings from risk premium are large enough (which requires that the manager's risk-aversion is not too small), options can be less costly than stock to induce managers to exert effort. Thus, the model provides an explanation for the use of options. The calibration of a standard principal-agent model shows that options should not be a part of an optimal contract (Dittmann and Maug (2007)). Thus it is puzzling that why options are used so widely in practice. A general argument is based on the risk-taking incentives. That is, options can encourage managers to take more risk, through which the mean of firm value will also be raised up. Since shareholders can diversify their risk in markets, they will be happy with risk-taking to some extent. Thus, it can justify the use of options in executive compensation (Feltham and Wu (2001), Dittmann and Yu (2008)). However, several papers (Carpenter (2000), Ross (2004)) notice that it is not always true that giving options to agents will make them more willing to take risk. Bettis, Bizjak, and Lemmon (2005) show that managers exercise their options earlier when volatility increases. Hayes, Lemmon and Qiu (2010) also find little evidence to support the view that the convexity inherent in option-based compensation is used to reduce risk-related agency problems between managers and shareholders. Therefore, this paper provides another way for people to look at the role of options.

Secondly, I show that when the target effort goes up, it becomes more difficult to induce managers to perceive a low risk. This is because inducing a manager to exert higher efforts requires granting her higher incentives. Since managers are risk-averse, they tend to perceive a higher risk when they are granted more incentives. Therefore, at some cut-off point of efforts, the risk premium paid to managers will move from low to high. As a result, shareholders will be reluctant to induce managers to exert higher efforts above this point, even if there are some improvements in managers' ability or in firms' investment opportunities. This can help to explain the use of benchmark pay documented by empirical studies (Bizjak, Lemmon, and Naveen (2008), Faulkender and Yang (2010)). The reason is: for firms with some similarities, although these firms could differ a little bit in some dimensions, the cost arising from ambiguity will align their target efforts and make their managers' contracts look similar.

Lastly, I show that when there is ambiguity in the mean of firm value, options can still be more efficient than stock. I prove that a manager always perceives a low mean of firm value with stock, but could perceive a high mean of firm value with options. The perception of a high firm value not only increases the subjective value of options to the manager, but also reduces the number of options required to induce effort. These two benefits can make options less costly than stock. I then consider the case where ambiguity comes from market returns. I compare four different instruments: non-indexed stock, indexed stock, non-indexed options and indexed options, [1] and prove that non-indexed options can be the most efficient way to induce effort. Thus, tying CEO pay to markets can be optimal. This helps to explain the pay-for-luck puzzle (Bertrand and Mullainathan (2001)).

3.2 Model

3.2.1 Basic Set-up without Ambiguity

Consider a one-period principal-agent model in which the principal are shareholders and the agent is a manager. At time $t = 0$, the manager exerts an effort a, which is unobservable to shareholders. I assume that a feasible set of actions is a bounded interval, i.e. $a \in [\underline{a}, \bar{a}]$. The upper bound $\bar{a}$ reflects the fact that there is a limit to the number of positive NPV projects that the manager can undertake or a limit to the number of hours that the manager can spend. Then at time $t = 1$, firm value is given by $x = a + l + \epsilon$,

[1] Indexed instruments mean that market returns are excluded from their payoffs.

where ϵ is a noise term, which is normally distributed with mean zero and variance σ^2. The noise is realized after the manager exerting effort. l is a constant term which can be considered as the base value of firm or the expected market return (also called luck). [2]

Then the mean and variance of an option with exercise price k are

$$m_k = \sigma[\phi - (1-\Phi)\eta], \tag{3.1}$$

$$\sigma_k^2 = \sigma^2[(1-\Phi-\phi^2) - \phi(2\Phi-1)\eta + \Phi(1-\Phi)\eta^2]. \tag{3.2}$$

where $\eta = \frac{k-a-l}{\sigma}$, and

$$\phi = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\eta^2\right),$$

$$\Phi = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right)dt$$

Lemma 3.2.1. *m_k and σ_k^2 are increasing in a, l and σ, and decreasing in k.*

The main objective of the paper is to compare linear contracts with non-linear contracts in the presence of ambiguity, so I consider a stock-based contract as a linear contract and let $m_{-\infty} = a + l$ and $\sigma^2_{-\infty} = \sigma^2$ to be the mean and variance of stock.

Since the manager's action is unobservable, shareholders can only tie the manager's pay to the outcome of firm value to implement the target effort a. At time $t = 0$, the contract is signed. For tractability, I restrict contracts to be either stock-based or option-based. An option-based contract consists of a base salary α and β_k options with exercise price k. The manager's payoff with the option-based contract is

$$w = \alpha + \beta_k \max(x-k, 0).$$

For a stock-based contract with a base salary α and $\beta_{-\infty}$ shares, the payoff is

$$w = \alpha + \beta_{-\infty} x.$$

[2]In Section 3.2.3, I will allow l to be an ambiguous term.

The manager is risk-averse and her expected utility is given by the following mean-variance preference:

$$E[u] = E[w] - \frac{1}{2}\lambda Var[w] - g(a) \tag{3.3}$$

where $g(a) = \frac{1}{2}a^2$ is the cost function if the manager exerts effort a and λ is the measure for the manager's risk-aversion.

The shareholders are risk-neutral and want to maximize the expected firm value net of CEO pay:

$$\max_{\alpha,\beta_k,\hat{a}} E[x - w] = \hat{a} + l - \alpha - \beta_k m_k$$

s.t. IR constraint: $E[u(a)] \geq 0$ [3] and IC constraint: $a \in \arg\max_{\hat{a}} E[u(\hat{a})]$.

Grossman and Hart (1983) provide a two-stage approach to solve the principal-agent problem. The first stage is to find the minimum cost $c(a)$ for implementing the target effort a. Then in the second step, the principal will search for the optimal target effort over all possible pairs of effort a and minimized cost $c(a)$. In this paper, I focus on the first stage and compare the cost of a stock-based contract to implement the target effort with that of an option-based contract. [4]

First-order approach

The managers' expected utility is

$$\begin{aligned} E[u] &= E[w] - \frac{1}{2}\lambda Var[w] - g(a) \\ &= \alpha + \beta_k m_k - \frac{1}{2}\lambda \beta_k^2 \sigma_k^2 - \frac{1}{2}a^2. \end{aligned}$$

In general, the above utility function is not concave in the manager's effort a, hence it might be invalid to replace the IC constraint by the first-order condition. In Appendix, I provide some conditions under which the use of the first-order condition is valid. Here I assume that those conditions hold and replace

[3] Here I normalize the reservation utility to be 0.

[4] Edmans and Gabaix (2010) prove a maximum effort principle, which says that if the productivity of a manager's effort is sufficiently large, then the maximum effort $\bar{a}$ is optimal. Therefore, we can skip the second step if we apply the maximum effort principle.

the IC constraint by the first-order condition:

$$\beta_k \frac{\partial m_k}{\partial a} - \frac{1}{2}\lambda\beta_k^2 \frac{\partial \sigma_k^2}{\partial a} - a = 0. \tag{3.4}$$

Note that α does not affect the manager's effort choice, so the IR constraint must be binding:

$$\alpha + \beta_k m_k - \frac{1}{2}\lambda\beta_k^2\sigma_k^2 - \frac{1}{2}a^2 = 0$$

Thus the expected CEO pay (i.e., the cost of implementing a) can be rewritten as

$$E[w] = \alpha + \beta_k m_k = \frac{1}{2}\lambda\beta_k^2\sigma_k^2 + \frac{1}{2}a^2. \tag{3.5}$$

Formula (3.5) is the cost of an option-based contract. It consists of two parts: the risk premium paid to managers and the cost of managers' effort. It can also be shown that the cost of a linear contract (i.e., a stock-based contract) corresponds to (3.5) with $k = -\infty$. [5] Since the target effort a is given, $\frac{1}{2}a^2$ is just a constant. Hence I drop $\frac{1}{2}a^2$ from (3.5) and refer to $\frac{1}{2}\lambda\beta_k^2\sigma_k^2$ as the cost of implementing the target effort a. Denote it by

$$c_k = \frac{1}{2}\lambda\beta_k^2\sigma_k^2 \tag{3.6}$$

We can see that when the target effort is fixed, the cost of implementing efforts is equal to the risk premium paid to managers. If the manager is risk-neutral (i.e., $\lambda = 0$), then there is no difference between stocked-based contracts and option-based contracts. However, if the manager is risk-averse, the following proposition shows that the risk premium of a stock-based contract is smaller.

Proposition 3.2.2. *The minimum number of options required to induce the target effort a is*

$$\beta_k = \frac{(1-\Phi) - \sqrt{\Delta}}{2\lambda\Phi m_k} = \frac{2a}{(1-\Phi) + \sqrt{\Delta}} \geq a,$$

where $\Delta = (1-\Phi)^2 - 4a\lambda\Phi m_k$. β_k increases in the manager's risk aversion λ and the option's exercise price k.

[5] That is why I use $m_{-\infty}$ and $\sigma_{-\infty}^2$ to denote the mean and variance of stock.

Moreover, the total cost of inducing effort a is

$$c_k = \frac{1}{2}\lambda\beta_k^2\sigma_k^2 = \frac{2\lambda a^2\sigma^2 y_1}{(1+\sqrt{1-4a\lambda\sigma y_2})^2},$$

where

$$\begin{aligned} y_1 &= \frac{(1-\Phi-\phi^2)-\phi(2\Phi-1)\eta+\Phi(1-\Phi)\eta^2}{(1-\Phi)^2}, \\ y_2 &= \frac{\Phi[\phi-(1-\Phi)\eta]}{(1-\Phi)^2}. \end{aligned}$$

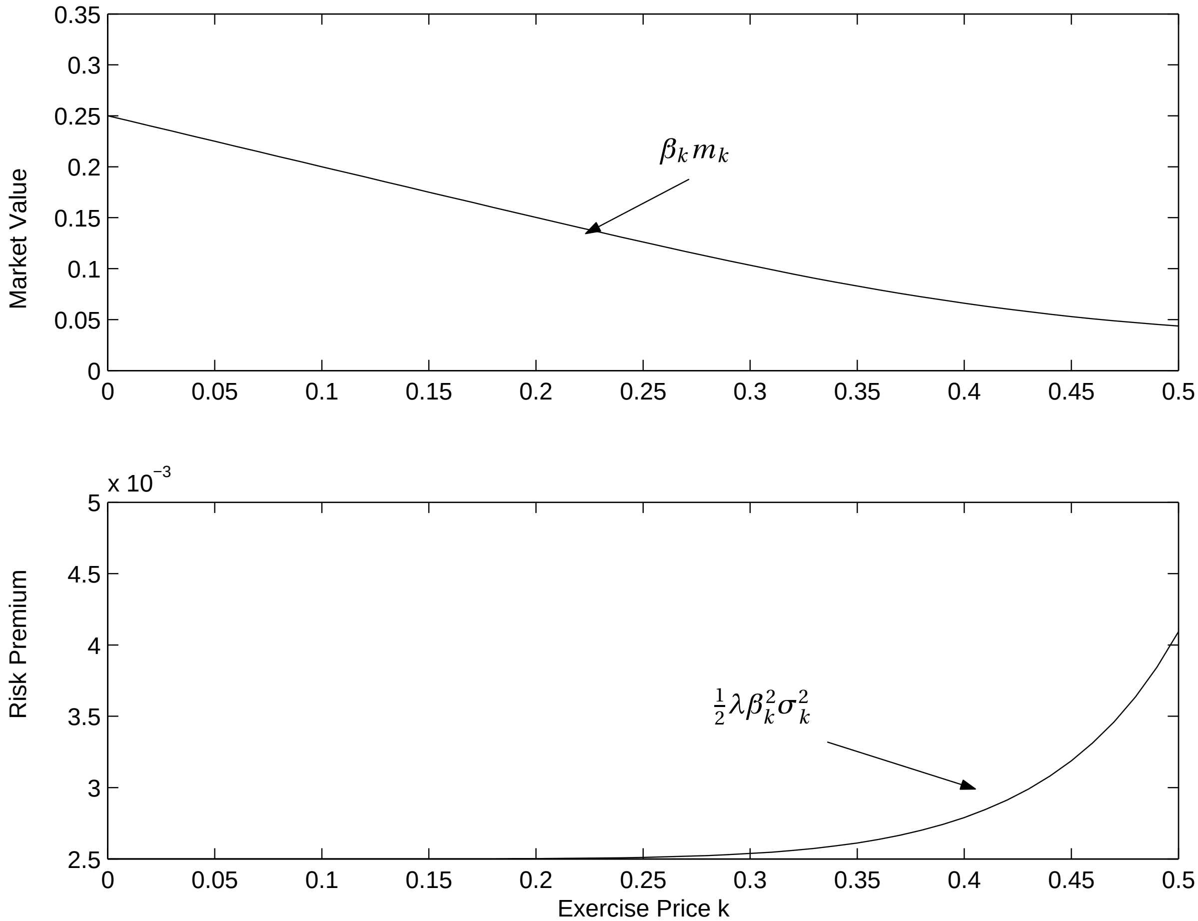
The cost c_k increases in k, hence stock-based contracts are less costly than option-based contracts;

Proof. See Appendix. □

When people compare stock to options, there are two perspectives. On one hand, options are expensive since they are riskier, e.g., \$100 of options are worth less to an undiversified risk-averse manager than \$100 of stock; but on the other hand, options usually can provide more incentives to induce effort than stock. If the base salary is rigid, the cost of a stock-based or an option-based contract equals to the market value of stock or options. In this case, options are less costly to provide the same level of incentives. This is illustrated in the top graph of Figure 3.1, which shows that the market value is decreasing in exercise price. However, if the base salary is adjustable, since stock usually provides more subjective values to a manager than options, a stock-based contract can save more on the base salary than an option-based contract. As a result, we have shown above that the cost of a contract equals to the risk premium of the contract. Now stock dominates options as shown in Proposition 3.2.2 and the bottom graph of Figure 3.1. The intuition here is the same as in Dittmann and Maug (2007).

In fact, the result in Proposition 3.2.2 is also consistent with the linearity result of Holmstrom and Milgrom (1987). In Holmstrom and Milgrom (1987), it is a continuous time model over the time interval $[0,1]$. They make the following assumptions in the model: 1) the agent is risk-averse with an exponential utility; 2) the agent chooses efforts continuously over the time interval $[0,1]$ to control the drift vector of a Brownian motion $\{x(\tau): 0 \le \tau \le 1\}$; 3) the agent can observe the accumulated performance before acting. Then they show that an optimal contract specifies that the agent will choose a constant effort a

Figure 3.1: Market Value and Risk Premium of An Option-Based Contract



This picture describes a numerical example showing that how the market value and the risk premium of an option-based contract change w.r.t the change of exercise price k. In the example, we set that the target effort $a = 0.5$; $l = 0$; $\sigma = 0.1$; $\lambda = 2$.

over time, regardless of the history at any time τ, and the agent's wage has the form $w = \alpha + \beta x(1)$, which is a linear function of the final outcome regardless of any intermediate outcomes. Then we can see that although the two models are different (one is discrete, one is continuous), they coincide at time 0 and 1. At time 0 the manager exerts an effort a which determines the mean of firm value; at $t = 1$ the manager's payoff is realized and only depends on the firm value at time 1. If the manager with an exponential utility function is granted a linear contract, then the certainty-equivalent of her utility is exactly given by (3.3). Hence if the result in Proposition 3.2.2 is not true, then we should be able to provide a counter example to the Holmstrom and Milgrom (1987)'s linearity result. Thus the result in Proposition 3.2.2 is not surprising from this point of view.

3.2.2 Ambiguity in Firm Risk

In the basic model, firm value x is normally distributed with mean $a + l$ and variance σ^2. A manager infers her expected utility from this objective distribution. However, a question arises that does the manager know firm risk σ^2 exactly? Especially, for new firms or new industries, since the available information is little, can the manager evaluate future risk correctly? If not, then there may be some ambiguous information on firm risk. Then instead of one number for sure, she may only know a set of possible variances $\{\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2\}$ for firm risk. In this section, I study how the presence of this ambiguity affect managers' contracts.

Following the literature on ambiguity-aversion, I assume that the manager is both risk-averse and ambiguity-averse. When facing ambiguity, she will first minimize her utility under possible distributions (i.e. her utility is inferred in the worst case) and then choose an effort to maximize her utility that has been minimized under the possible distributions. That is, the manager's objective is

$$\max_{\hat{a}} \min_{\sigma^2 \in \{\sigma_1^2, \cdots, \sigma_n^2\}} E[w] - \frac{1}{2}\lambda Var[w] - g(\hat{a}) \tag{3.7}$$

Gilboa and Schmeidler (1989) provide an axiomatic basis for this "max-min" preference. For simplicity, I assume that $n = 2$ (i.e. there are only two possible values for firm risk, low risk or high risk) and $\sigma_1^2 < \sigma_2^2$.

For the shareholders, I assume that their perception of risk is σ_s, where σ_s can be σ_1 or σ_2. [6] In the following proposition, we can see that either $\sigma_s = \sigma_1$ or $\sigma_s = \sigma_2$ does not affect the results. The main forces driving the results are the assumption that the manager is ambiguity-averse.

Given the target effort a, the shareholders' objective is to minimize $\alpha + \beta_k m_k(a, \sigma_s)$. Suppose the manager's perception of risk is σ_m. Then to induce the manager to exert effort a, we must have $\beta_k = \beta_k(a, \sigma_m)$. Also as before we must have the binding IR constraint (because α does not affect the manager's effort decision), that is

$$\alpha + \beta_k(a, \sigma_m) m_k(a, \sigma_m) - \frac{1}{2}\lambda \beta_k^2(a, \sigma_m)\sigma_k^2(a, \sigma_m) - \frac{1}{2}a^2 = 0$$

Then we have that the expected CEO pay from shareholders' perspective is

$$\alpha + \beta_k(a, \sigma_m) m_k(a, \sigma_s) = \frac{1}{2}\lambda \beta_k^2(a, \sigma_m)\sigma_k^2(a, \sigma_m) + \beta_k(a, \sigma_m)[m_k(a, \sigma_s) - m_k(a, \sigma_m)] + \frac{1}{2}a^2$$

Since $\frac{1}{2}a^2$ is just a constant, dropping it from the expression above, we obtain that the cost of implementing the target effort is $\frac{1}{2}\lambda \beta_k^2(a, \sigma_m)\sigma_k^2(a, \sigma_m) + \beta_k(a, \sigma_m)[m_k(a, \sigma_s) - m_k(a, \sigma_m)]$. I use $\hat{c}_k$ to denote this new cost, i.e.

$$\hat{c}_k(a, \sigma_m) = \frac{1}{2}\lambda \beta_k^2(a, \sigma_m)\sigma_k^2(a, \sigma_m) + \beta_k(a, \sigma_m)[m_k(a, \sigma_s) - m_k(a, \sigma_m)].$$

If the manager's perception of risk is always the same as shareholders', i.e., $\sigma_m = \sigma_s$, then everything is the same as in the case without ambiguity and Proposition 3.2.2 shows that stock-based contracts are less costly. However, because of the convexity of options' payoffs, options can induce managers to perceive a lower risk than stock. In the following lemma, I show that if a manager's risk-aversion λ is not too big, then options will induce the manager to perceive a low risk, which is different from the case of stock. (i.e., the manager perceives a high risk with stock but perceives a low risk with options). The intuition is as follows: the mean and variance of stock are given by $m_{-\infty} = a + l$ and $\sigma^2_{-\infty} = \sigma^2$. So a rise of σ has no effect on the mean value of stock, but increases the variance of stock. A risk-averse manager will always perceive the high risk σ_2^2 when she is granted shares of stock. However, for options, a rise of σ increases

[6] It can be shown that if shareholders are ambiguity-averse, then $\sigma_s = \sigma_2$.

both the mean and variance of options. For a risk-averse manager, while an increase in the variance of options decreases her utility, an increase in the mean value of options also has a positive effect on her utility. So whether the manager perceives a high risk or low risk depends on which effect dominates. If the rise of σ has a larger effect on the mean value of options and the manager is not too risk-averse, then her utility can be minimized at the low risk. In that case, she will perceive the low risk σ_1^2.

