The London School of Economics and Political Science

Essays on Adaptation, Innovation Incentives and Compensation Structure

Yiqing LU

A thesis submitted to the Department of Finance of the London School of Economics for the degree of Doctor of Philosophy, London, May 2015

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 34,757 words.

I confirm that Chapter 3 is jointly co-authored with Peggy Huang and Moqi Xu.

Abstract

This thesis explores both theoretically and empirically how firms design employees' compensation contracts to motivate them to work and to adapt to external changes under an informed principal framework. The first chapter analyzes how a principal, privately informed about the changing market condition, structures the agent's incentive contract to inform and motivate her to adapt. The results show that a failure to overturn employees' belief about the changing market condition could lead to insufficient adaptation. Further, a more pressing market condition induces earlier adaptation and greater information revelation. Finally, the compensation structure underpinning insufficient adaptation imposes a legacy problem due to excessive use of long-term incentives, which restrains the reconfiguration of the contract in place. Based on the first chapter, the second chapter aims to explain asymmetric contractual adjustment of CEO compensation, only upward but not downward. I argue that a principal, privately informed about the firm's changing productive efficiency, uses contracts to provide the agent with not only working incentives but also information about her productivity. The principal commits to a back-loaded compensation plan with an increasing salary or with an increasing short-term performance pay. Such rigid contracts achieve greater efficiency by inducing more efforts from the agent through profit sharing. The third chapter, co-authored with Peggy Huang and Moqi Xu, finds CEO contracts explicitly account for subjective reviews in a new dataset of CEO contracts and stated reasons for compensation changes. Our results suggest that firms prefer to keep early R&D successes from the public and thus raise salaries for early R&D success not yet realized in performance measures. Consistent with this explanation, standalone salary increases predict better long-run portfolio and stock returns, but only following positive subjective evaluations and in firms with high R&D investment.

Acknowledgements

I would first like to thank my supervisors, Daniel Ferreira and Kathy Yuan, for all the advice and guidance that they have given me over the past years. In the many discussions that I have had with Daniel, he inspired me to think freely and evaluated my ideas critically. His guidance is present on every page, in particular, in the chapter on Adaptation, which would have been impossible without his encouragement. Kathy read my drafts and gave me numerous incisive comments to improve my papers and presentations. She encouraged me to be an adventurous researcher. I cannot thank her enough for her patience and support.

I am also immensely grateful to Ulf Axelson, Vicente Cuñat, Daniel Paravisini and Moqi Xu. They helped me bring my model closer to reality and repeatedly paraphrased my ideas in their own words, which greatly enriched my understanding of my own work. I am especially thankful to Moqi for guiding me on our empirical project and helping me improve my writing skills. I have also been lucky to have had numerous discussions with visiting and permanent faculty members at the London School of Economics. My heartfelt thanks go to Mike Burkart, Maria Bustamante, Jon Danielsson, Amil Dasgupta, Juanita Gonzalez-Uribe, Christian Julliard, Carsten Kengeter, Joel Shapiro and David Webb. Finally, I would like to thank Igor Makarov and Giacinta Cestone for agreeing to be my thesis examiners and providing detailed comments and suggestions.

Completing this thesis was a challenge, and I would therefore like to thank my colleagues and friends for helping me through the many stumbling blocks. Svetlana Bryzgalova, Huaizhi Chen, Jason Donaldson, Yi Fan, Sergei Glebkin, Shiyang Huang, John Kuong, Isabelle Roland, Ji Shen, Jing Zeng and Cheng Zhang made this process much more enjoyable than it would otherwise have been.

My biggest thanks go to my family. I am forever indebted to my parents for their unconditional love and faith in me, even though my trips back home were too few and far between. Chenhui, my beloved husband, has always been a great companion throughout my research journey. I thank him for being always proud of me. As I realize that this Acknowledgement is perhaps the only section in which I am allowed to reason with emotion, I want to thank myself for finally giving myself the chance of writing this Acknowledgement.