Lemma 3.2.3. *Suppose the manager is both risk-averse and ambiguity-averse.*

1) The manager will always perceive a high risk σ_2^2 if she is granted stock;

2) If the manager is granted $\beta_k = \beta_k(a, \sigma_1)$ options with exercise price k, then she will perceive the low risk σ_1^2 and exert the target effort a if

$$\lambda < \frac{2}{\beta_k} \frac{m_k(a, \sigma_2) - m_k(a, \sigma_1)}{\sigma_k^2(a, \sigma_2) - \sigma_k^2(a, \sigma_1)} \tag{3.8}$$

Proof. See Appendix. □

Now let's analyze the effect of changing the manager's perception of firm risk. Recall that under ambiguity, the new cost of inducing the target effort a is

$$\hat{c}_k(a, \sigma_m) = \frac{1}{2}\lambda \beta_k^2(a, \sigma_m)\sigma_k^2(a, \sigma_m) + \beta_k(a, \sigma_m)[m_k(a, \sigma_s) - m_k(a, \sigma_m)].$$

We can see that the new cost $\hat{c}_k(a, \sigma_m)$ consists of two parts: $c_k(a, \sigma_m)$ and $\beta_k(a, \sigma)[m_k(a, \sigma_s) - m_k(a, \sigma_m)]$. The first part is the risk premium paid to the manager for the risk that she is bearing, which is the same as in the previous section; but in the presence of ambiguity, we have a new part in the cost, which is the different valuations of options between shareholders and a manager. Lemma 3.2.3 shows that the manager will always perceive the high risk with a stock-based contract, thus the cost of a stock-based contract is

$$\begin{aligned} \hat{c}_{-\infty}(a, \sigma_m) &= \frac{1}{2}\lambda \beta_{-\infty}^2(a, \sigma_2)\sigma_{-\infty}^2(a, \sigma_2) + \beta_{-\infty}(a, \sigma_2)[m_{-\infty}(a, \sigma_s) - m_{-\infty}(a, \sigma_2)] \\ &= \frac{1}{2}\lambda a^2 \sigma_2^2, \end{aligned}$$

where the second equality follows from the fact that $m_{-\infty}(a, \sigma_s) = m_{-\infty}(a, \sigma_2) = a + l$.

On the other hand, if the manager perceives the low risk with options of exercise price k, then the cost of the option-based contract is

$$\hat{c}_k(a,\sigma_m)=\frac{1}{2}\lambda\beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)+\beta_k(a,\sigma_1)[m_k(a,\sigma_s)-m_k(a,\sigma_1)].$$

Note that the shareholders' valuation of options is increasing in their perception of risk σ_s, thus the cost $\hat{c}_k$ is increasing in σ_s. Hence let us consider the worst case where shareholders perceive the high risk σ_2. If the manager perceives a lower risk with options, then there are two effects on the cost of the option-based contract. First, it increases the second part of the cost $\hat{c}_k$, because shareholders value options more than the manager. Second, perceiving a lower risk by the manager reduces the risk premium paid to her (i.e., the first part of the cost). Thus the total cost of an option-based contract may be larger or smaller than the cost of a stock-based contract, depending on which effect dominates. Intuitively, if σ_1 is small enough and λ is an intermediate value, then it is possible that using options is less costly to induce effort. The reason is that when a manager perceives firm risk, there are two forces - higher risk increases the market value of options, but it also reduces the value of options to the manager because he is risk-averse. The more risk-averse he is, the greater the second effect, and so a very risk-averse manager will perceive a high risk. Thus, we need risk aversion to be small enough to induce the manager to perceive a low risk with options (see Lemma 3.2.3). But, if a manager's risk aversion is too small, risk has little effect on the manager's valuation of her compensation contract (she does not care about it) and so changing the manager's perception of firm risk from high to low does not change the result that stock is optimal. Thus, we need the manager's risk aversion to be large enough, such that a change in the manager's perception of firm risk can reduce the cost of option-based contracts sufficiently large and therefore render options optimal. The following proposition shows that options are less costly than stock for an intermediate value of λ.

Proposition 3.2.4. *If the manager's risk aversion λ satisfies*

$$\frac{2\beta_k(a,\sigma_1)[m_k(a,\sigma_2)-m_k(a,\sigma_1)]}{a^2\sigma_2^2-\beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)}<\lambda<\frac{2}{\beta_k(a,\sigma_1)}\frac{m_k(a,\sigma_2)-m_k(a,\sigma_1)}{\sigma_k^2(a,\sigma_2)-\sigma_k^2(a,\sigma_1)}.$$

then it is less costly to induce effort a by using options with exercise price k than using stock.

Corollary 1. *Given $k < a + l$, if*

$$\frac{2[m_k(a,\sigma_2) - (a + l - k)]}{a\sigma_2^2} < \lambda < \frac{2[m_k(a,\sigma_2) - (a + l - k)]}{a\sigma_k^2(a,\sigma_2)}, \tag{3.9}$$

then there exists a cut-off $\hat{\sigma}(\lambda, k)$ such that when $\sigma_1 < \hat{\sigma}(\lambda, k)$, it is cheaper to induce effort a by using options with exercise price k than using stock.

Remark: since $\sigma_k^2(a,\sigma_2)$ is always less than $\sigma_2^2 = \sigma_{-\infty}^2(a,\sigma_2)$ by Lemma 3.2.1, the condition (3.9) can be satisfied.

An inertia property of compensation contracts under ambiguity

Given the target effort, options can be more efficient than stock because they induce a manager to perceive a low risk, which reduces the risk premium paid to the manager. Then the question is can we always use an option-based contract to induce the manager to perceive a low risk? Intuitively, when the target effort goes up, shareholders need to give a manager more incentives to induce effort. Because the manager is risk-averse, he is more likely to perceive a higher risk when she is granted more incentives. As a result, it becomes more difficult to make the manager perceive a low risk when the target effort increases.

Lemma 3.2.5. *Given two effort levels $a_1 < a_2$, if people can use a contract consisting of stock and options to implement a_2 and keep the manager perceive a low risk, then they can also use a contract with stock and options to implement a_1 and keep the manager perceive the low risk as well.*

Lemma 3.2.5 implies that there could exist a cut-off a_0 such that for any effort level below a_0, shareholders can implement efforts with a contract that keeps the manager to perceive the low risk, In this case, they only need to pay the manager a low risk premium. However, for any effort level above a_0, the incentives to induce the target effort is so high that the manager will perceive the high risk. In this case, a high risk premium is required. This implies that there could be a discontinuous jump in the cost of compensation contracts when the target effort moves from below a_0 to above a_0. Suppose the current target effort is a_0, then even if there are some small improvements in the manager's ability (e.g. the manger gains more experiences or the manager's effort becomes more productive) or in the firm's investment opportunities (e.g. a few more positive NPV projects come out), shareholders may be reluctant to take advantage of these improvements. If they want to take advantage of improvements, they will have to grant the manager more incentives, but granting more incentives will move the manager's perception of risk from low to high. As a result, there will be a large increase in the risk premium that has to be paid to the manager. If this cost is large enough, it will prevent shareholders from increasing the manager's incentives. Therefore the manager's compensation contract will exhibit an inertia property around the point a_0. The inertia property of compensation contracts may help to explain why benchmark pay is so widely used in practice (e.g. Bizjak, Lemmon and Naveen (2008) and Faulkender and Yang (2010)). Standard principal-agent models usually require that contracts should change according to parameters. For example, high productive CEOs or low risk-averse CEOs can be induced to exert effort more easily, so shareholders should grant them more incentives to induce them to exert higher target effort. But in reality, contracts appear to be quite standard and benchmark pay is widely used. Faulkender and Yang (2010) find that CEO compensation of the compensation peers explains CEO compensation at the sample firm beyond traditional pay determinants and proxies for CEO labor market conditions. They also find that size, industry and similar visibility are important determinants of selecting peer firms in benchmarking. For firms in the same industry with similar size or visibility, they probably have similar target efforts for their CEOs, though the target efforts may differ a little bit. According to Lemma 3.2.5, the cost arising from ambiguity would align their target efforts and small changes in parameters would not change the target efforts and the corresponding contracts.

In addition, it also implies that shareholders may be reluctant to grant a manager high incentives, because high incentives induce the manager to perceive a high risk which imposes a high risk premium for the contracts. Thus low CEO incentives (e.g. Jensen and Murphy (1990)) may be desirable.

3.2.3 Ambiguity in Expected Firm Value

In this section, I study the case where there is ambiguity about the mean of firm value. A firm's future performance is affected by market conditions or industrial conditions. However, if the information about the future market or industrial conditions is ambiguous, then the expectation on firm value is also ambiguous. Recall that the mean of firm value is given by $a + l$, where a is the manager's effort and l is a constant which may reflect the base value of firm or the expectation on market returns or industrial returns. To introduce ambiguity in this part, I assume that l is an ambiguous term. People do not know the exact value of l, but only know a set of possible values $\{l_1, \cdots, l_n\}$ for l. To simplify the analysis, I assume that firm risk σ has no ambiguity. In this section, I focus on the case where the mean of firm value is ambiguous.

As in section 3.2.2, I assume that the manager is both risk-averse and ambiguity-averse. His objective is

$$\max_{\hat{a}} \min_{l \in \{l_1, \cdots, l_n\}} E[w] - \frac{1}{2}\lambda Var[w] - g(\hat{a}) = \max_{\hat{a}} \min_{l \in \{l_1, \cdots, l_n\}} \alpha + \beta_k m_k(\hat{a}, l) - \frac{1}{2}\lambda \beta_k^2 \sigma_k^2(\hat{a}, l) - \frac{1}{2}\hat{a}^2.$$

For simplicity, I also assume that $n = 2$ and $l_1 < l_2$.

Suppose l_s is the shareholders' perception of l. [7] As in the previous section, I will show that the shareholders' perception is not crucial to the results in this section. Let l_m be the manager's perception of l. Then to induce the target effort a, the number of options granted to the manager must be $\beta_k = \beta_k(a, l_m)$. We also have the binding IR constraint as follows:

$$\alpha + \beta_k m_k(a, l_m) - \frac{1}{2}\lambda \beta_k^2 \sigma_k^2(a, l_m) - \frac{1}{2}a^2 = 0.$$

Then the cost of the compensation contract from the shareholders' perspective is

$$\begin{aligned} & \alpha + \beta_k m_k(a, l_s) \\ = \ & \frac{1}{2}\lambda \beta_k^2 \sigma_k^2(a, l_m) + \beta_k[m_k(a, l_s) - m_k(a, l_m)] + \frac{1}{2}a^2 \end{aligned}$$

Dropping the constant term $\frac{1}{2}a^2$, I define that

$$\bar{c}_k(a, l_m) = \frac{1}{2}\lambda \beta_k^2 \sigma_k^2(a, l_m) + \beta_k[m_k(a, l_s) - m_k(a, l_m)]$$

is the cost of inducing the target effort a. Similar to the cost $\hat{c}_k$ in section 3.2.2, the cost $\bar{c}_k$ also contains two parts: the first part is the risk premium paid to the manager and the second part $\beta_k[m_k(a, l_s) -$

[7] It can be shown that $l_s = l_1$ for most of cases if shareholders are ambiguity-averse.

$m_k(a, l_m)]$ is the different valuations of the options between the shareholders and the manager.

If the manager is granted shares of stock, then he will always perceive a low l. Since a higher l increases the expected value of stock, but has no effect on the variance of stock, the manager's utility is minimized at low l and thus an ambiguity-averse manager will perceive a low l. However, the situation is different if the manager is granted options. When the manager perceives the return l with options, a higher l increases the market value of options, but it also increases the variance of options, which reduces the value of options to a risk-averse manager. The more risk-averse she is, the greater the second effect, and so a high risk-averse manager would perceive a higher return l_2. In the following lemma, I show that the manager will always perceive a low l with stock, but may perceive a high l with options.

Lemma 3.2.6. *1) A risk-averse and ambiguity-averse manager will always perceive l_1 if she is granted shares of stock;*

2) If she is granted $\beta_k = \beta_k(a, l_2)$ options with exercise price k, then she will perceive l_2 and exert effort a if

$$\lambda > \frac{2}{\beta_k} \frac{m_k(a, l_2) - m_k(a, l_1)}{\sigma_k^2(a, l_2) - \sigma_k^2(a, l_1)} \tag{3.10}$$

Now let us compare the cost of implementing a by using stock and options respectively. With a stock-based contract, the manager will always perceive a low l, hence the cost is

$$\begin{aligned} \bar{c}_{-\infty}(a, l_1) &= \frac{1}{2}\lambda\beta_{-\infty}^2\sigma_0^2(a, l_1) + \beta_{-\infty}[m_{-\infty}(a, l_s) - m_{-\infty}(a, l_1)] \\ &= \frac{1}{2}\lambda a^2\sigma^2 + a[m_{-\infty}(a, l_s) - m_{-\infty}(a, l_1)]. \end{aligned} \tag{3.11}$$

For options, suppose the manager's perception of l is l_2 (i.e. the condition (3.10) is satisfied), then the cost will be

$$\bar{c}_k(a, l_2) = \frac{1}{2}\lambda\beta_k(a, l_2)^2\sigma_k^2(a, l_2) + \beta_k(a, l_2)[m_k(a, l_s) - m_k(a, l_2)].$$

Although the perception of a high l increases the variance of each option (since $\sigma_k^2(a, l)$ is increasing in l), there are two benefits from the perception of a high l. First, if the manager perceives a high l, the options are more valuable to her and it becomes easier to induce the target effort. The reason is that suppose $a + l_1 < k < a + l_2$. If the manager perceives l_1, since $a + l_1 < k$, even she exerts effort a, the options are still more likely to expire with value zero. So she will be less willing to exert effort a. But if the manager perceives l_2, then she will be more likely to exert effort a, because $a + l_2 > k$ and she will think that the options are more likely to have a positive payoff if she exerts effort a. Therefore shareholders only need to grant less options to the manager (i.e. $\beta_k(a,l)$ is decreasing in l) to implement the target effort. As a result, although the perception of a high l increases the variance of each option, the total risk premium is reduced due to the reduction in the number of options granted to the manager. [8] Second, the perception of a high l also increases the subjective value of options to a manager. Thus the base salary paid to the manager is not as high as in the base case where there is no ambiguity. If the two benefits are large enough, then the cost of an option-based contract can be smaller. The following proposition proves the result.

Proposition 3.2.7. *If the manager's risk aversion λ satisfies*

$$\frac{2}{\beta_k(a,l_2)} \frac{m_k(a,l_2) - m_k(a,l_1)}{\sigma_k^2(a,l_2) - \sigma_k^2(a,l_1)} < \lambda < \frac{2a[m_k(a,l_2) - m_k(a,l_1)]}{\beta_k^2(a,l_2)\sigma_k^2(a,l_2) - a^2\sigma^2}, \tag{3.12}$$

then it is less costly to implement effort a by using options with exercise price k than by using stock.

Pay for luck

The model in this section is also helpful to explain the pay-for-luck puzzle. Bertrand and Mullainathan (2001) documented that CEOs are rewarded for general market upswings beyond CEO's control, i.e., paid for luck. This is puzzling because the market return is beyond managers' control and is also volatile, but since the manager is risk-averse, tying her pay to luck seems only increasing the risk premium required to compensate the manager.

However if market return is ambiguous and managers are ambiguous-averse, then we can apply the result above to explain the puzzle. Let us consider l as the expectation of market return. I assume that the market return is normally distributed with mean value l and variance σ_l^2, where the variance is given, but the mean is ambiguous with two possible values l_1, l_2. Also I assume that market return is uncorrelated with idiosyncratic risk ϵ. The variance of idiosyncratic risk is $\sigma^2 - \sigma_l^2$, and the total variance of firm value

[8] It is easy to apply the result in Proposition 3.2.2 to show that $\frac{1}{2}\lambda\beta_k(a,l)^2\sigma_k^2(a,l)$ is decreasing in l.

is still σ^2. The following lemma shows that to implement the target effort a, the cost of using indexed stock (i.e., no pay for luck) is lower than the cost of using non-indexed stock or indexed options.

Lemma 3.2.8. *The cost of implementing the target effort a is*

1) $\frac{1}{2}\lambda a^2\sigma^2 + a[m_{-\infty}(a,l_s) - m_{-\infty}(a,l_1)]$*, if the manager is granted non-indexed stock; or*

2) c_k *evaluated at the variance* $\sigma^2 - \sigma_l^2$ *(where* c_k *is defined in (3.6)), if the manager is granted indexed options;*

3) $\frac{1}{2}\lambda a^2(\sigma^2 - \sigma_l^2)$*, if the manager is granted indexed stock.*

So it is less costly to implement effort using indexed stock than using non-indexed stock or indexed options.

If shareholders are optimistic about market return, i.e. their perception of l is high, (i.e., l_2), then they would like to grab the market return as much as possible. Thus indexed stock should be the best choice. But if shareholders' perception of market return is low, then paying non-indexed options can be better if managers perceive a high market return with non-indexed options, because the non-indexed options look more valuable to managers than to shareholders. The different valuations of options between managers and shareholders will reduce the cost of non-indexed options from shareholders' perspective. As a result, the optimal way to implement effort is using non-indexed options.

Proposition 3.2.9. *1) If shareholders' perception of market return is high, i.e.* $l_s = l_2$*, then it is less costly to implement the target effort by using indexed stock than using non-indexed options.*

2) If shareholders' perception of market return is low, i.e. $l_s = l_1$ *and the manager's risk-aversion* λ *satisfies*

$$\frac{2}{\beta_k(a,l_2)}\frac{m_k(a,l_2) - m_k(a,l_1)}{\sigma_k^2(a,l_2) - \sigma_k^2(a,l_1)} < \lambda < \frac{2\beta_k(a,l_2)[m_k(a,l_2) - m_k(a,l_1)]}{\beta_k^2(a,l_2)\sigma_k^2(a,l_2) - a^2(\sigma^2 - \sigma_l^2)}, \tag{3.13}$$

then among the four ways to pay the manager: 1) indexed stock; 2) non-indexed stock; 3) indexed options 4) non-indexed options, the least costly way to implement the target effort is using non-indexed options.

3.3 Conclusion

Standard models on executive compensation usually assume that the distribution on future firm value is known. This paper attempts to study the implications on CEO compensation if there are multiple possible distributions on future firm value (For example, due to the lack of information, this is quite possible for new firms and new industries). Following the literature on ambiguity-aversion, I assume that managers are ambiguity-averse, i.e., they will pick the worst case to evaluate their contracts and utilities. The paper finds that when there are multiple possible distributions on firm value, different types of contracts (e.g., linear contracts VS non-linear contracts) can induce managers to evaluate firm value and their contracts using different distributions. As a result, this can change the cost of contracts in different directions. In particular, I find that in the model, when there is no ambiguity, stock is less costly than options to induce effort. However, if there is ambiguity in firm risk and the manager is risk-averse and ambiguity-averse, options can be more efficient than stock. When the manager is granted more incentives, she tends to perceive a higher risk. Hence, shareholders are reluctant to grant a manager high incentives and take advantage of some small improvements, such as improvements in the manager's ability or firm's investment opportunities. Thus, compensation contracts exhibit an inertia property and benchmark pay is reasonable. Finally, I study the case where there is ambiguity in the mean of firm value. I also show that options can be more efficient. Especially, if the ambiguity is from market return, Tying CEO pay to the market can be optimal, which helps to explain the pay-for-luck puzzle.

Appendix A

Supplement to Relative Wealth Concerns and Executive Compensation

A.1 Proofs for Chapter 1

Derivation of (1.3)

$H_i = r_i \int_0^1 w_k dk = r_i[\int_0^1 (\alpha_k + \beta_k \pi_k a_k) dk + \int_0^1 \beta_k \tilde{m} dk + \int_0^1 \beta_k \eta_k dk] = r_i[W + M\tilde{m} + \int_0^1 \beta_k \eta_k dk]$. Since $(\eta_k)_{k\in[0,1]}$ are independent of each other, and the variance of $\beta_k \eta_k$ is bounded by $\sigma^2_{\eta k}$ which is bounded by $\sigma^2_{\max}$, by Kolmogorov's strong law of large numbers, $\int_0^1 \beta_k \eta_k dk$ converges to 0 almost surely. So $H_i = r_i[W + M\tilde{m}]$.

Proof of Proposition 1.2.2

Given a linear contract $w = \alpha + \beta V$, if a CEO exerts an effort a, then her expected utility is

$$U = \alpha + \beta \pi a - rW - \frac{1}{2}\lambda\left[(\beta - rM)^2 \sigma_m^2 + \beta^2 \sigma_\eta^2\right] - \frac{1}{2}a^2.$$

Taking the first-order derivative w.r.t a yields that the optimal effort taken by the CEO is $a = \beta\pi$. Note that the base salary α does not affect the CEO's effort decision, it must be set to make the participation constraint bind, i.e.

$$\alpha + \beta \pi a - rW - \frac{1}{2}\lambda\left[(\beta - rM)^2 \sigma_m^2 + \beta^2 \sigma_\eta^2\right] - \frac{1}{2}a^2 = \underline{u}. \tag{A.1}$$

The shareholders' objective is to maximize the expected firm value net of CEO pay. Combining with the binding IR constraint (A.1), we can obtain that the shareholders' objective is

$$\max_{\beta} \beta\pi^2 - \left[rW + \frac{1}{2}\lambda \left[(\beta - rM)^2 \sigma_m^2 + \beta^2 \sigma_\eta^2 \right] + \frac{1}{2}\beta^2\pi^2 + \underline{u} \right]$$

Taking the first-order derivative w.r.t β yields that the optimal CEO incentives are

$$\beta^* = \min\left(\frac{\pi^2 + \lambda r M \sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}, 1 \right).$$

So if $rM\sigma_m^2 \geq \sigma_m^2 + \sigma_\eta^2$, then $\beta^* = 1$.

If $rM\sigma_m^2 < \sigma_m^2 + \sigma_\eta^2$, $\beta^* = \frac{\pi^2+\lambda r M\sigma_m^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$ is an optimal interior solution. From the solution of β^*, it is straight-forward to check that β^* is increasing in π, r, and M; decreasing in λ and σ_η^2. For the relation between β^* and σ_m^2,

$$\frac{\partial \beta^*}{\partial \sigma_m^2} = \frac{\lambda}{[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]^2} \left[r\frac{\partial(M\sigma_m^2)}{\partial \sigma_m^2}[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)] - (\pi^2 + \lambda r M \sigma_m^2) \right].$$

So if $r\frac{\partial(M\sigma_m^2)}{\partial\sigma_m^2} > \frac{\pi^2+\lambda r M\sigma_m^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$, then β^* is increasing in σ_m^2; if $r\frac{\partial(M\sigma_m^2)}{\partial\sigma_m^2} < \frac{\pi^2+\lambda r M\sigma_m^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$, then β^* is decreasing in σ_m^2.

Plugging the solution of β^* into (A.1) yields that

$$\alpha^* = \frac{(\pi^2 - \lambda r M \sigma_m^2)(\pi^2 + \lambda r M \sigma_m^2)}{2[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]} - \frac{\pi^2(\pi^2 + \lambda r M \sigma_m^2)^2}{[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]^2} + rW + \frac{1}{2}\lambda r^2 M^2 \sigma_m^2 + \underline{u}.$$

The expected CEO pay is

$$E[w] = \alpha^* + \beta^*\pi a = \frac{\pi^4}{2[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]} + rW + \frac{1}{2}\lambda r^2 M^2 \sigma_m^2 \left[1 - \frac{\lambda \sigma_m^2}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)} \right] + \underline{u},$$

which is increasing in W and M.

Proof of Lemma 1.2.4

Proof of necessity

Suppose an (hiring) equilibrium exists and let M be the aggregate CEO incentives and W be the aggregate expected CEO pay in the equilibrium. Then by Proposition 1.2.2, for each firm i, the optimal CEO incentives are given by $\beta_i^* = \frac{\pi_i^2+\lambda_i r_i M\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}$. And the expected CEO pay is $E[w_i] = \frac{\pi_i^4}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]} + r_i W + \frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2\left[1-\frac{\lambda_i\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\right] + \underline{u}_i$. Since $M = \int_0^1 \beta_i^* di$ and $W = \int_0^1 E[w_i]di$, we can solve that M has a unique solution $M = \frac{\int_0^1 \frac{\pi_i^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}di}{1-\int_0^1 \frac{\lambda_i r_i\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}di}$; and $W = \frac{\int_0^1\left[\frac{\pi_i^4}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]}+\frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2\left[1-\frac{\lambda_i\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\right]+\underline{u}_i\right]di}{1-\int_0^1 r_i di}$. Since $W > 0$ and $M > 0$, we must have $\int_0^1 r_i di < 1$. To ensure that each firm will hire its CEO, i.e. an interior solution exists, we must have $\beta_i^* \le 1$ for each $i \in [0,1]$, which is equivalent to $r_i M\sigma_m^2 \le \sigma_m^2 + \sigma_{\eta i}^2$ for each $i \in [0,1]$. Also in each firm, shareholders' payoff must be nonnegative, i.e. (1.6) is satisfied.

Proof of sufficiency

Let

$$
\begin{aligned}
M &= \frac{\int_0^1 \frac{\pi_i^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}di}{1-\int_0^1 \frac{\lambda_i r_i\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}di}, \\
W &= \frac{\int_0^1\left[\frac{\pi_i^4}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]}+\frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2\left[1-\frac{\lambda_i\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}\right]+\underline{u}_i\right]di}{1-\int_0^1 r_i di}.
\end{aligned}
$$

Since $\int_0^1 r_i di < 1$, $W > 0$ and $M > 0$. Then we can check that

$$
\begin{cases}
\alpha_i^* = \frac{(\pi_i^2-\lambda_i r_i M\sigma_m^2)(\pi_i^2+\lambda_i r_i M\sigma_m^2)}{2[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]} - \frac{\pi_i^2(\pi_i^2+\lambda_i r_i M\sigma_m^2)^2}{[\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)]^2} + r_i W + \frac{1}{2}\lambda_i r_i^2 M^2\sigma_m^2 + \underline{u}_i, \\
\beta_i^* = \frac{\pi_i^2+\lambda_i r_i M\sigma_m^2}{\pi_i^2+\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)} \le 1, \\
a_i^* = \pi_i \beta_i^*.
\end{cases}
$$

is the unique equilibrium that satisfies the properties in Definition 1.2.3.

Proof of Lemma 1.2.5

The shareholders' payoffs are

$$
\frac{\pi^2(\pi^2+\lambda r M\sigma_m^2)}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)]} - \left\{\frac{\pi^4}{2[\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)]} + rW + \frac{1}{2}\lambda r^2 M^2\sigma_m^2\left[1-\frac{\lambda\sigma_m^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right] + \underline{u}\right\},
$$

which is quadratic in r. Therefore it is easy to see that if $\frac{\lambda M\sigma_m^2\pi^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)} - W \le 0$, $r^* = 0$; if $\frac{\lambda M\sigma_m^2\pi^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)} - W > 0$,

$r^* = \frac{\lambda M \sigma_m^2 \pi^2 - W[\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]}{\lambda M^2 \sigma_m^2 (\pi^2 + \lambda \sigma_\eta^2)} > 0.$

Proof of Lemma 1.2.6

Proof of necessity

Given any $\hat{r} < r$, if a CEO of type r chooses the contract $(\alpha(\hat{r}), \beta(\hat{r}))$ and exerts an effort a, then her relative payoff is

$$w - H = \alpha(\hat{r}) + \beta(\hat{r})\pi a + (\beta(\hat{r}) - rM)\tilde{m} + \beta(\hat{r})\eta - rW.$$

Then we can calculated that her expected utility is given by

$$\begin{aligned} U(\hat{r}|r) &= E[w-H] - \frac{1}{2}\lambda Var[w-H] - \frac{1}{2}a^2 \\ &= \alpha(\hat{r}) + \beta(\hat{r})\pi a - rE[W] \\ &- \frac{1}{2}\lambda\left[(\beta(\hat{r}) - rE[M])^2\sigma_m^2 + \beta(\hat{r})^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right] - \frac{1}{2}a^2. \end{aligned}$$

Thus the optimal effort taken by the CEO is $a = \beta(\hat{r})\pi$ and her expected utility is

$$\begin{aligned} U(\hat{r}|r) &= \alpha(\hat{r}) + \frac{1}{2}\beta(\hat{r})^2\pi^2 - rE[W] \\ &- \frac{1}{2}\lambda\left[(\beta(\hat{r}) - rE[M])^2\sigma_m^2 + \beta(\hat{r})^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right]. \end{aligned} \tag{A.2}$$

By the IC constraint, $U(\hat{r}|r) \le U(r|r)$ implies that

$$\alpha(r) - \alpha(\hat{r}) \ge \frac{1}{2}[\pi^2 - \lambda(\sigma_m^2 + \sigma_\eta^2)](\beta(\hat{r})^2 - \beta(r)^2) + \lambda r E[M]\sigma_m^2(\beta(\hat{r}) - \beta(r)). \tag{A.3}$$

Similarly, $U(\hat{r}|\hat{r}) \ge U(r|\hat{r})$ implies that

$$\alpha(r) - \alpha(\hat{r}) \le \frac{1}{2}[\pi^2 - \lambda(\sigma_m^2 + \sigma_\eta^2)](\beta(\hat{r})^2 - \beta(r)^2) + \lambda \hat{r} E[M]\sigma_m^2(\beta(\hat{r}) - \beta(r)). \tag{A.4}$$

Thus, combining (A.3) and (A.4), and also note that $\hat{r} < r$, we must have $\beta(\hat{r}) \le \beta(r)$ for any $\hat{r} < r$. So $\frac{\partial \beta(r)}{\partial r} \ge 0$. Dividing the both sides of (A.3) and (A.4) by $r - \hat{r}$ and letting $\hat{r} \to r$ yield

$$\lim_{\hat{r} \to r} \frac{\alpha(r) - \alpha(\hat{r})}{r - \hat{r}} = -\left\{[\pi^2 - \lambda(\sigma_m^2 + \sigma_\eta^2)]\beta(r) + \lambda r E[M]\sigma_m^2\right\}\frac{\partial \beta(r)}{\partial r},$$

i.e. $\frac{\partial \alpha(r)}{\partial r} = -\left\{[\pi^2 - \lambda(\sigma_m^2 + \sigma_\eta^2)]\beta(r) + \lambda r E[M]\sigma_m^2\right\} \frac{\partial \beta(r)}{\partial r}$.

Proof of Sufficiency

Given the two conditions in the lemma, we can calculate that

$$\frac{\partial U(\hat{r}|r)}{\partial \hat{r}} = \lambda E[M]\sigma_m^2 (r - \hat{r}) \frac{\partial \beta(r)}{\partial r}.$$

Then it is easy to see that $\frac{\partial U(\hat{r}|r)}{\partial \hat{r}}$ is positive when $\hat{r} < r$ and negative when $\hat{r} > r$. So $U(\hat{r}|r)$ is maximized at $\hat{r} = r$. In other words, $U(\hat{r}|r) \le U(r|r)$ for any $\hat{r}$ and r, i.e. IC constraint holds.

Proof of Proposition 1.2.7

Suppose $U(r_0|r_0) = \min_{r \in [0, r_{\max}]} U(r|r)$, then we have $U(r_0|r_0) = \underline{u}$. Since from (A.2), we can obtain that

$$\frac{dU(r|r)}{dr} = \lambda E[M]\sigma_m^2 \beta(r) - E[W] - \lambda r \left(E[M^2]\sigma_m^2 + Var[W]\right). \tag{A.5}$$

For brevity, let $Y(r) = E[W] + \lambda r \left(E[M^2]\sigma_m^2 + Var[W]\right)$. Then the CEO of type r has obtain a utility $U(r|r) = \underline{u} - \int_r^{r_0} \left[\lambda E[M]\sigma_m^2 \beta(r) - Y(r)\right] dr$. Then the expected CEO pay is

$$\begin{aligned} E[w(r)] &= rE[W] + \frac{1}{2}\lambda \left[(\beta(r) - rE[M])^2\sigma_m^2 + \beta(r)^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right] \\ &\quad + \frac{1}{2}\beta(r)^2\pi^2 + U(r|r) \\ &= rE[W] + \frac{1}{2}\lambda \left[(\beta(r) - rE[M])^2\sigma_m^2 + \beta(r)^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right] \qquad \text{(A.6)} \\ &\quad + \frac{1}{2}\beta(r)^2\pi^2 + \underline{u} - \int_r^{r_0} \left[\lambda E[M]\sigma_m^2 \beta(r) - Y(r)\right] dr. \end{aligned}$$

Then the shareholders' objective is to maximize the expected firm value net of CEO pay across CEOs of all possible types, i.e. $\max\int_0^{r_{\max}} f(r)\left[\beta(r)\pi^2 - E[w(r)]\right]dr$, which is equal to

$$\begin{aligned}
& \int_0^{r_0} f(r)\Big\{\beta(r)\pi^2 - rE[W] - \frac{1}{2}\beta(r)^2\pi^2 - \underline{u} + \frac{F(r)}{f(r)}\left[\lambda E[M]\sigma_m^2\beta(r) - Y(r)\right] \\
- & \frac{1}{2}\lambda\left[(\beta(r) - rE[M])^2\sigma_m^2 + \beta(r)^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right]\Big\}dr \\
+ & \int_{r_0}^{r_{\max}} f(r)\Big\{\beta(r)\pi^2 - rE[W] - \frac{1}{2}\beta(r)^2\pi^2 - \underline{u} - \frac{1-F(r)}{f(r)}\left[\lambda E[M]\sigma_m^2\beta(r) - Y(r)\right] \\
- & \frac{1}{2}\lambda\left[(\beta(r) - rE[M])^2\sigma_m^2 + \beta(r)^2\sigma_\eta^2 + r^2 Var[M]\sigma_m^2 + r^2 Var[W]\right]\Big\}dr
\end{aligned}$$

Proof of Part 1)

If $r_0 = 0$, then the pointwise optimization of the above objective function yields that the derivative w.r.t $\beta(r)$ is

$$\pi^2 + \lambda\sigma_m^2 E[M]\left[r - \frac{1-F(r)}{f(r)}\right] - [\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)]\beta(r).$$

Let $\beta_1(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r - \frac{1-F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$. Then the objective function is increasing in $\beta(r)$ if $\beta(r) < \beta_1(r)$; and decreasing in $\beta(r)$ if $\beta(r) > \beta_1(r)$; So the optimal incentives in this case are $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r - \frac{1-F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$.

Proof of Part 2)

Similarly, if $r_0 = r_{\max}$, let $\beta_2(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r + \frac{F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$, then the objective function is increasing in $\beta(r)$ if $\beta(r) < \beta_2(r)$; and decreasing in $\beta(r)$ if $\beta(r) > \beta_2(r)$; So the optimal incentives in this case are $\beta^*(r) = \frac{\pi^2 + \lambda\sigma_m^2 E[M]\left[r + \frac{F(r)}{f(r)}\right]}{\pi^2 + \lambda(\sigma_m^2 + \sigma_\eta^2)}$.

Proof of Part 3)

If $0 < r_0 < r_{\max}$, then without IC constraint, for $r < r_0$, the optimal incentives are $\beta^*(r) = \beta_2(r)$; and $r > r_0$, the optimal incentives are $\beta^*(r) = \beta_1(r)$. However, the IC constraint requires that $\beta^*(r)$ is increasing in r, and also note that $\beta_1(r) < \beta_2(r)$, so in this case, we cannot get the first-best solution.

Suppose $\beta^*(r)|_{r\in[0,r_{\max}]}$ is the solution. I first show that for $r \geq r_0$, the optimal incentives must have the following form: there exists $r_2 \geq r_0$ such that $\beta^*(r)$ is constant for $r_0 \leq r \leq r_2$ and $\beta^*(r) = \beta_1(r)$ for $r \geq r_2$. If $\beta^*(r) < \beta_1(r)$ for for any $r \geq r_0$, then the shareholders' payoffs can be improved by increasing $\beta^*(r)$ to $\beta_1(r)$ for each $r \geq r_0$. Thus let r_2 be the minimum $r \in [r_0, r_{\max}]$ that satisfies $\beta^*(r_2) = \beta_1(r_2)$. Then similar argument implies that $\beta^*(r) = \beta_1(r)$ for $r \geq r_2$. For $r_0 \leq r \leq r_2$, by the definition of r_2, we must have either 1) $\beta^*(r) < \beta_1(r)$ for any $r_0 \leq r \leq r_2$; or 2) $\beta^*(r) > \beta_1(r)$ for any $r_0 \leq r \leq r_2$. In the first case, the shareholders' payoffs can be improved by increasing $\beta^*(r)$ to $\beta_1(r)$. In the second case, if there exists $r_0 \le r_3 < r_2$ such that $\beta^*(r_3) < \beta^*(r_2)$, then the shareholders' payoffs can be improved by reducing $\beta^*(r)$ to $\max(\beta^*(r_3), \beta_1(r))$ for $r \in [r_3, r_2]$. So we must have $\beta^*(r)$ is constant for $r_0 \le r \le r_2$ and $\beta^*(r) = \beta_1(r)$ for $r \ge r_2$.

Similarly, for $r \le r_0$, the optimal incentives must have the following form: there exists $r_1 \le r_0$ such that $\beta^*(r) = \beta_2(r)$ for $0 \le r \le r_1$ and $\beta^*(r)$ is constant for $r_1 \le r \le r_0$. If $\lim_{r\to r_0-} \beta^*(r) < \lim_{r\to r_0+} \beta^*(r)$, then the shareholders' payoffs can be improved by reducing $\beta^*(r)$ around right side of r_0 a little bit. So there must exist $r_1 \le r_0 \le r_2$ such that $\beta^*(r) = \beta_2(r)$ for $0 \le r \le r_1$, $\beta^*(r)$ is constant for $r_1 \le r \le r_2$, and $\beta^*(r) = \beta_1(r)$ for $r \ge r_2$.

Proof of Example 1

Firstly, notice that by Kolmogorov's strong law of large numbers, M converges to $E[M]$ and W converges to $E[W]$ almost surely. Thus I treat M and W as contants in the following proof and $Var[M] = Var[W] = 0$. Now I will characterize the two equilibria as follows:

Equilibrium 1

Suppose $r_0 = 0$ (i.e. the CEO of type $r = 0$ is at the reservation utility), then by Proposition 1.2.7, the optimal incentives for firm i are $\beta_{i1}(r) = \frac{\pi^2+\lambda\sigma_m^2 M_1(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$. Aggregating all the CEO incentives yields that

$$M_1 = \int_0^{r_{\max}} \beta_{i1}(r)di = E_r\left[\frac{\pi^2+\lambda\sigma_m^2 M_1(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right] = \frac{\pi^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}.$$

From the proof of Proposition 1.2.7, the expected CEO pay in each firm is given by (A.6). Then aggregating the expected CEO pay from all the firms yields that

$$W_1 = \frac{1}{2}[\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)]\left[1-\frac{1}{3}\left(\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)^2\right]M_1^2 + \underline{u}.$$

$r_0 = 0$ means that $U(r|r) \ge U(0|0)$ for any $r \in [0, r_{\max}]$. From (A.5), this is equivalent to

$$\int_0^r \left[\lambda M_1^2\sigma_m^2\left(1+\frac{\lambda\sigma_m^2(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right) - W_1 - \lambda r M_1^2\sigma_m^2\right]dr \ge 0, \text{ for any } r \in [0, r_{\max}].$$

Finally, the above condition can be satisfied as long as

$$\lambda M_1^2\sigma_m^2 \cdot \min\left(1-\frac{1}{2}r_{\max}, 1-\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right) - W_1 \geq 0.$$

Equilibrium 2

Suppose $r_0 = r_{\max}$ (i.e. the CEO of type $r = r_{\max}$ is at the reservation utility), then by Proposition 1.2.7, the optimal incentives for firm i are $\beta_{i2}(r) = \frac{\pi^2+2r\lambda\sigma_m^2 M_2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}$. Aggregating all the CEO incentives yields that

$$M_2 = \frac{\pi^2}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)-\lambda\sigma_m^2 r_{\max}}.$$

Then $\beta_{i2}(r) = \left(1+\frac{\lambda\sigma_m^2(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)M_2$. Similarly, we can compute that the sum of the expected CEO pay from all the firms is

$$W_2 = \frac{1}{1-r_{\max}}\left[\frac{1}{2}[\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)]\left[1-\frac{\lambda\sigma_m^2 r_{\max}(2-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}-\frac{1}{3}\left(\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)^2\right]M_2^2+\underline{u}\right].$$

$r_0 = r_{\max}$ means that $U(r|r) \geq U(r_{\max}|r_{\max})$ for any $r \in [0, r_{\max}]$. From (A.5), this is equivalent to

$$\int_r^{r_{\max}}\left[\lambda M_2^2\sigma_m^2\left(1+\frac{\lambda\sigma_m^2(2r-r_{\max})}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right)-W_1-\lambda r M_2^2\sigma_m^2\right]dr \leq 0, \text{ for any } r\in[0,r_{\max}].$$

The above condition can be satisfied as long as

$$\lambda M_2^2\sigma_m^2 \cdot \max\left(1-\frac{1}{2}r_{\max}, 1-r_{\max}+\frac{\lambda\sigma_m^2 r_{\max}}{\pi^2+\lambda(\sigma_m^2+\sigma_\eta^2)}\right) - W_2 \leq 0.$$

Proof of Proposition 1.2.8

Similar to the proof of Proposition 1.2.2.