Contents

Chapter 1

Informing and Motivating Adaptation

1.1 Introduction

As firms operate in a changing environment, a given strategy is unlikely to remain optimal forever. This imposes tremendous challenges for firms to adapt. In a wide array of industries, including disk drives (Christensen, 1993; Lerner, 1997), drugs (Guedj and Scharfstein, 2004), photolithography (Henderson, 1993), photography (Gavetti, Henderson, and Giorgi, 2004) and steel (Collinson and Wilson, 2006), many of the world's leading firms have faltered over decades because they failed to respond to factors such as globalization of markets and rapidly evolving technology.

This paper studies the link between adaptation and incentive contracts and explores the difficulty of informing and motivating adaptation when senior managers are able to identify market changes. The analysis highlights the adverse role of information asymmetry between top managers and employees in obstructing successful adaptation. Top managers need to credibly communicate and instill their visions of market trends to employees in order to foster successful adaptation. There are scenarios in which communication becomes impossible or the information that concerns employees is soft in nature. When unaware of market conditions and the manager's vision, employees may find adaptation too costly and uncertain, leading to adaptation failure.

Conventional explanations for insufficient adaptation include sales cannibalization (Arrow, 1962), internal resistance (Dow and Perotti, 2010), coordination (Dessein and Santos, 2006; Rantakari, 2008), commitment (Rotemberg and Saloner, 1994) and iterative learning (Argote, 2012; Schreyögg and Sydow, 2011). A fundamental difference between other explanations and this paper lies in the premise of the origin of adaptation. While other papers consider adaptive changes initiated by specialized employees and divisional managers as in (Mintzberg and Waters, 1985), this paper takes the view of Bennis and Nanus (1985) and Quigley (1993) that visionary senior managers are the primary driving force of adaptation in organizations.[1,2]

Instilling a vision to employees and motivating them to adapt is extremely challenging. According to Kaplan and Henderson (2005), a lack of incentives for employees to cope with changes is one of the many reasons leading to adaptation failure. Nippon Steel, the world's largest steelmaker in 1970s, tried to counter the effects of the recession in 1990s by following a strategy of diversification in technological innovations. However, its researchers kept undertaking customer-induced innovation, as the career prospect of becoming professional specialists who undertake science-based innovation was too uncertain (Collinson and Wilson, 2006). Anderson Consulting and Kodak also experienced difficulties in overcoming their employees' belief in the efficacy of their new strategies.[3,4]

This paper also seeks to understand the extent to which a changing market shapes the adaptation dynamics of a firm, which is an issue that has not been the subject of formal economic analysis. I construct a new framework to study the evolution of firm adaptation and incentive systems under an informed principal setting. I show that insufficient adaptation, as the residue of successful implementation of an adaptive strategy in the past, besets the pursuit of a new strategy. Depending on the market condition, a firm's ability to adapt to market changes in the short run differs from its ability to adapt in the long run - early and late adaptation are both likely to happen. Moreover, adaptation is path-dependent in the sense that past success can either foster or suppress successful adaptation in future. Lastly, the optimal contract imposes a legacy problem, which restrains the reconfiguration of the incentive system in place.

The basic model presented in this paper consists of two periods and two players

[1]For example, CEOs could be hired for their vision in exploiting new potentials dormant in the market. Quigley (1993) conducted a survey on managers in 20 countries. 95% of them say that the most important CEO trait is the ability to convey a strong sense of vision to employees. Board directors have experiences and connections in multiple industries (Larcker, Saslow, and Tayan, 2014; Casal and Caspar, 2014), and the top management team could access confidential client and market data which is however not accessible to employees.

[2]Effective communication of the changing market condition to employees is without doubt a key component of successful adaptation (Covin and Kilmann, 1990; Lewis, 2006).

[3]In an attempt to generate further growth in its core IT business, Anderson Consulting experienced difficulty overcoming its employees' belief in the efficacy of the new strategy and consequently adopted an incentive system that were much closer in form to existing arrangements (Kaplan and Henderson, 2005).