Proof of Lemma 1.3.1

Given the contract $w = \alpha + \beta V$, the CEO's utility is given by

$$\begin{aligned} U &= E[w-H] - \frac{1}{2}\lambda Var[w-H] - \frac{1}{2}a^2 \\ &= \alpha + \beta\pi a - rW - \frac{1}{2}\lambda[(\beta a - rM)^2\sigma_m^2 + (\beta a)^2\sigma_\eta^2] - \frac{1}{2}a^2. \end{aligned}$$

U is concave in a, so taking the derivative w.r.t a yields that the optimal effort is

$$a = \frac{\beta(\pi + \lambda\sigma_m^2 rM)}{\lambda(\sigma_m^2 + \sigma_\eta^2)\beta^2 + 1}.$$

Then the comparative statics are straight-forward.

Proof of Proposition 1.3.2

By Lemma 1.3.1, the CEO will exert an effort $a = \frac{\beta(\pi+\lambda rM\sigma_m^2)}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1}$ if he is granted a contract $w = \alpha + \beta V$. Note that the base salary α does not affect the CEO's effort decision, it must be set to make the participation constraint bind, that is

$$\alpha + \beta\pi a - rW - \frac{1}{2}\lambda[(\beta a - rM)^2\sigma_m^2 + (\beta a)^2\sigma_\eta^2] - \frac{1}{2}a^2 = \underline{u}.$$

So we have

$$E[\text{CEO pay}] = \alpha + \beta\pi a = rW + \frac{1}{2}\lambda[(\beta a - rM)^2\sigma_m^2 + (\beta a)^2\sigma_\eta^2] + \frac{1}{2}a^2 + \underline{u}.$$

Then the shareholders' payoffs are

$$\begin{aligned} E[V-w] &= \pi a - (\alpha + \beta\pi a) \\ &= \pi a - \left[rW + \frac{1}{2}\lambda[(\beta a - rM)^2\sigma_m^2 + (\beta a)^2\sigma_\eta^2] + \frac{1}{2}a^2 + \underline{u}\right] \\ &= \frac{\beta\pi(\pi + \lambda rM\sigma_m^2)}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1} - \frac{1}{2}\frac{\beta^2(\pi - \lambda rM\sigma_m^2)(\pi + \lambda rM\sigma_m^2)}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1} - rW - \frac{1}{2}\lambda(rM)^2\sigma_m^2 - \underline{u}. \end{aligned}$$

Taking the derivative w.r.t β yields that

$$\frac{\partial E[V-w]}{\partial\beta} = \frac{\pi(\pi + \lambda rM\sigma_m^2)}{[\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1]^2}\left[-\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2 - \left(1 - \frac{\lambda rM\sigma_m^2}{\pi}\right)\beta + 1\right].$$

$\frac{\partial E[V-w]}{\partial \beta}$ is increasing in β when $\beta < \frac{2}{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)+\sqrt{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)}}$ and decreasing in β when $\beta > \frac{2}{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)+\sqrt{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)}}$, therefore the optimal incentives are

$$\beta^* = \min\left(\frac{2}{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)+\sqrt{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)}}, 1\right).$$

If $rM\sigma_m^2 \geq h(\sigma_m^2+\sigma_\eta^2)$, then $\beta^*=1$.

If $rM\sigma_m^2 < h(\sigma_m^2+\sigma_\eta^2)$, then $\beta^* = \frac{2}{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)+\sqrt{\left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)^2+4\lambda(\sigma_m^2+\sigma_\eta^2)}}$. For the comparative statics, let $F(\beta,q) = -\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2 - \left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)\beta + 1$, where q refers to basic parameters. Then $\frac{\partial F}{\partial \beta}|_{\beta^*} < 0$, so from $\frac{\partial F}{\partial \beta}\frac{\partial \beta}{\partial q} + \frac{\partial F}{\partial q} = 0$, $\frac{\partial \beta}{\partial q}|_{\beta^*}$ has the same sign with $\frac{\partial F}{\partial q}|_{\beta^*}$. Then it is straight-forward to derive the results in part 1). For part 2), $\frac{\partial F}{\partial \sigma_m^2}|_{\beta^*} = -\lambda\beta^{*2} + \frac{\lambda r\beta^*}{\pi}\left(\frac{\partial M}{\partial \sigma_m^2}\sigma_m^2 + M\right)$. So when $rM + r\sigma_m^2\frac{\partial M}{\partial \sigma_m^2} < \pi\beta^*$, β^* is decreasing in σ_m^2; when $rM + r\sigma_m^2\frac{\partial M}{\partial \sigma_m^2} > \pi\beta^*$, β^* is increasing in σ_m^2. Part 3) is obvious.

Proof of Lemma 1.3.3

Let $\sigma^2 = \sigma_m^2 + \sigma_\eta^2$. From lemma 1.3.1, we can calculate that

$$\frac{da}{dr} = \frac{\lambda\beta\sigma_m^2 M}{\lambda\sigma^2\beta^2+1} + (\pi + \lambda\sigma_m^2 rM)\frac{1-\lambda\sigma^2\beta^2}{(\lambda\sigma^2\beta^2+1)^2}\frac{\partial \beta}{\partial r}.$$

Since $-\lambda\sigma^2\beta^2 - \left(1-\frac{\lambda r M\sigma_m^2}{\pi}\right)\beta + 1 = 0$,

$$\frac{\partial \beta}{\partial r} = \frac{\frac{\lambda M\sigma_m^2\beta}{\pi}}{2\lambda\sigma^2\beta + \left(1-\frac{\lambda r M\sigma_m^2}{h}\right)} = \frac{\frac{\lambda M\sigma_m^2\beta^2}{\pi}}{\lambda\sigma^2\beta^2+1}.$$

So

$$\frac{da}{dr} = \frac{\lambda M\beta\sigma_m^2}{\lambda\sigma^2\beta^2+1}\left[1 + (\pi+\lambda\sigma_m^2 rM)\frac{\beta}{\pi}\frac{1-\lambda\sigma^2\beta^2}{(\lambda\sigma^2\beta^2+1)^2}\right].$$

Note that $(\pi+\lambda\sigma_m^2 rM)\frac{\beta}{\pi} = \lambda\sigma^2\beta^2 + 2\beta - 1$,

$$\frac{da}{dr} = \frac{\lambda M\beta\sigma_m^2}{(\lambda\sigma^2\beta^2+1)^3}\left[(\lambda\sigma^2\beta^2+1)^2 - (\lambda\sigma^2\beta^2+2\beta-1)(1-\lambda\sigma^2\beta^2)\right] > 0.$$

So $\frac{da}{dr} > 0$. By the similar argument, $\frac{da}{dM} > 0$.

Proof of Lemma 1.3.5

The procedure of solving an equilibrium is as follows:

1) Guess a value for $M = \int_0^1 a_i \beta_i di$, the sum of the sensitivity of CEO pay to the aggregate shock.

2) Then for each firm i, we can solve the optimal incentives and CEO's effort from Proposition 1.3.2 and Lemma 1.3.1. Specifically, given M, we have $\beta_i(M) = \frac{2}{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)+\sqrt{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)^2+4\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}}$ and $a_i = \frac{\beta_i(M)(\pi_i+\lambda_i\sigma_m^2 r_i M)}{\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)\beta_i(M)^2+1}$.

3) Plugging the expressions for $\beta_i(M)$ and a_i into $M = \int_0^1 a_i \beta_i di$, we can solve for M.

4) Plugging the solution of M into $\beta_i(M) = \frac{2}{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)+\sqrt{\left(1-\frac{\lambda_i \sigma_m^2 r_i M}{\pi_i}\right)^2+4\lambda_i(\sigma_m^2+\sigma_{\eta i}^2)}}$ to check that the incentives are not greater than 1 (ensured by (1.11)).

5) Bounded efforts: let $a_k = \max_{i\in[0,1]} a_i$. Then by Lemma 1.3.1, $a_k = \frac{\beta_k(\pi_k+\lambda_k r_k M\sigma_m^2)}{\lambda_k(\sigma_m^2+\sigma_{\eta k}^2)\beta_k^2+1}$. Since $M = \int_0^1 a_i\beta_i di \leq a_k \int_0^1 \beta_i di \leq a_k$, $a_k \leq \frac{\beta_k(\pi_k+\lambda_k r_k a_k \sigma_m^2)}{\lambda_k(\sigma_m^2+\sigma_{\eta k}^2)\beta_k^2+1}$. Note that $\beta_k \geq \frac{2}{1+\sqrt{1+4\lambda_k(\sigma_m^2+\sigma_{\eta k}^2)}}$, so by (1.10), $a_k \leq \frac{\beta_k \pi_k}{\lambda_k\beta_k[\beta_k(\sigma_m^2+\sigma_{\eta k}^2)-r_k\sigma_m^2]+1} \leq \pi_k$.

The procedure above is summarized in Lemma 1.3.5.

Proof of Example 2

Let β and a denote the optimal incentives and CEO's effort for firms in $[0,k]$. Let $\bar{\beta}$ and $\bar{a}$ denote the optimal incentives and CEO's effort for firms in $[1-k,1]$. Then $M = \int_0^1 a_i\beta_i di = ka\beta + (1-k)\bar{a}\bar{\beta}$. For brevity, let $\sigma^2 = \sigma_m^2 + \sigma_\eta^2$ and $\bar{\sigma}^2 = \sigma_m^2 + \bar{\sigma}_\eta^2$. By Lemma 1.3.1, $a = \frac{\beta(\pi+\lambda\sigma_m^2 rM)}{\lambda\sigma^2\beta^2+1}$, and $\bar{a} = \frac{\bar{\beta}(\pi+\lambda\sigma_m^2 \bar{r}M)}{\lambda\bar{\sigma}^2\bar{\beta}^2+1}$. Then

$$M = ka\beta + (1-k)\bar{a}\bar{\beta} = \frac{k\beta^2(\pi+\lambda\sigma_m^2 rM)}{\lambda\sigma^2\beta^2+1} + \frac{(1-k)\bar{\beta}^2(\pi+\lambda\sigma_m^2\bar{r}M)}{\lambda\bar{\sigma}^2\bar{\beta}^2+1},$$

which implies that

$$M = \frac{\frac{k\beta^2\pi}{\lambda\sigma^2\beta^2+1} + \frac{(1-k)\bar{\beta}^2\pi}{\lambda\bar{\sigma}^2\bar{\beta}^2+1}}{1-\lambda\sigma_m^2\left[\frac{rk\beta^2}{\lambda\sigma^2\beta^2+1} + \frac{\bar{r}(1-k)\bar{\beta}^2}{\lambda\bar{\sigma}^2\bar{\beta}^2+1}\right]}.$$

By the proof of Proposition 1.3.2, β satisfies that

$$-\lambda\sigma^2\beta^2 - \left(1-\frac{\lambda\sigma_m^2 rM}{\pi}\right)\beta + 1 = 0. \tag{A.7}$$

$\bar{\beta}$ satisfies that

$$-\lambda\bar{\sigma}^2\bar{\beta}^2 - \left(1 - \frac{\lambda\sigma_m^2\bar{r}M}{\pi}\right)\bar{\beta} + 1 = 0. \tag{A.8}$$

From (A.7) and (A.8), we obtain that

$$\frac{\bar{r}}{r} = \frac{\lambda\bar{\sigma}^2\bar{\beta} + 1 - 1/\bar{\beta}}{\lambda\sigma^2\beta + 1 - 1/\beta}. \tag{A.9}$$

Thus we can solve $\bar{\beta}$ in terms of β, i.e. $\bar{\beta} = \bar{\beta}(\beta)$. Then β is a solution to

$$-\lambda\sigma^2\beta^2 - \left(1 - \frac{\lambda\sigma_m^2 r M(\beta)}{\pi}\right)\beta + 1 = 0.$$

Let $F(\beta)$ denote the left side of the above equation. It is obvious that $F(0) > 0$. Note that $\bar{\beta}(1) = 1$ by (A.9) and the condition 2) in the lemma (i.e. $\lambda r\bar{\sigma}^2 = \lambda\bar{r}\sigma^2$), then it is easy to check that $F(1) > 0$ if and only if $\sigma_m^2(1-k) > \frac{\sigma^2}{r} - \sigma_m^2 k$. Similarly, $F(\sqrt{\frac{1}{\lambda\sigma^2}}) < 0$ if and only if

$$\sigma_m^2(1-k)\left(\lambda\sigma^2 + \lambda\sigma^2\frac{\bar{r}}{r}\right)\frac{1}{\lambda\sigma^2\frac{\bar{r}}{r} + \frac{1}{\bar{\beta}(\sqrt{\frac{1}{\lambda\sigma^2}})^2}} < \frac{\sigma^2}{r} - \sigma_m^2 k. \tag{A.10}$$

From (A.9), we can solve that $\bar{\beta}(\sqrt{\frac{1}{\lambda\sigma^2}}) = \frac{(\bar{r}-r)+\sqrt{(\bar{r}-r)^2+4\lambda\sigma^2\bar{r}r}}{2\lambda\sigma^2\bar{r}}$. Plugging it into (A.10) yields that $F(\sqrt{\frac{1}{\lambda\sigma^2}}) <$ 0 if and only if $\frac{r+\bar{r}}{2\bar{r}}\left(1 + \frac{\bar{r}-r}{\sqrt{(\bar{r}-r)^2+4\lambda\sigma^2\bar{r}r}}\right) < \frac{\sigma_m^2+\sigma_\eta^2-\sigma_m^2 rk}{\sigma_m^2 r(1-k)}$. Hence given the four conditions in the lemma, we have $F(0) > 0$, $F(\sqrt{\frac{1}{\lambda\sigma^2}}) < 0$, and $F(1) > 0$. Thus for firms in $[0,k]$, there is an equilibrium in which β is less than $\sqrt{\frac{1}{\lambda\sigma^2}}$; and there is another equilibrium in which β is greater than $\sqrt{\frac{1}{\lambda\sigma^2}}$.

Proof of Lemma 1.4.1

Since firms are identical, we can assume that in equilibrium, for firms in $[0,k]$, the shareholders are using an option-based contract $(\alpha_O + \beta_O \max(e_1(\tilde{m}+\eta), 0))$ to induce their CEOs' efforts; for firms in $[k,1]$, the shareholders are using a stock-based contract $\alpha_S + \beta_S(\pi + e_1(\tilde{m}+\eta))$ to induce their CEOs' efforts. Then the sum of all CEOs' wages is $C = \int_0^1 w_i\,di = k\alpha_O + (1-k)\alpha_S + k\beta_O\int_0^1 \max(e_1(\tilde{m}+\eta_i), 0)\,di + (1-k)\beta_S(\pi + e_1\tilde{m})$ (because $\int_0^1 \eta_i\,di = 0$).

Then for any representative firm (i.e. I drop subscript i for brevity), if shareholders use a stock contract, then if the CEO exerts an effort at a cost c, her relative payoff is $w_S - rC = \alpha_S + \beta_S(\pi + e_1(\tilde{m}+\eta)) - rC$. Since $\tilde{m} = \sigma_m$, or $-\sigma_m$ with probability $\frac{1}{2}$ for each, $C = C^+ = k\alpha_O + (1-k)\alpha_S + k\beta_O E_\eta[\max(e_1(\sigma_m +$

$\eta),0)] + (1-k)\beta_S(\pi + e_1\sigma_m)]$ with probability $\frac{1}{2}$ and $C = C^- = k\alpha_O + (1-k)\alpha_S + k\beta_O E_\eta[\max(e_1(-\sigma_m + \eta),0)] + (1-k)\beta_S(\pi - e_1\sigma_m)]$ with probability $\frac{1}{2}$. Then we can calculate that the mean of $w_S - rC$ is $\alpha_S + \beta_S\pi - r\frac{C^+ + C^-}{2}$, and the variance of $w_S - rC$ is $\left(\beta_S e_1\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2$. Therefore the CEO's expected utility under working is $\alpha_S + \beta_S\pi - r\frac{C^+ + C^-}{2} - \frac{1}{2}\lambda\left[\left(\beta_S e_1\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right] - c$. Similarly, if the CEO does not work and given that all the other CEOs are working, then her relative payoff is $\alpha_S + \beta_S e_0(\tilde{m} + \eta) - rC$. It can be calculated that her expected utility under shirking is $\alpha_S - r\frac{C^+ + C^-}{2} - \frac{1}{2}\lambda\left[\left(\beta_S e_0\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right]$. So to guarantee that the CEO is working, we must have

$$\begin{aligned}
& \alpha_S + \beta_S\pi - r\frac{C^+ + C^-}{2} - \frac{1}{2}\lambda\left[\left(\beta_S e_1\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right] - c \\
\geq\ & \alpha_S - r\frac{C^+ + C^-}{2} - \frac{1}{2}\lambda\left[\left(\beta_S e_0\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right],
\end{aligned}$$

which can be simplified to $\beta_S\pi - \frac{1}{2}\lambda\beta_S^2\sigma_\eta^2(e_1^2 - e_0^2) - \frac{1}{2}\lambda\beta_S\sigma_m(e_1 - e_0)[\beta_S(e_1 + e_0)\sigma_m - r(C^+ - C^-)] - c \geq 0$. The base salary α_S must be set to make the participation constraint binding, thus the expected CEO pay is equal to $E[w_S] = r\frac{C^+ + C^-}{2} + \frac{1}{2}\lambda\left[\left(\beta_S e_1\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right] + c + \underline{u}$. Therefore, if the shareholders use stock, then their objective is

$$\min_{\beta_S} r\frac{C^+ + C^-}{2} + \frac{1}{2}\lambda\left[\left(\beta_S e_1\sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \beta_S^2 e_1^2\sigma_\eta^2\right] + c + \underline{u},$$

subject to $\beta_S\pi - \frac{1}{2}\lambda\beta_S^2\sigma_\eta^2(e_1^2 - e_0^2) - \frac{1}{2}\lambda\beta_S\sigma_m(e_1 - e_0)[\beta_S(e_1 + e_0)\sigma_m - r(C^+ - C^-)] - c \geq 0$.

Due to the assumption that shareholders will use stock compensation if there is no difference between stock and options, and since each firm is identical, we must have that $k = 0$ or $k = 1$. If all the firms use stock (i.e. $k = 0$), then we have $C^+ - C^- = 2\beta_S e_1\sigma_m$. Then it is easy to see that the optimal $\beta_S^* \geq 0$ must be the minimum solution to the constraint $\beta_S\pi - \frac{1}{2}\lambda\beta_S^2\sigma_\eta^2(e_1^2 - e_0^2) - \frac{1}{2}\lambda\beta_S\sigma_m(e_1 - e_0)[\beta_S(e_1 + e_0)\sigma_m - r(C^+ - C^-)] - c \geq 0$, which yields that $\beta_S^* = \frac{2c}{\pi + \sqrt{\pi^2 - 2c\lambda[(e_1^2 - e_0^2 - 2re_1(e_1 - e_0))\sigma_m^2 + (e_1^2 - e_0^2)\sigma_\eta^2]}}$. Also note that $\frac{C^+ + C^-}{2} = \alpha_S + \beta_S\pi = E[w_S]$, so we get that the cost of stock is $E[w_S] = \frac{1}{1-r}\left[\frac{1}{2}\lambda[(\beta_S^* e_1\sigma_m)^2(1-r)^2 + (\beta_S^* e_1\sigma_\eta)^2] + c + \underline{u}\right]$.

If shareholders use options, then we will calculate the cost of options in two cases: $\sigma_\eta \leq \sigma_m$ or $\sigma_\eta > \sigma_m$.

Case 1: $\sigma_\eta \leq \sigma_m$.

In this case, if the CEO works, then we can compute that the mean of $w_O - rC$ is $\alpha_O + \frac{1}{2}\beta_O e_1 \sigma_m - r\frac{C^+ + C^-}{2}$, and the variance of $w_O - rC$ is $\left(\frac{1}{2}\beta_O e_1 \sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \frac{1}{2}(\beta_O e_1 \sigma_\eta)^2$. Similarly, we can obtain that shareholders' objective is

$$\min_{\beta_O} r\frac{C^+ + C^-}{2} + \frac{1}{2}\lambda\left[\left(\frac{1}{2}\beta_O e_1 \sigma_m - r\frac{C^+ - C^-}{2}\right)^2 + \frac{1}{2}(\beta_O e_1 \sigma_\eta)^2\right] + c + \underline{u},$$

subject to $\frac{1}{2}\beta_O e_1 \sigma_m - \frac{1}{2}\lambda\frac{1}{2}(\beta_O e_1 \sigma_\eta)^2 - \frac{1}{2}\lambda\frac{1}{2}\beta_O e_1 \sigma_m\left(\frac{1}{2}\beta_O e_1 \sigma_m - r(C^+ - C^-)\right) - c \geq 0$.

If all the firms use options (i.e. $k = 1$), then $C^+ - C^- = \beta_O e_1 \sigma_m$ and $\frac{C^+ + C^-}{2} = E[w_O]$. So similarly we can derive the optimal number of options is $\frac{1}{2}\beta_O^* = \frac{2c}{e_1\sigma_m + \sqrt{(e_1\sigma_m)^2 - 2c\lambda[e_1^2(1-2r)\sigma_m^2 + 2e_1^2\sigma_\eta^2]}}$, and the cost of options is $E[w_O] = \frac{1}{1-r}\left[\frac{1}{2}\lambda\left[\left(\frac{1}{2}\beta_O^* e_1 \sigma_m\right)^2(1-r)^2 + 2\left(\frac{1}{2}\beta_O^* e_1 \sigma_\eta\right)^2\right] + c + \underline{u}\right]$.

Case 2: $\sigma_\eta > \sigma_m$.

Similarly, we have that the optimal number of options is $\frac{1}{2}\beta_O^* = \frac{2c}{e_1\sigma_\eta + \sqrt{(e_1\sigma_\eta)^2 - 2c\lambda[e_1^2(2-2r)\sigma_m^2 + e_1^2\sigma_\eta^2]}}$, and the cost of options is $E[w_O] = \frac{1}{1-r}\left[\frac{1}{2}\lambda\left[\left(\frac{1}{2}\beta_O^* e_1 \sigma_m\right)^2(1 + (1-r)^2) + \left(\frac{1}{2}\beta_O^* e_1 \sigma_\eta\right)^2\right] + c + \underline{u}\right]$.

Combining the two cases, we can conclude that the cost of options is

$$E[w_O] = \frac{1}{1-r}\left[\frac{1}{2}\lambda\left[\left(\frac{1}{2}\beta_O^* e_1 \sigma_m\right)^2(1-r)^2 + \left(\frac{1}{2}\beta_O^* e_1 \sigma_\eta\right)^2 + \left(\frac{1}{2}\beta_O^* e_1 \sigma_{\min}\right)^2\right] + c + \underline{u}\right],$$

where $\frac{1}{2}\beta_O^* = \frac{2c}{e_1\sigma_{\max} + \sqrt{(e_1\sigma_{\max})^2 - 2c\lambda[e_1^2(1-2r)\sigma_m^2 + e_1^2\sigma_\eta^2 + e_1^2\sigma_{\min}^2]}}$, $\sigma_{\min} = \min(\sigma_m, \sigma_\eta)$, $\sigma_{\max} = \max(\sigma_m, \sigma_\eta)$.

Then it is easy to see that the ratio of the risk premium of one unit of stock to the risk premium of one unit of options is $\kappa(r) = \frac{4[(1-r)^2\sigma_m^2 + \sigma_\eta^2]}{(1-r)^2\sigma_m^2 + \sigma_\eta^2 + \sigma_{\min}^2}$.

Comparing the cost of stock with the cost of options, there exists $\hat{\pi}$ such that all the firms will use stock if and only if $\pi \geq \hat{\pi}$.

$$\hat{\pi} = \frac{1}{4}e_1\sigma_{\max}\sqrt{\kappa(r)}\left[\left(1 + \frac{4Q}{\kappa(r)P}\right) + \left(1 - \frac{4Q}{\kappa(r)P}\right)\sqrt{1 - \frac{4Pc}{e_1^2\sigma_{\max}^2}}\right], \tag{A.11}$$

where $P = \frac{1}{2}\lambda[e_1^2\sigma_m^2(1-2r) + e_1^2\sigma_\eta^2 + e_1^2\sigma_{\min}^2]$, $Q = \frac{1}{2}\lambda[(e_1^2 - e_0^2 - 2re_1(e_1 - e_0))\sigma_m^2 + (e_1^2 - e_0^2)\sigma_\eta^2]$.

Proof of Part 1)

As $\frac{\sigma_\eta}{\sigma_m}$ goes to 0, $\frac{\partial \kappa(r)}{\partial r}$ goes to 0. Thus it can be calculated that

$$\frac{\partial \hat{\pi}}{\partial r} = \frac{e_0 \sigma_m}{2(1+\sqrt{1-2c\lambda(1-2r)})^2 \sqrt{1-2c\lambda(1-2r)}} \left(\sqrt{1-2c\lambda(1-2r)} + 1 - 2\lambda c + 2\lambda c r + \frac{e_0 \lambda c}{e_1} \right) > 0.$$

Proof of Part 2)

Note that $\left(1+\frac{4Q}{\kappa(r)P}\right)+\left(1-\frac{4Q}{\kappa(r)P}\right)\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}$ can be rewritten as

$$D(r) = 1 + \sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}} + \frac{16Qc}{\kappa(r)e_1^2\sigma_{\max}^2} \frac{1}{1+\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}.$$

So the sign of $\frac{\partial \hat{\pi}}{\partial r}$ is the same as the sign of $\frac{\partial \sqrt{\kappa(r)}D(r)}{\partial r} = \frac{1}{2\sqrt{\kappa(r)}}\frac{\partial \kappa(r)}{\partial r}D(r) + \sqrt{\kappa(r)}\frac{\partial D(r)}{\partial r}$. Note that $\frac{\partial \kappa(r)}{\partial r} < 0$ and $\frac{r}{1+\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}$ is increasing in r. We can obtain that

$$\begin{aligned}
\frac{\partial \sqrt{\kappa(r)}D(r)}{\partial r} &< \frac{1}{2\sqrt{\kappa(r)}}\frac{\partial \kappa(r)}{\partial r}\left[1+\frac{16Qc}{\kappa(r)e_1^2\sigma_{\max}^2}\frac{1}{1+\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}\right] \\
&+ \sqrt{\kappa(r)}\left[\frac{2\lambda c\sigma_m^2}{\sigma_{\max}^2}\frac{1}{\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}} - \frac{16Qc}{\kappa(r)^2e_1^2\sigma_{\max}^2}\frac{1}{1+\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}\frac{\partial \kappa(r)}{\partial r}\right] \\
&= \frac{1}{2\sqrt{\kappa(r)}}\frac{\partial \kappa(r)}{\partial r}\left[1-\frac{16Qc}{\kappa(r)e_1^2\sigma_{\max}^2}\frac{1}{1+\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}\right] + \sqrt{\kappa(r)}\frac{2\lambda c\sigma_m^2}{\sigma_{\max}^2}\frac{1}{\sqrt{1-\frac{4Pc}{e_1^2\sigma_{\max}^2}}}
\end{aligned}$$

Since $2 \le \kappa(r) \le 4$, and $Q \le \frac{1}{2}\lambda(e_1^2-e_0^2)(\sigma_m^2+\sigma_\eta^2)$, $\frac{\partial \sqrt{\kappa(r)}D(r)}{\partial r} < 0$ if the following condition is satisfied.

$$\left(\frac{1}{2}-4\lambda c\right)\frac{(1-r)\sigma_m^2\sigma_{\min}^2}{[(1-r)^2\sigma_m^2+\sigma_\eta^2+\sigma_{\min}^2][(1-r)^2\sigma_m^2+\sigma_\eta^2]} > \frac{\lambda c}{\sqrt{1-6\lambda c}}.$$

Suppose that $\sigma_m^2 \ge \sigma_\eta^2$, then the above condition is satisfied if $\frac{(1-r)\sigma_\eta^2}{2[(1-r)^2\sigma_m^2+\sigma_\eta^2]} > X$, where $X = \frac{2\lambda c}{(1-8\lambda c)\sqrt{1-6\lambda c}}$. This can be simplified to $\sigma_\eta^2 > \frac{2X(1-r)^2\sigma_m^2}{(1-r)-4X}$.

Proof of Part 3)

As $\frac{\sigma_m}{\sigma_\eta}$ goes to 0, $\kappa(r) \to 4$, $\frac{P}{e_1^2\sigma_{\max}^2} \to \frac{1}{2}\lambda$, and $\frac{Q}{P} \to \frac{e_1^2-e_0^2}{e_1^2}$. So from (A.11), we can see that $\frac{\partial \hat{\pi}}{\partial r}$ goes to 0 as well.

A.2 Two Firms with RWCs

In the main part of the paper, I have assumed that a continuum of firms and CEOs exist. Each firm represents an infinitesimal portion of the continuum, so the change of one CEO's incentives will not affect other CEOs' actions. This nice property leads to the result that CEO incentives are increasing in a CEO's RWCs r (Proposition 1.3.2), because the rise of a CEO's RWCs increases the CEO's effort and also helps to reduce the risk premium associated with the aggregate shock. However, if a firm's compensation policy affects some other CEOs' actions largely, then the result can be changed. In this section, I consider the case with two firms. If the two CEOs compare compensation with each other, then the increase of one CEO's incentives can increase the other CEO's effort. The rise of the other CEO's effort will increase her own income, which in turn reduces the target CEO's relative payoff because of the RWCs. As a result, the relation between CEO incentives and RWCs may change.

Suppose there are two firms with both CEOs having RWCs: firm 1 and firm 2. I use $\overline{\text{parameters}}$ for firm 2. For example, the measure of a CEO's risk aversion is λ for firm 1 and $\bar{\lambda}$ for firm 2. CEOs' efforts increase both the mean and variance of firm value, which is given by (1.8). CEO 1's utility with RWCs is

$$U = E[u, H] = E[w - H] - \frac{1}{2}\lambda Var[w - H] - \frac{1}{2}a^2,$$

where $H = r\bar{w} = r\left(\bar{\alpha} + \bar{\beta}\left[\bar{\pi}\bar{a} + \bar{a}(\tilde{m} + \bar{\eta})\right]\right)$ denotes CEO 1's RWCs. A larger r means that the CEO 1 cares more about CEO 2's wage. Similarly, CEO 2's utility with RWCs is

$$\bar{U} = E[\bar{u}, \bar{H}] = E[\bar{w} - \bar{H}] - \frac{1}{2}\bar{\lambda} Var[\bar{w} - \bar{H}] - \frac{1}{2}\bar{a}^2,$$

where $\bar{H} = \bar{r}w = \bar{r}\left(\alpha + \beta\left[\pi a + a(\tilde{m} + \eta)\right]\right)$ denotes CEO 2's RWCs. A larger $\bar{r}$ means that the CEO 2 cares more about CEO 1's wage.

The equilibrium is a subgame perfect equilibrium. In the first stage ($t = 0$), each firm will choose an optimal contract given the other firm's contract. In the second stage ($t = 1$), after the contracts are signed, each CEO will exert an optimal effort given the other CEO's contract and effort. I will solve the problem using backward induction.

First, I consider the CEOs' effort decisions. Suppose that CEO 1 is granted β shares of stock, and CEO 2 is granted $\bar{\beta}$ shares of stock. Given CEO 2's effort $\bar{a}$, a similar derivation as in Lemma 1.3.1 shows that CEO 1's effort is given by

$$a = \frac{\beta(\pi + \lambda\sigma_m^2 r\bar{a}\bar{\beta})}{\lambda(\sigma_m^2 + \sigma_\eta^2)\beta^2 + 1}. \tag{A.12}$$

Similarly, given CEO 1's effort a, CEO 2's effort is given by

$$\bar{a} = \frac{\bar{\beta}(\bar{\pi} + \bar{\lambda}\sigma_m^2 \bar{r} a\beta)}{\bar{\lambda}(\sigma_m^2 + \sigma_{\bar{\eta}}^2)\bar{\beta}^2 + 1}. \tag{A.13}$$

I focus on solving the principal-agent problem for firm 1. The solution for firm 2, then, is straight-forward by symmetry.

Lemma A.2.1. *Given the two contracts $w = \alpha + \beta V$ and $\bar{w} = \bar{\alpha} + \bar{\beta}\bar{V}$,*

1) the optimal effort taken by CEO 1 is

$$a = \frac{\beta(\pi + \lambda\sigma_m^2 A)}{\lambda(\sigma_m^2 + \sigma_\eta^2 - \sigma_m^2 B)\beta^2 + 1},$$

where $A = \frac{r\bar{\beta}^2\bar{\pi}}{\bar{\lambda}(\sigma_m^2+\sigma_{\bar{\eta}}^2)\bar{\beta}^2+1}$ and $B = \frac{r\bar{r}\bar{\lambda}\bar{\beta}^2\sigma_m^2}{\bar{\lambda}(\sigma_m^2+\sigma_{\bar{\eta}}^2)\bar{\beta}^2+1}$. The optimal effort a is increasing in r and $\bar{r}$.

2) CEO 2's expected pay increase in CEO 1's incentives, i.e. $\frac{\partial E[\bar{w}]}{\partial \beta} > 0$. Thus, $\frac{\partial E[H]}{\partial \beta} > 0$.

The increase of r increases CEO 1's concerns about CEO 2's payoff; hence it increases CEO 1's incentives to exert a higher (risk-taking) action to catch up with CEO 2's exposure to aggregate shock. Similarly, the increase of $\bar{r}$ increases CEO 2's effort, and thus increases CEO 2's exposure to the aggregate shock. Since CEO 1 cares about CEO 2's payoff, then the increase of CEO 2's exposure to aggregate shock incentivizes CEO 1 to exert a higher effort. For the shareholders in firm 1, when they grant higher incentives to CEO 1, it will increase CEO 1's exposure to aggregate shock, thus it will increase CEO 2's incentives to exert effort. Higher incentives granted to CEO 1 will also induce CEO 2 to exert a higher effort (by (A.13), we see that $\bar{a}$ is increasing in β), which in turn will affect CEO 1's utility through her RWCs. Lemma A.2.1 shows that the rise of CEO 1's incentives increases CEO 2's expected payoff and hence reduces CEO 1's relative payoff. As a result, the shareholders in firm 1 may have to increase CEO pay due to the participation constraint. So the optimal CEO incentives may decrease in RWCs. In other words, RWCs can lead to low CEO incentives, which has been documented in empirical studies (Jensen and Murphy (1990)).

Proposition A.2.2. *The optimal linear contract $w = \alpha^* + \beta^* V$ for firm 1 is that β^* is a solution to the following equation,*

$$\pi - \frac{2\lambda(\sigma^2 - \sigma_m^2 B)\pi\beta^2}{\lambda(\sigma^2 - \sigma_m^2 B)\beta^2 + 1} - \frac{\beta(\pi - \lambda\Gamma\sigma_m^2)}{\lambda\sigma^2\beta^2 + 1} + \left[\left(\frac{(\lambda\sigma_m^2)^2\beta^2}{\lambda\sigma^2\beta^2 + 1} - \lambda\bar{\sigma}^2\right)\Gamma - \bar{\pi}\right]\frac{2B\beta}{\lambda(\sigma^2 - \sigma_m^2 B)\beta^2 + 1} = 0,$$

where $\Gamma = r\bar{a}\bar{\beta} = A + Ba\beta = A + \frac{\beta^2(\pi + \lambda\sigma_m^2 A)B}{\lambda(\sigma_m^2 + \sigma_\eta^2 - \sigma_m^2 B)\beta^2 + 1}$, *$A$ and B are defined in Lemma A.2.1. α^* is set to make the participation constraint binding.*

Moreover, if $(2\lambda\sigma_m^2 + 1)\pi < \left(2 - \frac{\lambda}{\bar{r}\bar{\lambda}}\right)\bar{\pi}$, *then* $\frac{\partial\beta^*}{\partial r}|_{r=0} < 0$. *This implies that if $\bar{\pi}$ and $\bar{r}$ are big, then for small r, β^* is decreasing in r.*

The intuition behind Proposition A.2.2 is that when CEO 1's RWCs increase (i.e., r goes up), it increases CEO 1's effort and also helps to reduce the risk premium associated with aggregate shock. These two effects would push up the optimal incentives for CEO 1. However, if $\bar{r}$ is big, then CEO 2 cares a lot about CEO 1's payoff, then increasing CEO 1's incentives will increase CEO 2's effort significantly. Since the productivity of CEO 2's effort is big (i.e., $\bar{\pi}$ is big), the increase in CEO 2's effort will increase CEO 2's pay greatly. Then because CEO 1 also cares about CEO 2's payoff, the shareholders may have to pay a high premium for CEO 1's RWCs. Thus increasing CEO 1's incentives may lead to a large increase in CEO 1's pay. If this cost is too big, then there will be a negative relation between CEO incentives and r.

Proof of Proposition A.2.2

Given a linear contract $w = \alpha + \beta V$ for CEO 1 and $\bar{w} = \bar{\alpha} + \bar{\beta}\bar{V}$ for CEO 2, CEO 1's effort is $a = \frac{\beta(\pi + \lambda\sigma_m^2 A)}{\lambda(\sigma_m^2 + \sigma_\eta^2 - \sigma_m^2 B)\beta^2 + 1}$ by Lemma A.2.1. Also note that CEO 1's RWCs are

$$H = r\bar{w} = r\bar{\alpha} + \Gamma(\bar{\pi} + \tilde{m} + \tilde{\eta}),$$

where $\Gamma = r\bar{a}\bar{\beta} = A + Ba\beta = A + \frac{\beta^2(\pi + \lambda\sigma_m^2 A)B}{\lambda(\sigma_m^2 + \sigma_\eta^2 - \sigma_m^2 B)\beta^2 + 1}$ and recall that $A = \frac{r\bar{\beta}^2\bar{\pi}}{\bar{\lambda}(\sigma_m^2 + \sigma_{\bar{\eta}}^2)\bar{\beta}^2 + 1}$ and $B = \frac{r\bar{r}\bar{\lambda}\bar{\beta}^2\sigma_m^2}{\bar{\lambda}(\sigma_m^2 + \sigma_{\bar{\eta}}^2)\bar{\beta}^2 + 1}$.

Combining with the binding participation constraint, the shareholders' objective in firm 1 is to maximize

$$\frac{\beta\pi(\pi+\lambda\sigma_m^2 A)}{\lambda(\sigma_m^2+\sigma_\eta^2-\sigma_m^2 B)\beta^2+1}-\frac{1}{2}\frac{\beta^2(\pi-\lambda\Gamma\sigma_m^2)(\pi+\lambda\Gamma\sigma_m^2)}{\lambda(\sigma_m^2+\sigma_\eta^2)\beta^2+1}-\frac{1}{2}\lambda(\sigma_m^2+\sigma_{\bar\eta}^2)\Gamma^2-\bar\pi\Gamma-r\bar\alpha.$$

For brevity, let $\sigma^2=(\sigma_m^2+\sigma_\eta^2)$ and $\bar\sigma^2=(\sigma_m^2+\sigma_{\bar\eta}^2)$. Taking the first-order condition w.r.t β yields that

$$\begin{aligned}&\frac{\pi(\pi+\lambda\sigma_m^2 A)}{\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1}-\frac{2\lambda(\sigma^2-\sigma_m^2 B)\pi(\pi+\lambda\sigma_m^2 A)\beta^2}{[\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1]^2}-\frac{\beta(\pi-\lambda\Gamma\sigma_m^2)(\pi+\lambda\Gamma\sigma_m^2)}{(\lambda\sigma^2\beta^2+1)^2}\\ +\quad&\left[\left(\frac{(\lambda\sigma_m^2)^2\beta^2}{\lambda\sigma^2\beta^2+1}-\lambda\bar\sigma^2\right)\Gamma-\bar\pi\right]\frac{2(\pi+\lambda\sigma_m^2 A)B\beta}{[\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1]^2}=0.\end{aligned}$$

Note that $a=\frac{\beta(\pi+\lambda\sigma_m^2 A)}{\lambda(\sigma_m^2+\sigma_\eta^2-\sigma_m^2 B)\beta^2+1}=\frac{\beta(\pi+\lambda\sigma_m^2\Gamma)}{\lambda\sigma^2\beta^2+1}$, the above equation can be simplified to

$$\pi-\frac{2\lambda(\sigma^2-\sigma_m^2 B)\pi\beta^2}{\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1}-\frac{\beta(\pi-\lambda\Gamma\sigma_m^2)}{\lambda\sigma^2\beta^2+1}+\left[\left(\frac{(\lambda\sigma_m^2)^2\beta^2}{\lambda\sigma^2\beta^2+1}-\lambda\bar\sigma^2\right)\Gamma-\bar\pi\right]\frac{2B\beta}{\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1}=0.$$

Let $F(\beta,r)$ denote the left side of the above equation. Then since $\frac{\partial F}{\partial\beta}<0$, $\frac{\partial\beta}{\partial r}$ has the same sign with $\frac{\partial F}{\partial r}$. We can calculate that

$$\begin{aligned}\frac{\partial F}{\partial r}|_{r=0} &= \frac{2\lambda\pi\beta^2}{(\lambda\sigma^2\beta^2+1)^2}\frac{\bar r\bar\lambda\bar\beta^2(\sigma_m^2)^2}{\bar\lambda\bar\sigma^2\bar\beta^2+1}+\frac{\lambda\sigma_m^2\beta}{\lambda\sigma^2\beta^2+1}\left[\frac{\bar\beta^2\bar\pi}{\bar\lambda\bar\sigma^2\bar\beta^2+1}+\frac{\beta^2\pi}{\lambda\sigma^2\beta^2+1}\frac{\bar r\bar\lambda\bar\beta^2\sigma_m^2}{\bar\lambda\bar\sigma^2\bar\beta^2+1}\right]\\ &+\left(\frac{(\lambda\sigma_m^2)^2\beta^2}{\lambda\sigma^2\beta^2+1}-\lambda\bar\sigma^2\right)\frac{\partial}{\partial r}\left[\frac{2\Gamma B\beta}{\lambda(\sigma^2-\sigma_m^2 B)\beta^2+1}\right]-\frac{2\bar\pi\beta}{\lambda\sigma^2\beta^2+1}\frac{\bar r\bar\lambda\bar\beta^2\sigma_m^2}{\bar\lambda\bar\sigma^2\bar\beta^2+1}\\ &< \frac{\bar r\bar\lambda\bar\beta^2\sigma_m^2}{\bar\lambda\bar\sigma^2\bar\beta^2+1}\frac{\beta}{\lambda\sigma^2\beta^2+1}\left[\frac{2\lambda\sigma_m^2\pi\beta}{\lambda\sigma^2\beta^2+1}+\frac{\bar\pi\lambda}{\bar r\bar\lambda}+\frac{\lambda\sigma_m^2\pi\beta^2}{\lambda\sigma^2\beta^2+1}-2\bar\pi\right]\\ &< \frac{\bar r\bar\lambda\bar\beta^2\sigma_m^2}{\bar\lambda\bar\sigma^2\bar\beta^2+1}\frac{\beta}{\lambda\sigma^2\beta^2+1}\left[2\lambda\sigma_m^2\pi+\frac{\bar\pi\lambda}{\bar r\bar\lambda}+\pi-2\bar\pi\right].\end{aligned}$$

So if $(2\lambda\sigma_m^2+1)\pi<\left(2-\frac{\lambda}{\bar r\bar\lambda}\right)\bar\pi$, then $\frac{\partial F}{\partial r}|_{r=0}<0$. Thus $\frac{\partial\beta}{\partial r}|_{r=0}<0$.