[4]Kodak, which in 1976 had a 90% market share of photographic film sales in the United States, began to struggle financially in the late 1990s due to its slowness in transitioning from chemical to digital photography. As one industry executive commented, "Fisher (CEO of Kodak) has been able to change the culture at the very top. But he hasn't been able to change the huge mass of middle managers, and they just don't understand this [digital] world." (Gavetti et al., 2004)

- a principal (or the firm) and an agent (or the employee). The principal privately knows the initial market condition and how it changes in the second period. In each period, the agent implements one of two strategies - a market-insensitive and a market-sensitive strategy. The success probability of the market-insensitive strategy is fixed and known to all, but that of the market-sensitive strategy depends on the market condition. The market-sensitive strategy is more costly to implement, but has a potentially higher probability of success if it fits the market. When the market condition is good (resp., bad), the market-sensitive strategy is more (resp., less) efficient than the market-insensitive one. Market conditions change over time in a persistent way; a good market today increases the chance of a good market tomorrow.

Information asymmetry is key to understanding insufficient adaptation. The thrust of the mechanism can be explained by two opposite forces affecting an employee's decision to adapt. First, the principal incurs a cost of salary to reveal the market condition; second, not revealing the market condition via a contract increases the incentive cost of motivating the employee to adapt. A failure to overturn employees' belief about a changing market condition leads to a failure to adapt. If information asymmetry constrains the firm from building a fully informative incentive system, the employee can only infer the market condition from past performance.

First, I show that if the principal and the agent are equally informed of the changing market, the agent implements the strategy that adapts to the current market condition. In the presence of information asymmetry, insufficient and path-dependent adaptation arises in equilibrium. In one equilibrium which I call early adaptation, the principal reveals her private information in the first period and following only good performance in the second period. In the other equilibrium which I call late adaptation, the principal reveals her private information only in the second period and following only bad performance. A firm in a deteriorating market condition sticks to the market-sensitive strategy adopted in the first period following bad (resp., good) performance in early (resp., late) adaptation. It fails to adopt the market-insensitive strategy which has become the more cost-effective strategy in the new market. Both equilibrium adaptation paths thus exhibit inertial adoption of the old strategy.

In early adaptation, a firm in a good market maximizes the benefit of information revelation in the first period if there is no information revelation following bad performance in the second period. Otherwise it would be mistaken as in a bad market following bad performance if it does not signal in the first period and incur a high incentive cost. Since a firm in a good market is less likely than a firm in a bad market to attain bad performance, it saves the first-period signalling cost by committing to information revelation only following good performance. In late adaptation, a firm under a good market condition does not reveal its private information in the first period. An employee makes a negative inference of the market condition following bad performance and a positive inference following good performance. Due to the persistent market changes, a firm in a good market saves more incentive cost if it reveals the market condition following bad performance than following good performance.

Second, early adaptation achieves greater information revelation and efficiency than late adaptation. Early adaptation arises if the distribution of a good market condition is poor or if the new strategy requires a drastic change in the implementation cost. Intuitively, a pressing market outlook and a drastic shift in market conditions force a firm to build an informative incentive system to reduce the incentive cost of adaptation. This is also in reminiscent of Schumpeterian view that economic downturns play a positive role in promoting long-run growth.

Third, my model suggests that a firm's incentive system is intimately interlinked with its adaptation path. The compensation structure that induces full adaptation consists of a non-decreasing fixed pay and a non-decreasing short-term performance-based pay. Equilibrium contracts under both early and late adaptation, however, are path-dependent and thus include long-term performance-based pay. In particular, a downward rigid contract is too costly to adopt in late adaptation. Moreover, the long-term commitment limits a firm in a non-deteriorating market to restructure its incentive system. Its employees would not give up the overly-paid incentive compensation if they were informed of the good market condition. Consequently, a firm which faces a worsening market keeps mimicking a firm in a good market in order to save incentive cost. Internal resistance, therefore, endogenously arises in the firm's design of an optimal incentive system.