Appendix B

Supplement to Inside Debt

Proof of Equation (2.3)

Rearranging Equations (2.1) and (2.2) yields

$$\begin{aligned} p_R V_{GR} - p_S V_{GS} &\leq (1-p_S)V_{BS} - (1-p_R)V_{BR} \\ p_R V_{GR} - p_S V_{GS} &\leq F(p_R - p_S) + \frac{1}{\alpha}[(p_S - p_R)(\beta F + J) + (1-p_S)\beta V_{BS} - (1-p_R)\beta V_{BR}]. \end{aligned}$$

Equating the right-hand sides of each inequality leads to (2.3). To prove that the denominator of (2.3) is positive, we have:

$$F(p_S - p_R) + (1-p_S)V_{BS} - (1-p_R)V_{BR} = (V_{BS} - V_{BR})(1-p_R) + (p_S - p_R)(F - V_{BS}) > 0.$$

Proof of Proposition 2.2.1

To find the cheapest contract that satisfies (2.3), we first calculate the cost of debt and the bonus. The firm is solvent with probability $p = p_R q + p_S(1-q)$. A bonus of J costs pJ; debt of β costs

$$\beta[pF + q(1-p_R)V_{BR} + (1-q)(1-p_S)V_{BS}].$$

Hence an incentive compatible contract will cost

$$W = pJ + \left[\alpha - \frac{J(p_S - p_R)}{F(p_S - p_R) + (1-p_S)V_{BS} - (1-p_R)V_{BR}}\right][pF + q(1-p_R)V_{BR} + (1-q)(1-p_S)V_{BS}],$$

where

$$\frac{\partial W}{\partial J} = p - \frac{(p_S - p_R)[pF + q(1-p_R)V_{BR} + (1-q)(1-p_S)V_{BS}]}{F(p_S - p_R) + (1-p_S)V_{BS} - (1-p_R)V_{BR}}.$$

Since the derivative is constant, we have a corner solution. The manager is paid entirely with debt if $\frac{\partial W}{\partial J} > 0$, i.e.,

$$p_R(1-p_S)V_{BS} > p_S(1-p_R)V_{BR},$$

and entirely with the bonus if $\frac{\partial W}{\partial J} < 0$.

Proof of Equation (2.6)

Rearranging (2.1) and (2.2) yields

$$\begin{aligned} p_R V_{GR} &\leq p_S V_{GS} + (1-p_S)V_{BS} - (1-p_R)V_{BR} \\ p_R V_{GR} &\leq p_R F - \frac{1}{\alpha} p_R J + \frac{\beta}{\alpha}\left[p_S V_{GS} + (1-p_S)V_{BS} - (1-p_R)V_{BR} - p_R F\right]. \end{aligned}$$

Equating the left-hand sides of each inequality leads to (2.6).

Proof of Lemma 2.3.1

The manager's objective function for effort is given by:

$$\alpha p g e_g + \beta (1-p) b e_b - \frac{1}{2}e_g^2 - \frac{1}{2}e_b^2$$

and differentiating with respect to e_g and e_b gives Equation (2.8).

Proof of Proposition 2.3.2

For conciseness, it is helpful to define the following:

$$\begin{aligned} X &= \frac{(1-p)^2}{p}\frac{(V_{BS}-V_{BR})^2}{V_{GRH}-V_{GS}} \\ Y &= p\left[\frac{V_{GRH}+V_{GS}}{2}-F\right]. \end{aligned}$$

From (2.10), the manager will choose S if and only if

$$V_{GR} < V_{GS} + \frac{\beta(1-p)}{\alpha p}(V_{BS}-V_{BR}).$$

Since $V_{GR} \sim U[V_{GS}, V_{GRH}]$, this occurs with probability $\min\left(1, \frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}\right)$. There are three cases to consider.

Case 1: $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} \leq 1$ *and at least one of* α *and* β *is strictly positive.* Here, both R and S are selected with strictly positive probability, and we need not worry about boundary cases. Then the values of equity and debt are given by:

$$\begin{aligned} E_0 &= p\frac{V_{GRH}+V_{GS}+\frac{\beta(1-p)}{\alpha p}(V_{BS}-V_{BR})}{2} \times \frac{V_{GRH}-V_{GS}-\frac{\beta(1-p)}{\alpha p}(V_{BS}-V_{BR})}{V_{GRH}-V_{GS}} \\ &\quad + pV_{GS} \times \frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} - pF + p^2g^2\alpha \\ &= Y - \frac{\beta^2}{2\alpha^2}X + p^2g^2\alpha. \qquad \text{(B.1)} \\ D_0 &= [pF+(1-p)V_{BR}] \times \frac{V_{GRH}-V_{GS}-\frac{\beta(1-p)}{\alpha p}(V_{BS}-V_{BR})}{V_{GRH}-V_{GS}} \\ &\quad + [pF+(1-p)V_{BS}] \times \frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} + (1-p)^2b^2\beta \\ &= pF+(1-p)V_{BR} + \frac{\beta}{\alpha}X + (1-p)^2b^2\beta. \qquad \text{(B.2)} \end{aligned}$$

Differentiating the objective function (2.7) with respect to β and α yields:

$$(\beta) \quad : \quad -(1-\alpha)\frac{\beta}{\alpha^2}X - D_0 + (1-\beta)\left[\frac{X}{\alpha}+(1-p)^2b^2\right] = 0 \qquad \text{(B.3)}$$

$$(\alpha) \quad : \quad -E_0 + (1-\alpha)\left[\frac{\beta^2}{\alpha^3}X + p^2g^2\right] + (1-\beta)\left[-\frac{\beta}{\alpha^2}X\right] = 0. \qquad \text{(B.4)}$$

Using the expressions for E_0 and D_0 in (B.1) and (B.2), the above first-order conditions become:

$$(\beta) \quad : \quad -[pF+(1-p)V_{BR}]+\left[\frac{1-\beta}{\alpha}-\frac{\beta}{\alpha^2}\right]X+(1-2\beta)(1-p)^2b^2=0 \tag{B.5}$$

$$(\alpha) \quad : \quad -Y+\left[\frac{\beta^2}{2\alpha^2}+\frac{\beta^2}{\alpha^3}-\frac{\beta}{\alpha^2}\right]X+(1-2\alpha)p^2g^2=0. \tag{B.6}$$

Case 2: $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}>1$ *and at least one of* α *and* β *is strictly positive.* In this case, S is always selected. Shareholders' payoff (2.7) is

$$(1-\alpha)(pV_{GS}-pF+p^2g^2\alpha)+(1-\beta)(pF+(1-p)V_{BS}+(1-p)^2b^2\beta).$$

Differentiating this yields

$$\alpha^* \; = \; \max\left(0,\frac{pg^2-(V_{GS}-F)}{2pg^2}\right) \tag{B.7}$$

$$\beta^* \; = \; \frac{(1-p)^2b^2-(pF+(1-p)V_{BS})}{2(1-p)^2b^2}. \tag{B.8}$$

There is no $\max(0,\cdot)$ function for β^* since $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}>1$ rules out $\beta^*=0$.

Case 3: $\alpha=\beta=0$. This must be considered separately from Cases 1 and 2 since the expression $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}$ is undefined. We assume that the manager takes the efficient project in this case.

Proof of part (i)

For Cases 2 and 3, it is immediate that $\alpha^*<\frac{1}{2}$ and $\beta^*<\frac{1}{2}$, so we only need to tackle Case 1.

First, we consider the case in which the optimum is on the boundaries of Case 1, and so the first-order conditions do not apply. The boundaries of Case 1 are $\{(\alpha,\beta):\beta=0,\alpha>0\}$ and $\{(\alpha,\beta):\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}=1\}$. On the boundary $\{(\alpha,\beta):\beta=0,\alpha>0\}$, shareholders' payoff is

$$(1-\alpha)(Y+p^2g^2\alpha)+pF+(1-p)V_{BR}.$$

Differentiating this yields

$$\begin{aligned}\alpha^* &= \frac{p^2g^2 - Y}{2p^2g^2}\\ \beta^* &= 0,\end{aligned}$$

and so $\alpha^* < \frac{1}{2}$ and $\beta^* < \frac{1}{2}$.

The second boundary can be rewritten $\{(\alpha, k) : \frac{k(1-p)}{p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} = 1\}$. Shareholders solve

$$\max_{\alpha,\beta}(1-\alpha)(pV_{GS} - pF + p^2g^2\alpha) + (1-\beta)(pF + (1-p)V_{BS} + (1-p)^2b^2\beta)$$

subject to the constraint

$$\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}} = 1.$$

Defining $A = \frac{p(V_{GRH}-V_{GS})}{(1-p)(V_{BS}-V_{BR})}$, we have

$$\begin{aligned}\alpha^* &= \frac{p^2g^2 - p(V_{GS}-F) + A[(1-p)^2b^2 - pF - (1-p)V_{BS}]}{2[p^2g^2 + A^2(1-p)^2b^2]}\\ \beta^* &= A\alpha^*.\end{aligned}$$

It is automatic that $\alpha^* < \frac{1}{2}$. A necessary condition for the optimum to be on this boundary is that

$$\frac{\partial}{\partial k}\left[(1-\alpha)(Y - \frac{1}{2}k^2X + p^2g^2\alpha) + (1-k\alpha)(pF + kX + (1-p)^2b^2k\alpha)\right] \geq 0$$

at $k = A$. This derivative at $k = A$ equals:

$$\begin{aligned}& -(1-\alpha)kX - \alpha[pF + kX + (1-p)^2b^2k\alpha] + (1-k\alpha)(X + (1-p)^2b^2\alpha)\\ \leq\ & -\alpha[pF + (1-p)(V_{BS}-V_{BR}) - (1-2k\alpha)(1-p)^2b^2]. \qquad \text{(B.9)}\end{aligned}$$

For the derivative to be non-negative, $1-2k\alpha$ must be positive and so $\beta^* = k^*\alpha^* < \frac{1}{2}$.

We now move to the interior of Case 1, which allows us to use first-order conditions. If $\beta \geq \frac{1}{2}$, then from (B.5) we must have $\frac{1-\beta}{\alpha} - \frac{\beta}{\alpha^2} > 0$. This yields

$$\alpha > \frac{\beta}{1-\beta} \geq \frac{1/2}{1-1/2} = 1.$$

This is a contradiction, so $\beta < \frac{1}{2}$.

Similarly, from (B.6), we have

$$\left[-Y + \frac{\beta^2}{2\alpha^2} X\right] + \left[\frac{\beta^2}{\alpha^3} - \frac{\beta}{\alpha^2}\right] X + (1-2\alpha) p^2 g^2 = 0.$$

From (B.1), the first term is the negative of the value of equity if $g = 0$. Since equity value must be positive, this first term must be negative. Thus if $\alpha \geq \frac{1}{2}$, we must have $\frac{\beta^2}{\alpha^3} - \frac{\beta}{\alpha^2} > 0$, i.e. $\alpha < \beta$. However, since $\beta < \frac{1}{2}$, this is inconsistent with $\alpha \geq \frac{1}{2}$. Hence $\alpha < \frac{1}{2}$.

Proof of part (ii)

First, we derive a sufficient condition to rule out Case 3 being optimal. Under Case 3, shareholders' payoff is

$$\frac{V_{GRH} + V_{GS}}{2} + \frac{1}{2} X + (1-p) V_{BR}.$$

To show that Case 3 is suboptimal, it is sufficient to prove that shareholders' payoff is lower than under Case 1 with an arbitrary contract – then it will definitely be lower than under Case 1 with the optimal contract. Consider the contract $\alpha = \beta = \epsilon$. Then, in Case 1, shareholders' payoff is

$$(1-\epsilon)\left[p^2 g^2 \epsilon + (1-p)^2 b^2 \epsilon + \frac{V_{GRH} + V_{GS}}{2} + \frac{1}{2} X + (1-p) V_{BR}\right].$$

The derivative of this function with respect to ϵ at $\epsilon = 0$ is

$$p^2 g^2 + (1-p)^2 b^2 - \frac{V_{GRH} + V_{GS}}{2} - \frac{1}{2} X - (1-p) V_{BR}.$$

Thus, to rule out $\alpha = \beta = 0$, it is sufficient to show that

$$p^2 g^2 + (1-p)^2 b^2 > \frac{V_{GRH} + V_{GS}}{2} + \frac{1}{2} X + (1-p) V_{BR}. \tag{B.10}$$

In addition, if $(1-p)^2b^2-[pF+(1-p)(V_{BS}-V_{BR})]<0$, then β^* in Equation (B.8) is negative and so Case 2 is not feasible. Hence, the optimum must be Case 1. Since Case 1 involves $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}\leq 1$, we have $\alpha^*>0$.

Proof of part (iii)

Condition (B.10) is sufficient to rule out Case 3. Case 2 requires $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}>1$ so $\beta^*>0$ automatically holds. We thus only need to consider Case 1. Differentiating the objective (2.7) with respect to β yields:

$$-\left[pF+(1-p)V_{BR}\right]+\left[\frac{1-\beta}{\alpha}-\frac{\beta}{\alpha^2}\right]X+(1-2\beta)(1-p)^2b^2.$$

At $\beta=0$, this becomes

$$-\left[pF+(1-p)V_{BR}\right]+\frac{1}{\alpha}X+(1-p)^2b^2.$$

Since $\alpha^*<\frac{1}{2}$, $2X>pF+(1-p)V_{BR}-(1-p)^2b^2$ is sufficient to guarantee that this derivative is positive and so $\beta^*>0$.

Proof of part (iv)

First, note that the condition $(1-p)^2b^2-[pF+(1-p)(V_{BS}-V_{BR})]\leq 0$ in part (ii) of the Proposition guarantees that (B.9), the derivative at the boundary where $\frac{\beta(1-p)}{\alpha p}\frac{V_{BS}-V_{BR}}{V_{GRH}-V_{GS}}=1$, is negative. Thus, the optimum is interior and so we can use first-order conditions.

Differentiating the objective (2.7) with respect to β, and treating α as a function of β, yields

$$\begin{aligned}&-Y\frac{\partial\alpha}{\partial\beta}-pF-(1-p)V_{BR}+X\left[\frac{\alpha-\frac{\partial\alpha}{\partial\beta}\beta}{\alpha^2}\left(1-\frac{\beta}{\alpha}\right)+\frac{\beta^2}{2\alpha^2}\frac{\partial\alpha}{\partial\beta}-\frac{\beta}{\alpha}\right]\\ +&\quad(1-2\alpha)p^2g^2\frac{\partial\alpha}{\partial\beta}+\left(1-2\beta\right)(1-p)^2b^2=0.\end{aligned}\tag{B.11}$$

If we solve for α from (B.6), then the solution for α is independent of b. Let h denote the left hand side of (B.11). Thus, the partial derivative of h with respect to b is $2\left(1-2\beta\right)(1-p)^2b$. From $h(\beta,b)=0$, we have $\frac{\partial h}{\partial\beta}\frac{\partial\beta}{\partial b}+\frac{\partial h}{\partial b}=0$; since $\frac{\partial h}{\partial\beta}<0$ at the optimum, the sign of $\frac{\partial\beta}{\partial b}$ is the same as the sign of $\frac{\partial h}{\partial b}$. The latter is positive since $\beta<\frac{1}{2}$. Thus, β^* is increasing in b. A similar analysis proves that α^* is increasing in g.

Proof of parts (v) and (vi)

Defining $k = \frac{\beta}{\alpha}$ and plugging into (B.5) and (B.6) yields the first-order conditions:

$$(\beta): -\left[pF + (1-p)V_{BR}\right] + \left[\frac{1-k\alpha}{\alpha} - \frac{k\alpha}{\alpha^2}\right] X + (1-2\alpha k)(1-p)^2 b^2 = 0 \tag{B.12}$$

$$(\alpha): -Y + \left[\frac{k^2}{2} + \frac{k^2}{\alpha} - \frac{k}{\alpha}\right] X + (1-2\alpha)p^2 g^2 = 0. \tag{B.13}$$

Multiplying (B.3) by k and adding it to (B.4) yields:

$$\alpha E_0 + \beta D_0 = (1-\beta)\beta(1-p)^2 b^2 + (1-\alpha)\alpha p^2 g^2. \tag{B.14}$$

Using the expressions for E_0 and D_0 in (B.1) and (B.2) yields

$$Y\alpha + \frac{1}{2}k^2\alpha X + \left[pF + (1-p)V_{BR}\right] k\alpha = p^2 g^2 \alpha(1-2\alpha) + (1-p)^2 b^2 (k\alpha)(1-2k\alpha), \tag{B.15}$$

and so

$$\alpha = \frac{p^2 g^2 + (1-p)^2 b^2 k - Y - \frac{1}{2}k^2 X - k\left[pF + (1-p)V_{BR}\right]}{2[p^2 g^2 + (1-p)^2 b^2 k^2]}. \tag{B.16}$$

Inserting (B.14) into (2.7) gives the shareholders' objective function as:

$$V_0 - [(1-\beta)\beta(1-p)^2 b^2 + (1-\alpha)\alpha p^2 g^2].$$

Differentiating this with respect to k, and treating α as a function of k, yields:

$$f(k,\theta) = X(1-k) + 2p^2 g^2 \alpha \frac{\partial \alpha}{\partial k} + 2(1-p)^2 b^2 k\alpha \frac{\partial (k\alpha)}{\partial k} = 0. \tag{B.17}$$

(All of these derivatives are being evaluated at $k = k^*$; we suppress "$|_{k=k^*}$" notation for brevity.) Divide the left side of the equation into two parts:

$$\begin{aligned} f_1 &= X(1-k) \\ f_2 &= 2p^2 g^2 \alpha \frac{\partial \alpha}{\partial k} + 2(1-p)^2 b^2 k\alpha \frac{\partial (k\alpha)}{\partial k}. \end{aligned}$$

The first part represents the effect of k on project selection; the second represents the effect on effort, i.e.

$f_2 = \frac{\partial \eta}{\partial k}$. From (B.17), it is easy to see that $k^* < 1$ if and only if $\frac{\partial \eta}{\partial k} < 0$.

We are interested in the relationship between k^* and a parameter θ. From $f(k,\theta) = 0$, we have $\frac{\partial f}{\partial k}\frac{\partial k}{\partial \theta} + \frac{\partial f}{\partial \theta} = 0$. Since $\frac{\partial f}{\partial k} < 0$ at a maximum, the sign of $\frac{\partial k}{\partial \theta}$ is the same as the sign of $\frac{\partial f}{\partial \theta}$. $\frac{\partial f_1}{\partial \theta}$ is simple to calculate; for $\frac{\partial f_2}{\partial \theta}$, we have

$$\frac{\partial f_2}{\partial \theta} = \frac{\partial}{\partial \theta}\left[\frac{\partial}{\partial k}(p^2 g^2 \alpha^2 + (1-p)^2 b^2 (k\alpha)^2)\right] = \frac{\partial}{\partial k}\left[\frac{\partial}{\partial \theta}(p^2 g^2 \alpha^2 + (1-p)^2 b^2 (k\alpha)^2)\right].$$

From (B.16), we have

$$p^2 g^2 \alpha^2 + (1-p)^2 b^2 (k\alpha)^2 = \frac{[p^2 g^2 + (1-p)^2 b^2 k - Y - \frac{1}{2}k^2 X - k\,[pF + (1-p)V_{BR}]]^2}{4[p^2 g^2 + (1-p)^2 b^2 k^2]},$$

and so

$$\frac{\partial f_2}{\partial \theta} = \frac{\partial}{\partial k}\frac{\partial}{\partial \theta}\left(\frac{[p^2 g^2 + (1-p)^2 b^2 k - Y - \frac{1}{2}k^2 X - k\,[pF + (1-p)V_{BR}]]^2}{4[p^2 g^2 + (1-p)^2 b^2 k^2]}\right).$$

Differentiating (B.16) with respect to k gives

$$\begin{aligned}
\frac{\partial \alpha}{\partial k} &= \frac{(1-p)^2 b^2 - kX - [pF + (1-p)V_{BR}]}{2[p^2 g^2 + (1-p)^2 b^2 k^2]} \\
&- \frac{p^2 g^2 + (1-p)^2 b^2 k - Y - \frac{1}{2}k^2 X - k\,[pF + (1-p)V_{BR}]}{2[p^2 g^2 + (1-p)^2 b^2 k^2]^2} \times 2(1-p)^2 b^2 k \\
&= \frac{(1-p)^2 b^2 - kX - [pF + (1-p)V_{BR}]}{2[p^2 g^2 + (1-p)^2 b^2 k^2]} - \frac{2(1-p)^2 b^2 k\alpha}{p^2 g^2 + (1-p)^2 b^2 k^2}. \qquad \text{(B.18)}
\end{aligned}$$

From (B.12) we have:

$$(1-p)^2 b^2 - kX - [pF + (1-p)V_{BR}] = \frac{k-1}{\alpha}X + 2k\alpha(1-p)^2 b^2.$$

Inserting this into (B.18) yields:

$$\frac{\partial \alpha}{\partial k} = \frac{(k-1)X}{2\alpha[p^2 g^2 + (1-p)^2 b^2 k^2]} - \frac{(1-p)^2 b^2 k\alpha}{p^2 g^2 + (1-p)^2 b^2 k^2}. \qquad \text{(B.19)}$$

We commence with the case of $pF + (1-p)V_{BR} > 2(1-p)b^2$, considered in part (v) of the Proposition. (B.12) yields $\frac{1-k\alpha}{\alpha} - \frac{k\alpha}{\alpha^2} > 0$. This implies $1 - k - k\alpha > 0$, and so $k^* < \frac{1}{1+\alpha} < 1$. From (B.19), we have $\frac{\partial \alpha}{\partial k} < 0$.

We now calculate $\frac{\partial f}{\partial \theta}$ for various parameters θ. For $\theta = V_{BS} - V_{BR}$, we have

$$\begin{aligned}\frac{\partial f_1}{\partial \theta} &= \frac{2X}{V_{BS}-V_{BR}}(1-k)\\ \frac{\partial f_2}{\partial \theta} &= -\frac{X}{V_{BS}-V_{BR}}\frac{\partial(\alpha k^2)}{\partial k}.\end{aligned}$$

Hence,

$$\frac{\partial f}{\partial \theta} = \frac{X}{V_{BS}-V_{BR}}\left[2(1-k) - \frac{\partial(\alpha k^2)}{\partial k}\right] = \frac{X}{V_{BS}-V_{BR}}\left[2(1-k-\alpha k) - k^2\frac{\partial \alpha}{\partial k}\right], \tag{B.20}$$

Since $(1-k-\alpha k) > 0$ and $\frac{\partial \alpha}{\partial k} < 0$, this is positive.

For $\theta = g$, $\frac{\partial f}{\partial \theta}$ depends on the sign of

$$(1-2\alpha)\frac{\partial \alpha}{\partial k}.$$

Since $\alpha < \frac{1}{2}$ and $\frac{\partial \alpha}{\partial k} < 0$, this is negative.

For $\theta = p$, it depends on the sign of

$$\begin{aligned}&\frac{1}{2}\frac{\partial X}{\partial p}\left[2(1-k) - \frac{\partial(\alpha k^2)}{\partial k}\right] - 2\alpha(1-p)b^2(1-2\alpha k) - (F - V_{BR})\alpha\\ +\quad & 2\frac{\partial \alpha}{\partial k}\left[(1-2\alpha)pg^2 - (1-p)b^2k(1-2\alpha k) - \frac{1}{2}\left[k(F-V_{BR}) + \left(\frac{V_{GRH}+V_{GS}}{2} - F\right)\right]\right].\end{aligned}$$

From (B.20), we know $2(1-k) - \frac{\partial(\alpha k^2)}{\partial k} > 0$; since also $\beta = \alpha k < \frac{1}{2}$, the first three terms of the above expression are negative. For the whole expression to be negative, it is sufficient to show that $(1-2\alpha)pg^2 - (1-p)b^2k(1-2\alpha k) - \frac{1}{2}\left[k(F-V_{BR}) + \left(\frac{V_{GRH}+V_{GS}}{2} - F\right)\right] > 0$ (since $\frac{\partial \alpha}{\partial k} < 0$). From (B.15), we know that

$$(1-2\alpha)pg^2 = \frac{1}{p}\left(Y + \frac{1}{2}k^2X\right) + \frac{1}{p}\left[pF + (1-p)V_{BR}\right]k - \frac{(1-p)^2}{p}b^2k(1-2k\alpha),$$

and so

$$\begin{aligned}&(1-2\alpha)pg^2 - (1-p)b^2k(1-2\alpha k) - \frac{1}{2}\left[k(F-V_{BR}) + \left(\frac{V_{GRH}+V_{GS}}{2} - F\right)\right]\\ =\quad & \frac{1}{p}\left(\frac{1}{2}Y + \frac{1}{2}k^2X\right) + k\left[\frac{1}{2}F + \frac{1-p}{p}V_{BR} + \frac{1}{2}V_{BR}\right] - \frac{1-p}{p}b^2k(1-2\alpha k).\end{aligned}$$

This is positive, since the first term is positive, and the combination of the second and third terms is positive because $pF + (1-p)V_{BR} > 2(1-p)b^2$. Thus k^* is decreasing in p.

We now turn to the case of $Y > \frac{1}{2}X + p^2 g^2$, considered in part (vi) of the Proposition. From (B.13), we have

$$\left(\frac{k^2}{\alpha} - \frac{k}{\alpha}\right) X = Y - \frac{k^2}{2}X - (1-2\alpha)p^2 g^2.$$

If $k \le 1$, then the right hand side of the above equation is positive, which implies that $\frac{k^2}{\alpha} - \frac{k}{\alpha} > 0$, i.e. $k > 1$. This is a contradiction, so we must have $k > 1$. We also have

$$\begin{aligned}
\frac{\partial(\alpha k)}{\partial k} &= \alpha + k\frac{\partial \alpha}{\partial k} \\
&= \alpha + \frac{k(k-1)X}{2\alpha[p^2 g^2 + (1-p)^2 b^2 k^2]} - \frac{(1-p)^2 b^2 k^2 \alpha}{p^2 g^2 + (1-p)^2 b^2 k^2} \\
&= \frac{k(k-1)X}{2\alpha[p^2 g^2 + (1-p)^2 b^2 k^2]} + \frac{p^2 g^2 \alpha}{p^2 g^2 + (1-p)^2 b^2 k^2} > 0 \text{ when } k > 1. \qquad \text{(B.21)}
\end{aligned}$$

Therefore, $\frac{\partial(\alpha k^2)}{\partial k}$ must be also positive. We also have $\frac{\partial(\alpha k)}{\partial k} > \frac{\partial \alpha}{\partial k}$. For $k > 1$, this is immediate by comparing (B.19) with (B.21). For $k < 1$, note that $\frac{\partial(k\alpha)}{\partial k} > \frac{\partial \alpha}{\partial k}$ if and only if $\alpha > \left(1 - \frac{\beta}{\alpha}\right)\frac{\partial \alpha}{\partial k}$. If $k < 1$, $\frac{\partial \alpha}{\partial k} < 0$ and so the right-hand side of this inequality is negative and less than the left-hand side.

From (B.20), k^* is decreasing in $V_{BS} - V_{BR}$. For $\theta = b$, $\frac{\partial f}{\partial \theta}$ depends on the sign of

$$(1-2\alpha k)\frac{\partial(\alpha k)}{\partial k},$$

which is positive, and so k^* is increasing in b. For $\theta = F$, $\frac{\partial f}{\partial \theta}$ depends on the sign of

$$\frac{\partial}{\partial k}(\alpha(1-k)) = \frac{\partial \alpha}{\partial k} - \frac{\partial(\alpha k)}{\partial k},$$

which is negative, and so k^* is decreasing in F.

Proof of part (vii)

Immediate from Equations (B.1) and (B.2).

Appendix C

Supplement to Incentive Contracting under Ambiguity-Aversion

Discussion of First-Order Approach

In general, the IC constraint cannot be replaced by the first-order condition. A sufficient condition for the validity of first-order approach (FOA) is that the agent's objective function is concave. Rogerson (1985) provides a sufficient condition for the validity of FOA: If MLRP (Monotone Likelihhood Ratio Property), together with CDFC (Convexity of the Distribution Function Condition), hold then the FOA is valid. But unfortunately, CDFC together with MLRP are very restrictive, it is even not true for normal distribution. Hence here I will provide some conditions under which the FOA is valid. Also these conditions are applicable to Section 3.2.2 and 3.2.3.

Lemma C.0.1. *1) If the manager's contract consists of a base salary and stock, then his utility function is concave in a. Thus the FOA is valid;*

2) Given l and $k < a + l$, the FOA is valid when σ is small;

3) Given k and σ, the FOA is valid when l is large.

Proof. Part 1):

The the mean and variance of stock are $m_{-\infty} = \hat{a} + l$ and $\sigma^2_{-\infty} = \sigma^2$, so with a stock-based contract, the manager's expected utility is

$$E[u] = \alpha + \beta_{-\infty}(\hat{a} + l) - \frac{1}{2}\lambda\beta_{-\infty}^2\sigma^2 - \frac{1}{2}\hat{a}^2,$$

which is obviously concave in $\hat{a}$.

Part 2):

For options with exercise price k, the manager's expected utility is

$$E[u] = \alpha + \beta_k m_k - \frac{1}{2}\lambda\beta_k^2\sigma_k^2 - \frac{1}{2}\hat{a}^2$$

Taking partial derivative w.r.t $\hat{a}$ yields that

$$\frac{\partial E[u]}{\partial \hat{a}} = \beta_k(1-\Phi) - \lambda\beta_k^2\Phi m_k - \hat{a}.$$

Similarly, the second derivatve of $E[u]$ w.r.t $\hat{a}$ is

$$\frac{\partial^2 E[u]}{\partial \hat{a}^2} = \beta_k\frac{\phi}{\sigma} - \lambda\beta_k^2\left[-\frac{\phi}{\sigma}m_k + \Phi(1-\Phi)\right] - 1.$$

If $k < \underline{a} + l$, then we have

$$\begin{aligned}
& \beta_k\frac{\phi}{\sigma} - \lambda\beta_k^2\left[-\frac{\phi}{\sigma}m_k + \Phi(1-\Phi)\right] - 1 \\
\leq\ & \beta_k\frac{\phi}{\sigma} + \lambda\beta_k^2\frac{\phi}{\sigma}m_k - 1 \\
\leq\ & \left[\beta_k + \lambda\beta_k^2(\bar{a}+l)\right]\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(k-\underline{a}-l)^2}{\sigma^2}\right) - 1 \\
\to\ & -1,\ \text{as}\ \sigma \to 0.
\end{aligned}$$

Thus when σ is small, then the second derivative is negative, which implies that the manager's objective is concave.

If $\underline{a}+l \leq k < a+l$, then given any ϵ, for any $\hat{a} \in [\underline{a}, k-l-\epsilon) \cup (k-l+\epsilon, \bar{a}]$, we have

$$\begin{aligned} & \beta_k \frac{\phi}{\sigma} - \lambda \beta_k^2 \left[-\frac{\phi}{\sigma} m_k + \Phi(1-\Phi) \right] - 1 \\ \leq \quad & \beta_k \frac{\phi}{\sigma} + \lambda \beta_k^2 \frac{\phi}{\sigma} m_k - 1 \\ \leq \quad & \left[\beta_k + \lambda \beta_k^2 (\bar{a}+l) \right] \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{\epsilon^2}{\sigma^2} \right) - 1 \\ \rightarrow \quad & -1, \text{ as } \sigma \rightarrow 0. \end{aligned}$$

On the other hand, note that $\beta_k \geq a$ by Proposition 3.2.2, so

$$\frac{\partial E[u]}{\partial \hat{a}}|_{\hat{a}=k-l-\epsilon} \rightarrow \beta_k + l + \epsilon - k > 0 \text{ as } \sigma \rightarrow 0.$$

Hence given a small ϵ, when σ is small, the manager's utility $E[u]$ is concave in $\hat{a}$ for $\hat{a} \in [\underline{a}, k-l-\epsilon) \cup$ $(k-l+\epsilon, \bar{a}]$ and increasing in $\hat{a}$ for $\hat{a} \in [\underline{a}, k-l-\epsilon)$. Also note that as $\sigma \rightarrow 0$,

$$E[u]|_{\hat{a}=k-l} = \alpha - \frac{1}{2}(k-l)^2 < \alpha + \beta_k l + \frac{1}{2}\beta_k^2 = E[u]|_{\hat{a}=a}.$$

Therefore, when ϵ is small,

$$E[u]|_{\hat{a} \in [k-l-\epsilon, k-l+\epsilon]} < E[u]|_{\hat{a}=a}.$$

Hence the maximum must be achieved in $(k-l+\epsilon, \bar{a}]$. Since $E[u]$ is concave in $(k-l+\epsilon, \bar{a}]$, the FOA is valid in this case.

Part 3)

When l is large, we have

$$\begin{aligned} & \beta_k \frac{\phi}{\sigma} - \lambda \beta_k^2 \left[-\frac{\phi}{\sigma} m_k + \Phi(1-\Phi) \right] - 1 \\ \leq \quad & \beta_k \frac{\phi}{\sigma} + \lambda \beta_k^2 \frac{\phi}{\sigma} m_k - 1 \\ \leq \quad & \left[\beta_k + \lambda \beta_k^2 (\bar{a}+l) \right] \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(k-a-l)^2}{\sigma^2} \right) - 1 \\ \rightarrow \quad & -1, \text{ as } l \rightarrow \infty. \end{aligned}$$

Thus, when l is large, the manager's utility is concave, and so the FOA is valid. $\square$

In Proposition 3.2.4, we need $k < a + l$ and a small σ_1 to justify the use of options. These conditions coincide with the conditions in Part 2). In Propostion 3.2.7, we need a large l_2 to justify the use of options, which also coincides with the condition in Part 3).

Proof of Proposition 3.2.2

Taking the partial derivtive of m_k and σ_k^2 w.r.t a yields that

$$\begin{aligned} \frac{\partial m_k}{\partial a} &= 1 - \Phi \\ \frac{\partial \sigma_k^2}{\partial a} &= 2\Phi m_k \end{aligned}$$

Then from the first-order condition (3.4), we can solve for the number of options β_k that are needed to induce the target effort a. That is

$$\beta_k = \frac{(1-\Phi) \pm \sqrt{\Delta}}{2\lambda\Phi m_k},$$

where $\Delta = (1-\Phi)^2 - 4a\lambda\Phi m_k$.

The shareholders will choose β_k optimally (i.e. as small as possible), so we have

$$\beta_k = \frac{(1-\Phi) - \sqrt{\Delta}}{2\lambda\Phi m_k} = \frac{2a}{(1-\Phi) + \sqrt{\Delta}}. \tag{C.1}$$

So the cost c_k is

$$c_k = \frac{1}{2}\lambda\beta_k^2\sigma_k^2 = \frac{2\lambda a^2\sigma^2 y_1}{(1+\sqrt{1-4a\lambda\sigma y_2})^2}$$

Then it can be shown that $y_1'(\eta) \geq 0$ and $y_2'(\eta) \geq 0$ (see Feltham and Wu (2001)). Note that $\frac{\partial\eta}{\partial k} = \frac{1}{\sigma} > 0$, so we have $\frac{\partial c_k}{\partial k} > 0$. Thus the cheapest way to induce the target effort a is using the stock.

Similarly, if $k \leq a + l$, then $\frac{\partial\eta}{\partial\sigma} = -\frac{k-a-l}{\sigma^2} \geq 0$ hence, c_k is increasing in σ in this case.

Note that from (C.1), we have

$$\beta_k = \frac{2a/(1-\Phi)}{1+\sqrt{1-4a\lambda\sigma y_2}}.$$

Thus it is straight-forward to see that β_k is increasing in λ and k, and $\beta_k \geq a$.

Proof of Lemma 3.2.3

The manager's expected utility is

$$E[u] = \alpha + \beta_k m_k - \frac{1}{2}\lambda\beta_k^2\sigma_k^2 - \frac{1}{2}a^2$$

In the case of low probability of bankruptcy, the mean and variance of stock are $m_{-\infty} = a + l$ and $\sigma_{-\infty}^2 = \sigma^2$. Thus we have

$$\frac{\partial E[u]}{\partial \sigma}|_{k=0} = -\beta_{-\infty}^2 \lambda\sigma < 0.$$

So the manager will always perceive high risk σ_2^2 if he is granted stock.

Then for options, let

$$\begin{aligned} A_1 &= \{\hat{a} : E[u(\hat{a})]|_{\sigma_m=\sigma_1} < E[u(\hat{a})]|_{\sigma_m=\sigma_2}\}, \\ A_2 &= \{\hat{a} : E[u(\hat{a})]|_{\sigma_m=\sigma_2} \le E[u(\hat{a})]|_{\sigma_m=\sigma_1}\}, \end{aligned}$$

where A_1 is the set of efforts for which the manager's expected utility is minimized under low risk and A_2 is the complement set of A_1. Then the manager's objective (3.7) can be rewritten as

$$\max\left(\max_{\hat{a}\in A_1} E[u(\hat{a})]|_{\sigma_m=\sigma_1}, \max_{\hat{a}\in A_2} E[u(\hat{a})]|_{\sigma_m=\sigma_2}\right).$$

By the condition (3.8), we have

$$\begin{aligned} & \alpha + \beta_k m_k(a,\sigma_1) - \frac{1}{2}\lambda\beta_k^2\sigma_k^2(a,\sigma_1) - \frac{1}{2}a^2 \\ < \;& \alpha + \beta_k m_k(a,\sigma_2) - \frac{1}{2}\lambda\beta_k^2\sigma_k^2(a,\sigma_2) - \frac{1}{2}a^2, \end{aligned}$$

so the target effort a belongs to the set A_1, i.e. $a \in A_1$. Note that $\beta_k = \beta_k(a,\sigma_1)$ is exactly the number of options to induce the effort a if the manager's perception of risk is σ_1. Hence we have

$$a \in \arg\max_{\hat{a}\in A_1\cup A_2} E[u]|_{\sigma_m=\sigma_1}.$$

On the other hand, for any $\hat{a} \in A_2$, we have

$$E[u(\hat{a})]|_{\sigma_m=\sigma_2} \leq E[u(\hat{a})]|_{\sigma_m=\sigma_1} < E[u(a)]|_{\sigma_m=\sigma_1}.$$

Therefore at the optimum, the manager will perceive low risk σ_1 and exert the target effort $a \in A_1$.