Lastly, the paper sheds light on policies that aim to facilitate and direct adaptive changes. My model suggests that selective policies can alleviate the information asymmetry problem through reducing the signalling cost. A government can either subsidize good performers or tax bad performers based on their after-compensation earnings. Such policies create differential effects on firms in different markets and reduce the cost of information revelation, giving rise to in full adaptation.

Related literature. The paper fits with an emerging body of theoretical and empirical literature that attempts to explain insufficient adaptation. Early examples are Hannan and Freeman (1984), Holmström (1989), and Rotemberg and Saloner (1994); more recent works include Manso (2011), Dessein and Santos (2006), Dow and Perotti (2010), Ferreira and Rezende (2007), Bolton, Brunnermeier, and Veldkamp (2013). These articles focus on related but different questions, such as compensation contracts, organization, human capital investment and managerial over-confidence.

This model is closely related to four different veins of theoretical literature:

(1) Incentive contracts and project choices. In the focus on incentive contracts and project choices in innovation, the paper is closer to Holmström (1989) and Manso (2011). Holmström (1989) explains why small firms are responsible for a disproportionate share of innovative research. It is shown that incentive costs associated with a given task depend on the total portfolio of tasks that a firm undertakes. Mixing hard-to-measure activities (innovation) with easy-to-measure activities (routine) is more costly for larger firms with a heterogeneous set of tasks. Manso (2011) characterizes the optimal contract that motivates its employee to conduct either exploration-based or exploitation-based innovation. In his setting, because neither the employee nor the firm knows the exact success probability of the exploration-based innovation, the firm needs to incentivise the employee to experiment with the exploratory project in order to learn its quality. Such a contract features the use of long-term incentives and high tolerance for short-term failure. Ederer and Manso (2013), in a controlled laboratory setting, show evidence that the combination of tolerance for early failure and reward for long-term success is effective in motivating innovation.

The construction of project choices in my paper is adapted from this body of literature but differs from it because of the focus on the roles of information asymmetry regarding market condition. In particular, the principal knows the underlying market condition that affects the success probability of the market-sensitive strategy. This model is concerned with the difficulty of information transmission from the top to the bottom in obstructing adaptation rather than learning the quality of the exploratory project.

(2) Organizational structure and adaptation. Dessein and Santos (2006) and Rantakari (2008) emphasize the importance of coordination and authority in influencing adaptation. In their setting, divisional managers (the agent) instead of the headquarter (the principal) have direct access to information about local market conditions. While their set-ups are motivated by the bottom-up view that small initiatives that spread throughout the organization are undertaken at the discretion of employees (Mintzberg and Waters, 1985), mine takes the top-down view as in Bennis and Nanus (1985) and Quigley (1993).

Relatedly, Dow and Perotti (2010) posits that firms fail to adapt to the changed circumstance because losers from the radical adjustment can credibly resist and oppose changes, lending support to the creation of new firms with no internal resistance. However, both the firm and agent understand the changes in their setting. Critically, their paper assumes that output is not verifiable and therefore remains silent on the issues of incentive contracts.

(3) Project-specific human capital investment. Another related literature on commitment and strategy specific investment also explains the difficulty of adopting adaptive strategies. Rotemberg and Saloner (1994) and Van den Steen (2005) show that managerial vision, or a bias towards a specific strategy, provides employees with more certainty that their strategy-specific investments will pay off. Mailath, Nocke, and Postlewaite (2004) consider mergers and the internalization of the negative externality. Ferreira and Rezende (2007) show that reputation concerned managers could use public disclosure of strategic plans as commitment to a specific strategy. Bolton et al. (2013) argue that a CEO's overconfidence regarding the quality of his initial information on the firm's optimal strategy serves as commitment. Those models posit that firms optimally stick to a narrower set of strategies to resolve the time-inconsistent problem and to encourage employees to invest in task specific skills.