Proof of Proposition 3.2.4

Note that if $k < a + l$,

$$\lim_{\sigma_1 \to 0} \frac{2}{\beta_k(a,\sigma_1)} \frac{m_k(a,\sigma_2) - m_k(a,\sigma_1)}{\sigma_k^2(a,\sigma_2) - \sigma_k^2(a,\sigma_1)} = \frac{2[m_k(a,\sigma_2) - (a + l - k)]}{a\sigma_k^2(a,\sigma_2)},$$

and

$$\lim_{\sigma_1 \to 0} \frac{2\beta_k(a,\sigma_1)[m_k(a,\sigma_2) - m_k(a,\sigma_1)]}{a^2\sigma_2^2 - \beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)} = \frac{2[m_k(a,\sigma_2) - (a + l - k)]}{a\sigma_2^2}.$$

Then by (3.9), there must exist a cut-off $\hat{\sigma}(\lambda)$ such that when $\sigma_1 < \hat{\sigma}(\lambda)$, we have

$$\frac{2\beta_k(a,\sigma_1)[m_k(a,\sigma_2) - m_k(a,\sigma_1)]}{a^2\sigma_2^2 - \beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)} < \lambda < \frac{2}{\beta_k(a,\sigma_1)} \frac{m_k(a,\sigma_2) - m_k(a,\sigma_1)}{\sigma_k^2(a,\sigma_2) - \sigma_k^2(a,\sigma_1)}.$$

Then by Lemma 3.2.3, the manager will perceive low risk σ_1 with the options of exercise price k and perceive high risk σ_2 with stock. Then we can derive that the cost of using stock is

$$\hat{c}_{-\infty} = \frac{1}{2}\lambda a^2\sigma_2^2 + a\,[m_{-\infty}(a,\sigma_s) - m_{-\infty}(a,\sigma_2)] = \frac{1}{2}\lambda a^2\sigma_2^2,$$

where the second equality is derived from the fact that $m_{-\infty}(a,\sigma_s) = m_{-\infty}(a,\sigma_2) = a + l$ because of low probability of bankruptcy. The cost of using options is

$$\begin{aligned}
\hat{c}_k &= \frac{1}{2}\lambda\beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1) + \beta_k(a,\sigma_1)\,[m_k(a,\sigma_s) - m_k(a,\sigma_1)] \\
&\leq \frac{1}{2}\lambda\beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1) + \beta_k(a,\sigma_1)\,[m_k(a,\sigma_2) - m_k(a,\sigma_1)].
\end{aligned}$$

Then since the condition

$$\lambda > \frac{2\beta_k(a,\sigma_1)[m_k(a,\sigma_2) - m_k(a,\sigma_1)]}{a^2\sigma_2^2 - \beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)}$$

implies that

$$\frac{1}{2}\lambda\beta_k^2(a,\sigma_1)\sigma_k^2(a,\sigma_1)+\beta_k(a,\sigma_1)\left[m_k(a,\sigma_2)-m_k(a,\sigma_1)\right]<\frac{1}{2}\lambda a^2\sigma_2^2,$$

we obtain that

$$\hat{c}_k(a,\sigma_1)<\hat{c}_{-\infty}(a,\sigma_2),$$

i.e. it is cheaper to implement a by using the options.

Proof of Lemma 3.2.6

If the manager is granted stock, then his expected utility is

$$E[u]=\alpha+\beta_{-\infty}(a+l)-\frac{1}{2}\lambda\beta_{-\infty}^2\sigma^2-\frac{1}{2}a^2,$$

which is obviously minimized at $l=l_1$. Hence the manager will perceive l_1 with stock.

For options, let

$$\begin{aligned} A_1 &= \{\hat{a}: E[u(\hat{a})]|_{l_m=l_1}\le E[u(\hat{a})]|_{l_m=l_2}\},\\ A_2 &= \{\hat{a}: E[u(\hat{a})]|_{l_m=l_2}< E[u(\hat{a})]|_{l_m=l_1}\},\end{aligned}$$

where A_2 is the set of efforts for which the manager's expected utility is minimized under high expected value of l and A_1 is the complement set of A_2. Then the manager's objective can be rewritten as

$$\max\left(\max_{\hat{a}\in A_1}E[u(\hat{a})]|_{l_m=l_1},\max_{\hat{a}\in A_2}E[u(\hat{a})]|_{l_m=l_2}\right).$$

By the condition (3.10), we have

$$\begin{aligned} &\alpha+\beta_k m_k(a,l_1)-\frac{1}{2}\lambda\beta_k^2\sigma_k^2(a,l_1)-\frac{1}{2}a^2\\ > \;&\alpha+\beta_k m_k(a,l_2)-\frac{1}{2}\lambda\beta_k^2\sigma_k^2(a,l_2)-\frac{1}{2}a^2,\end{aligned}$$

so the target effort a belongs to the set A_2, i.e. $a\in A_2$. Note that $\beta_k=\beta_k(a,l_2)$ is exactly the number of options to induce the effort a if the manager's perception of the return l is l_2. Hence we have

$$a \in \arg\max_{\hat{a}\in A_1\cup A_2} E[u]|_{l_m=l_2}.$$

On the other hand, for any $\hat{a} \in A_1$, we have

$$E[u(\hat{a})]|_{l_m=l^1} \leq E[u(\hat{a})]|_{l_m=l_2} < E[u(a)]|_{l_m=l_2}.$$

Therefore at the optimum, the manager will perceive l_2 and exert the target effort $a \in A_2$.

Proof of Proposition 3.2.7

By Lemma 3.2.6, the left inequality in (3.12) implies that the manager will perceive high value l_2 if he is granted options with exercise price k. The right inequality in (3.12) will imply that $\bar{c}_{-\infty}(a,l_1) > \bar{c}_k(a,l_2)$, i.e. the cost of the stock-based contract is larger than the cost of the option-based contract. The proof is as follows:

If $l_s = l_1$, then the cost of the stock-based contract is

$$\bar{c}_{-\infty}(a,l_1) = \frac{1}{2}\lambda a^2\sigma^2,$$

and the cost of the option-based contract is

$$\bar{c}_k(a,l_2) = \frac{1}{2}\lambda\beta_k(a,l_2)^2\sigma_k^2(a,l_2) + \beta_k(a,l_2)[m_k(a,l_1) - m_k(a,l_2)].$$

Then $\bar{c}_{-\infty}(a,l_1) > \bar{c}_k(a,l_2)$ is equivalent to

$$\lambda < \frac{2\beta_k(a,l_2)[m_k(a,l_2) - m_k(a,l_1)]}{\beta_k^2(a,l_2)\sigma_k^2(a,l_2) - a^2\sigma^2}.$$

Note that $\beta_k(a,l_2) \geq a$, so the right inequality in (3.12) is sufficient to ensure that $\bar{c}_{-\infty}(a,l_1) > \bar{c}_k(a,l_2)$.

The proof for $l_s = l_2$ is similar.

Proof of Lemma 3.2.8

The cost of the non-indexed stock compensation has been derived in (3.11). For the cost of the indexed stock or option compensation, since the ambiguity has been removed from the manager's payoff, it is the same as the case in the basic model. But since the risk associated with the market return is also removed from the manager's payoff, the cost is just c_k in (3.6) which is evaluated at the variance $\sigma^2 - \sigma_l^2$. In particular, the cost of the indexed stock is $c_0(\sigma^2 - \sigma_l^2) = \frac{1}{2}\lambda a^2(\sigma^2 - \sigma_l^2)$.

Proposition 3.2.2 shows that $c_0(\sigma^2 - \sigma_l^2) < c_k(\sigma^2 - \sigma_l^2)$. It is also obvious that $\frac{1}{2}\lambda a^2(\sigma^2 - \sigma_l^2) < \frac{1}{2}\lambda a^2\sigma^2 + a[m_{-\infty}(a, l_s) - m_{-\infty}(a, l_1)]$. So it is cheaper to implement the target effort using indexed stock than using indexed options or non-indexed stock.

Proof of Proposition 3.2.9

1) If $l_s = l_2$, then the cost of using non-indexed options is

$$\bar{c}_k(a, l_2) = \frac{1}{2}\lambda\beta_k(a, l_2)^2\sigma_k^2(a, l_2),$$

which is greater $\frac{1}{2}\lambda a^2\sigma^2$ by Proposition 3.2.2. So indexed stock is better than non-indexed options.

2) If $l_s = l_1$, then the cost of using non-indexed options is

$$\bar{c}_k(a, l_2) = \frac{1}{2}\lambda\beta_k(a, l_2)^2\sigma_k^2(a, l_2) + \beta_k(a, l_2)[m_k(a, l_1) - m_k(a, l_2)].$$

Following the proof of Proposition 3.2.7, we can show that it is optimal to grant the manager non-indxed options.

Bibliography

[1] Abel, Andrew, 1990, “Asset Prices under Habit Formation and Catching Up with the Jones,” *American Economic Review*, 80, 38-42.

[2] Agrawal, Anup and Nandu Nagarajan, 1990, “Corporate Capital Structure, Agency Costs, and Ownership Control: The Case of All-Equity Firms, ” *Journal of Finance*, 45, 1325-1331.

[3] Aggarwal, Rajesh and Andrew Samwick, 2006, “Empire-Builders and Shirkers: Investment, Firm-Performance, and Managerial Incentives, ” *Journal of Corporate Finance*, 12, 489-515.

[4] Ahn, David, Syngjoo Choi, Douglas Gale and Shachar Kariv, 2009, “Estimating Ambiguity-Aversion in a Portfolio Choice Experiment,” Working Paper.

[5] Bebchuk, Lucian and Robert Jackson, 2005, “Executive Pensions, ” *Journal of Corporation Law*, 30, 823-855.

[6] Bebchuk, Lucian and Holger Spamann, 2009 “Regulating Bankers’ Pay, ” *Georgetown Law Journal*, forthcoming.

[7] Bereskin, Frederick L. and David C. Cicero, 2011, “CEO Compensation Contagion: Evidence from an Exogenous Shock,” Working Paper.

[8] Bertrand, Marianne and Sendhil Mullainathan, 2001, “Are CEOs Rewarded for Luck? The Ones Without Principals Are, ” *Quarterly Journal of Economics*, 116, 901-932.

[9] Bettis, J. Carr, John M. Bizjak and Michael L. Lemmon, 2005, “Exercise Behavior, Valuation, and the Incentives Effects of Employee Stock Options,” *Journal of Financial Economics* 76, 445-470.

[10] Biais, Bruno and Catherine Casamatta, 1999, “Optimal Leverage and Aggregate Investment, ” *Journal of Finance*, 54, 1291-1323.

[11] Bizjak, John M., James A. Brickley and Jeffrey L. Coles, 1993, “Stock-based Incentive Compensation and Investment Behavior,” *Journal of Accounting and Economics* 16, 349-372.

[12] Bizjak, John M., Michael Lemmon and Lalitha Naveen, 2008, “Has the Use Peer Groups Contributed to Higher Pay and Less Efficient Compensation,” *Journal of Financial Economics* 90, 152-168.

[13] Bolton, Patrick, José Scheinkman and Wei Xiong, 2006, “Executive Compensation and Short-Termist Behaviour in Speculative Markets,” *Review of Economic Studies* 73, 577-610.

[14] Bouwman, Christa H.S., 2011, “The Geography of Executive Compensation,” Working Paper.

[15] Brander, James and Michel Poitevin, 1992, “Managerial Compensation and The Agency Costs of Debt Finance, ” *Managerial and Decision Economics*, 13, 55-64.

[16] Cabrales, Antonio, Antoni Calvó-Armengol and Nicola Pavoni, 2007, “Social Preferences, Skill Segregation, and Wage Dynamics,” *Review of Economic Studies* 74, 1-33.

[17] Calcagno, Riccardo and L. Renneboog, 2007, “The Incentive to Give Incentives: On the Relative Seniority of Debt Claims and Managerial Compensation, ” *Journal of Banking and Finance*, 31, 1795-1815.

[18] Campbell, John Y. and John H., Cochrane, 1999, “By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 107, 205-251.

[19] Carpenter, Jennifer, 2000, “Does Option Compensation Increase Managerial Risk Appetite,” *Journal of Finance* 55, 2311-2331.

[20] Coles, Jeffrey L., Naveen D. Daniel and Lalitha Naveen, 2006, “Managerial Incentives and Risk-Taking,” *Journal of Financial Economics* 79, 431-468.

[21] Constantinides, George M., 1990, “Habit Formation: A Resolution of the Equity Premium Puzzle,” *Journal of Political Economy*, 98, 519-543.

[22] DeFusco, Richard, Johnson Robert and Thomas Zorn, 1990, "The Effect of Executive Stock Option Plans On Stockholders and Bondholders, " *Journal of Finance*, 45, 617-627.

[23] DeMarzo, Peter M., Ron Kaniel and Ilan Kremer, 2008, "Relative Wealth Concerns and Financial Bubbles," *Review of Financial Studies* 21, 19-50.

[24] Diamond, Douglas, 1989, "Reputation acquisition in debt markets, " *Journal of Political Economy*, 97, 828-62.

[25] Dittmann, Ingolf and Ernst Maug, 2007, "Lower Salaries and No Options? On the Optimal Structure of Executive Pay, " *Journal of Finance* 62, 303-343.

[26] Dittmann, Ingolf, Ernst Maug and Oliver Spalt, 2010, "Sticks or Carrots? Optimal CEO Compensation When Managers Are Loss-Averse, " *Journal of Finance*, forthcoming.

[27] Dittmann, Ingolf and Ko-Chia Yu, 2008, "Providing Effort and Risk-Taking Incentives: The Optimal Compensation Structure for CEOs," Working Paper.

[28] Dybvig, Philip and Jaime Zender, 1991, "Capital Structure and Dividend Irrelevance with Asymmetric Information, " *Review of Financial Studies*, 4, 201-219.

[29] Easley, David and Maureen O'Hara, 2009, "Ambiguity and Nonparticipation: The Role Regulation," *Review of Financial Studies* 22, 1818-1843.

[30] Easley, David and Maureen O'Hara, 2010, "Liquidity and Valuation in an Uncertain World," *Journal of Financial Economics* 97, 1-11.

[31] Edmans, Alex and Xavier Gabaix, 2011, "The Effect of Risk on the CEO Market," *Review of Financial Studies* 24, 2822-2863.

[32] Edmans, Alex and Xavier Gabaix, 2011, "Tractability in Incentive Contracting," *Review of Financial Studies* 24, 2865-2894.

[33] Ellsberg, D., 1961, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics* 75, 643-669.

[34] Epstein, Larry G. and Martin Schneider, 2008, "Ambiguity, Information Quality, and Asset Pricing," *Journal of Finance* 63, 197-228.

[35] Epstein, Larry G. and Tan Wang, 1994, "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica* 62, 283-322.

[36] Faulkender, Michael and Jun Yang, 2010, "Inside the Black Box: The Role and Composition of Compensation Peer Groups," *Journal of Financial Economics* 96, 257-270.

[37] Feltham, Gerald A. and Martin G. H. Wu, 2001, "Incentive Efficiency of Stock versus Options," *Review of Accounting Studies* 6, 7-28.

[38] Friend, Irwin and Larry Lang, 1988, "An Empirical Test of The Impact of Managerial Self-Interest on Corporate Capital Structure, " *Journal of Finance*, 43, 271-281.

[39] Gabaix, Xavier and Augustin Landier, 2008, "Why has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123, 49-100.

[40] Galí, Jordi, 1994, "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices," *Journal of Money, Credit and Banking*, 26, 1-8.

[41] García, Diego and Günter Strobl, 2011, "Relative Wealth Concerns and Complementarities in Information Acquisition," *Review of Financial Studies* 24, 169-207.

[42] Gerakos, Joseph, 2007, "CEO Pensions: Disclosure, Managerial Power, and Optimal Contracting, " Working Paper.

[43] Gibbons, Robert and Kevin Murphy, 1992, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence, " *Journal of Political Economy*, 100, 468-505.

[44] Gilboa, Itzchak and David Schmeidler, 1989, "Maxmin Expected Utility with Nonunique Prior," *Journal of Mathematical Economics* 18, 141-153.

[45] Goel, Anand M. and Anjan V. Thakor, 2010, "Do Envious CEOs Cause Merger Waves?" *Review of Financial Studies* 23, 487-517.

[46] Gopalan, Radhakrishnan, Todd T. Milbourn and Fenghua Song, 2010, "Strategic flexibility and the optimality of pay for sector performance," *Review of Financial Studies* 23, 2060-2098.

[47] Gormley, Todd A., David A. Matsa and Todd Milbourn, 2011, "CEO Compensation and Corporate Risk-Taking: Evidence from a Natural Experiment," Working Paper.

[48] Grossman, Sanford and Oliver Hart, 1983, "An Analysis of the Principal-Agent Problem," *Econometrica* 51, 7-45.

[49] Guay, Wayne R., 1999, "The Sensitivity of Wealth to Equity Risk: An Analysis of the Magnitude and Determinants," *Journal of Financial Economics*, 53, 43-71.

[50] Habib, Michel and Bruce Johnsen, 2000, "The Private Placement of Debt and Outside Equity as An Information Revelation Mechanism, " *Review of Financial Studies*, 13, 1017-1055.

[51] Hayes, Rachel M., Michael Lemmon and Mingming Qiu, 2010, "Stock Options and Managerial Incentives for Risk-Taking: Evidence from FAS 123R," Working Paper.

[52] Hellwig, Martin, 2009, "A Reconsideration of The Jensen-Meckling Model of Outside Finance, " *Journal of Financial Intermediation*, 18, 495-525.

[53] Hirshleifer, David and Anjan Thakor, 1992, "Managerial Conservatism, Project Choice, and Debt, " *Review of Financial Studies*, 5, 437-470.

[54] Holmstrom, Bengt and Paul Milgrom, 1987, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica* 55, 303-328.

[55] Holmstrom, Bengt and Jean Tirole, 1993, "Market Liquidity and Performance Monitoring," *Journal of Political Economy* 101, 678-709.

[56] Illeditsch, Philipp, 2010, "Ambiguous Information, Portfolio Inertia, and Excess Volatility," Working Paper.

[57] Innes, Robert, 1990, "Limited Liability and Incentive Contracting with Ex-Ante Action Choices, " *Journal of Economic Theory*, 52, 45-67.

[58] Jensen, Michael and William Meckling, 1976, "Theory of The Firm: Managerial Behavior, Agency Costs, and Capital Structure, " *Journal of Financial Economics*, 3, 305-360.

[59] Jensen, Michael and Kevin Murphy, 1990, "Performance Pay and Top-Management Incentives," *Journal of Political Economy* 98, 225-264.

[60] Jin, Li, 2002, "CEO Compensation, Diversification, and Incentives," *Journal of Financial Economics* 66, 29-63.

[61] John, Teresa and Kose John, 1993, "Top-Management Compensation and Capital Structure, " *Journal of Finance*, 48, 949-974.

[62] Miglietta, Salvatore, 2008, "Incentive Contracts and Status-Concerned Agents," Working Paper.

[63] Miglietta, Salvatore, 2010, "Incentives and Relative-Wealth Concerns: Theory and Evidence," Working Paper.

[64] Murphy, Kevin, 1999, "Executive Compensation ," in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics* (New York: Elsevier/North-Holland).

[65] Myers, Stewart and Nicholas Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics*, 13, 187-221.

[66] Noe, Thomas H. and Michael J. Rebello, 2008, "To Each According to Her Luck and Power: Optimal Corporate Governance and Compensation Policy in a Dynamic World," Working Paper.

[67] Ortiz-Molina, Hernan, 2007, "Executive Compensation and Capital Structure: The Effects of Convertible Debt and Straight Debt, " *Journal of Accounting and Economics*, 43, 69-93.

[68] Ou-Yang, Hui, 2003, "Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem," *Review of Financial Studies* 16, 173-208.

[69] Oyer, Paul, 2004, "Why Do Firms Use Incentives That Have No Incentive Effects," *Journal of Finance* 59, 1619-1650.

[70] Prendergast, Canice, 2002, “The Tenuous Trade-Off between Risk and Incentives,” *Journal of Political Economy*, 110, 1071-1102.

[71] Ross, Stephen, 2004, “Incentives and the Duality of Risk Aversion and Riskiness,” *Journal of Finance* 59, 207-225.

[72] Roussanov, Nikolai, 2010, “Diversification and Its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status,” *Journal of Finance*, 65, 1755-1788.

[73] Shue, Kelly, 2011, “Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers,” Working Paper.

[74] Stoughton, Neal and Eli Talmor, 1999, “Managerial Bargaining Power in The Determination of Compensation Contracts and Corporate Investment, ” *International Economic Review*, 40, 69-93.

[75] Sundaram, Rangarajan and David Yermack, 2007, “Pay Me Later: Inside Debt and Its Role in Managerial Compensation, ” *Journal of Finance*, 62, 1551-1588.

[76] Sung Jaeyoung, 1995, “Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems,” *The RAND Journal of Economics* 26, 720-743.

[77] Wei, Chenyang and David Yermack, 2009, “Stockholder and bondholder reactions to revelations of large CEO inside debt holdings: an empirical analysis, ” Working Paper.

[78] Veblen, Thorstein, 1899, “The Theory of the Leisure Class: An Economic Study of Institutions,” New York: Penguin.