My paper is different in that it is the informational friction that makes a firm stick to the old strategy that no longer fits the changing market. Those models also remain silent on the role of incentive contracts in influencing adaptation.

(4) Informed principal models. The model is intellectually indebted to the literature of managerial compensation with an informed principal (Fuchs, 2013; Zábojník, 2014). Fuchs (2013) studies base salary as a signalling device but leaves the discussion of incentive pay aside. Zábojník (2014) explores the incentive effect of the subjective pay but restricts the attention to only separating equilibrium. In addition, both papers consider an informed principal with a constant type. In contrast to their frameworks, I analyze a changing information structure, which allows me to explore adaptation paths. Moreover, I examine the pooling equilibrium and explore the trade-off between informing and motivating adaptation. This brings new implications on compensation structure that would be otherwise impossible to attain under separating equilibrium.

In addition to the above four strands of literature, my work also complements the literature of evolutionary economics in understanding organizational inertia. Nelson and Winter (1982) point out that organizations tend to develop procedures and routines that once established are hard to change fundamentally. Hannan and Freeman (1984) argue that adaptation of organizational structures to environments occurs principally at the population level, with forms of organization replacing each other as conditions change. A major premise of this theory is that individual organizations are subject to strong inertial forces, but the logic of the very process producing organizational persistence remains by and large under-explored. Schreyögg and Sydow (2011) provide a review on the recent development in understanding the aforementioned process. For example, learning effects hold that the more often an operation is performed, the more efficiency will be achieved when operating subsequent iterations (Argote, 2012). My paper points out that information asymmetry could lead to organizational inertia due to failures in informing employees to take adaptive changes.

Last but not least, this paper is related to the macroeconomics literature on technology adoption and its speed. Acemoglu (2002) argues that the elasticity of substitution between different factors determine how technical change and factor prices respond to changes in relative supplies. Tinn (2010) presents a general equilibrium model which shows the importance of equity markets in facilitating the exit of entrepreneurs investing in technology. Firms trade off the"fear of unstable markets" with the "adoption to signal" force. Credit constraints may also be obstacles to fast technology adoption (Gertler and Rogoff, 1990; Aghion, Bacchetta, and Banerjee, 2004; Aghion, Comin, and Howitt, 2006). Different from this literature, my paper takes a micro approach to analyze the slowness in adaptation and posits that it can also arise as a result of a within-firm informational friction.

The rest of the paper is organized as follows. Section 2 describes the model. Two benchmark cases are discussed in section 3. Second 4 analyzes the equilibrium adaptation paths. Section 5 presents compensation structures underpinning different adaptation paths. Section 6 considers policies which facilitate adaptation. The last section concludes.

1.2 Model Setup

The model consists of two periods, period one and two ($t = 1$ and 2) and three dates (date 0, 1, and 2). There are two players, a principal and an agent.

1.2.1 Dynamic Environments

The market condition m_t, in each period, can be in one of two possible states, $m_t \in \{U, D\}$. U (resp., D) represents a good (resp., bad) market condition.

At date 0, the prior probabilities of m being U and D are α and $1-\alpha$ respectively, $0 < \alpha < 1$. The market condition might change in the second period. With β probability, a good market condition continues to be good, $Pr(m_2 = U|m_1 = U) = \beta$, and $0 < \beta < 1$. With $1 - \beta$ probability, a good market condition deteriorates, $Pr(m_2 = D|m_1 = U) = 1 - \beta$. Parameters α and β are known to both parties.

A bad market condition in the first period remains bad in the second period, $Pr(m_2 = D|m_1 = D) = 1$. Even though most of the results we obtain do not rely on this simplification, it greatly simplifies the analysis presented in this paper.[5]

[5] As long as the market change is persistent, the result of path-dependency in adaptation holds in a more general setting in which a bad market could improve in the second period. This assumption

To summarize, the market changes persistently - a good market today predicts a higher likelihood of a good market tomorrow than a bad market does.

1.2.2 Two Strategies

The modeling of two strategies is adapted from Manso (2011). Manso considers a classical two-period and two-armed bandit problem, in which the agent takes either an exploitative or an exploratory action in each period at a private cost. Both parties know the success probability of the exploitative action but not that of the exploratory one.

In my model, the agent can also choose one of two strategies $s_t \in \{s_s, s_i\}$ in each period. s_s is a market-sensitive strategy, and s_i a market-insensitive strategy. Implementing either of them yields a verifiable output y_t at the end of each period, $y_t \in \{0, 1\}$. y_1 and y_2 are independently distributed. The market-insensitive strategy generates a high output with probability θ_l, which remains constant and known to both parties over time .

The market-sensitive strategy exhibits more uncertainty. Its probability of success θ_t in period t depends on the market condition in period t. For instance, if a newly developed product meets consumer tastes, or a firm expands into a foreign market with a potential of rapid economic growth, then θ_t is high. Otherwise, it is low. To be more specific, θ_t can take two possible values $\theta_t \in \{\theta_l, \theta_h\}$, with $0 < \theta_l < \theta_h \leq 1$.[6] If $m_t = U$, then $\theta_t = \theta_h$, or $Pr(y_t = 1|m_t = U) = \theta_h$. If $m_t = D$, then $\theta_t = \theta_l$, or $Pr(y_t = 1|m_t = D) = \theta_l$. Because the market condition may change over time, the market-sensitive strategy which has a success probability of θ_h in the first period may become less productive in the second period.

If the principal and the agent are both uninformed of the underlying market condition, my model closely resembles the problem solved in Manso (2011) in the sense that the principal and the agent can only learn the market condition by experimenting with the market-sensitive strategy. Because the market does not affect the output generated by the market-insensitive strategy, adopting the market-insensitive strategy does not produce any informational value to the firm.

also provides a robust setting to study the excessive adoption of market-sensitive strategy, which is least likely to happen under this setting.

[6]Similar results hold if θ is smaller than θ_l in the bad state. However, the analysis is more complicated as the set of equilibrium strategies taken by the agent increases under pooling equilibrium. It also does not add new insights.

If s_s is implemented, the agent can make an inference from y_1 about m_1.

$$Pr(m_1 = U|y_1 = 1) = \frac{Pr(y_1 = 1|m_1 = U)Pr(m_1 = U)}{Pr(y_1 = 1)} = \frac{\alpha\theta_h}{\alpha\theta_h + (1-\alpha)\theta_l}$$

The probability of the a good market in the second period upon observing a high output is therefore:

$$Pr(m_2 = U|y_1 = 1) = \beta Pr(m_1 = U|y_1 = 1) = \frac{\alpha\beta\theta_h}{\alpha\theta_h + (1-\alpha)\theta_l}$$

And the probability of achieving a high output in the second period given a high output in the first period is:

$$\begin{aligned} Pr(y_2 = 1|y_1 = 1) &= \theta_h Pr(m_2 = U|y_1 = 1) + \theta_l(1 - Pr(m_2 = U|y_1 = 1)) \\ &= \frac{\alpha\beta\theta_h(\theta_h - \theta_l)}{\alpha\theta_h + (1-\alpha)\theta_l} + \theta_l \end{aligned}$$

Likewise, the probability of achieving a high output in the second period given a low output in the first period is:

$$\begin{aligned} Pr(y_2 = 1|y_1 = 0) &= \theta_h Pr(m_2 = U|y_1 = 0) + \theta_l(1 - Pr(m_2 = U|y_1 = 0)) \\ &= \frac{\alpha\beta(1-\theta_h)(\theta_h - \theta_l)}{\alpha(1-\theta_h) + (1-\alpha)(1-\theta_l)} + \theta_l \end{aligned}$$

The unconditional probability of achieving a high output in the second period is:

$$Pr(y_2 = 1) = \theta_h Pr(m_2 = U) + \theta_l Pr(m_2 = D) = \alpha\beta(\theta_h - \theta_l) + \theta_l$$

If the principal and the agent are symmetrically uninformed of the market condition, the transition matrix of successes corresponds to those assumed in Manso (2011) and Ferreira, Manso, and Silva (2014). To see this, one could easily verify that $Pr(y_2 = 1|y_1 = 0) < Pr(y_2 = 1) < Pr(y_2 = 1|y_1 = 1)$.[7] Intuitively, if the agent takes the market-sensitive strategy, a high output indicates a greater likelihood of a good market condition both today and tomorrow, therefore, a higher probability of a high output tomorrow. A low output indicates a higher likelihood of a bad market condition both today and tomorrow, therefore, a higher probability of a low output tomorrow. If the market change is not persistent ($\beta = 0$), then $Pr(y_2 = 1|y_1 = 0) = Pr(y_2 = 1) = Pr(y_2 = 1|y_1 = 1) = \theta_l$. In other words, learning from the past performance

[7] Manso (2011) and Ferreira et al. (2014) assume that $E(q|F) < E(q) < E(q|S)$. q is the probability of success of the innovative task, unknown to both the principal and the agent. F and S mean success and failure of the task implemented by the agent.

adds additional information to the firm's information set only if the market condition changes persistently. Otherwise, past performance is not indicative of future market conditions.

1.2.3 Incentive Problems

The principal hires the agent to implement a strategy. The market-insensitive (resp., market-sensitive) strategy requires a private effort cost of C_l (resp., C_h) to implement. The probability of achieving a high output, if the agent does not implement any strategy, is zero. As in Manso (2011), I assume that the principal does not observe the strategies implemented by the agent.[8] This assumption is meant to capture the difficulty for large organizations in ensuring implementation of the desired strategies. As pointed out by Mintzberg (1988), when implementing strategies, employees are sitting between past experiences (or knowledge) and future prospects. In other words, implementing strategies involves the work of minds, which is inherently unobservable. Adding to the difficulty is the separation of formulation of strategies done by senior managers and the implementation done by the many below. A lack of input-based measures, especially in uncertain environments, renders the monitoring of the agent's actions impossible.

An efficiency condition holds for the two strategies: $\frac{C_h}{\theta_h} < \frac{C_l}{\theta_l}$. While Manso (2011) considers both $C_l \geq C_h$ and $C_l < C_h$, I restrict attention to the case in which $C_l < C_h$ and thus rule out situations in which the market-sensitive strategy dominates the market-insensitive one in both strict and weak form.[9] These two conditions imply that the market-insensitive strategy is more efficient than the market-sensitive strategy under a bad market condition, but less under a good market condition. In other words, the market-sensitive strategy adapts to a good market condition while the market-insensitive strategy adapts to a bad market condition.[10]

1.2.4 An Informed Principal

The set-up of this paper departs from Manso (2011) in the following way. The principal is privately *informed* of the exact value of m_t in both periods. Therefore learning through experimentation has no informational value to the principal. The agent, how-

[8]This assumption is important because it implies that the principal will rely on output-based pay to provide incentives to the agent.

[9]If $C_l = C_h$, the market-sensitive strategy weakly dominates the market-insensitive strategy. In this case, equilibrium results still hold, but there is no efficiency loss in equilibrium.

[10]Note that the effort cost and success probabilities of these two strategies do not imply that the principal always prefers the market-sensitive strategy, as the principal has to internalize the incentive cost of motivating the agent to work.

ever, is uninformed of the market condition throughout. Although she could infer it from the past performance by implementing the market-sensitive strategy, the inference will not be as accurate as the private information that the principal has.

If neither the principal nor the agent is informed of the market condition, then it is more efficient to continue implementing the market-sensitive strategy following good performance and to switch back to the market-insensitive strategy following bad performance, a problem which has already been analyzed in Manso (2011) if I assume the same parameter values. However, in the setup of this paper, continuing implementing the market-sensitive strategy following good performance may not be efficient, as the market condition may deteriorate in the second period.

The economic setting I analyze in this paper thus emphasizes situations in which a visionary manager needs to credibly convey her private information to the employee and lead the ill-informed employee to adapt to market changes. Because the principal accurately knows the market condition, the strategy choice made by the agent in the second period depends not only on what she infers from past performance but also on information revealed by the principal. The focus of the contracting problem in this paper is not to encourage experimentation and learning but to facilitate information transmission from the top to the bottom in a hierarchical organization. The ultimate goal of such information transmission is to encourage adaptation.

1.2.5 Preferences

The principal and the agent are both risk neutral. For simplicity, I assume that the discount rate for future payoffs is zero. The principal maximizes the firm's profit, and the agent maximizes her compensation after deducting her effort disutility. She has zero initial wealth and is protected by limited liability. Her reservation utility is assumed to be zero over the entire time horizon, under any market conditions.

1.2.6 Contracts

Here I characterize long-term contracts in equilibrium. The principal offers the agent a contract $\mathcal{M}$ at date 0. The contract is a subset of $\mathbb{R}^4_+$, $\mathcal{M} \subseteq \mathbb{R}^4_+$. I call an element of $\mathcal{M}$ a compensation plan, i.e., contract $\mathcal{M}$ is a set of compensation plans. Let $w_{y_1y_2} = \{w_{00}, w_{10}, w_{01}, w_{11}\}$ denote a generic element of $\mathcal{M}$ (i.e., a compensation plan). The principal pays the agent w_{00} if $(y_1, y_2) = (0, 0)$, w_{10} if $(y_1, y_2) = (1, 0)$, w_{01} if $(y_1, y_2) = (0, 1)$, and w_{11} if $(y_1, y_2) = (1, 1)$. Limited liability constraints imply that all payments are non-negative.

While performance is contractible, neither the principal's private information nor

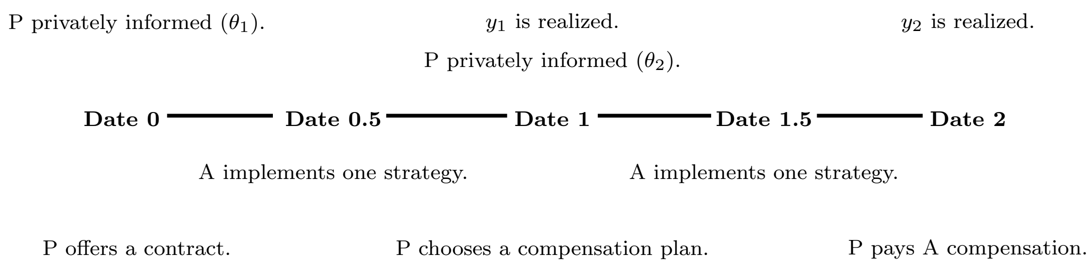


Figure 1.1: The Time-line

Note: A represents the agent; P represents the principal.

the agent's strategy choice is. Based on the private signal θ_1, the principal offers the agent a contract $\mathcal{M}$ at date 0. $\mathcal{M}$ may contain more than one compensation plan. After receiving the private signal θ_2 in the second period, the principal, at her sole discretion, chooses a single compensation plan $w_{y_1y_2}$ from $\mathcal{M}$. If the principal agrees to compensate the agent according to $w_{y_1y_2}$, then the agent is paid according to the compensation plan.

Figure 1.1 presents the time-line. At date 0, the principal is privately informed of the market condition m_1 and offers a contract $\mathcal{M}$ to the agent. The agent could leave or stay. If she leaves, she gets a reservation utility of zero. If she accepts the contract, she implements one strategy. At the end of the first period, two parties observe the realization of y_1.

At the beginning of the second period, after observing the market condition m_2, the principal chooses a single element $w_{y_1y_2}$ from the contract $\mathcal{M}$ and offers it to the agent. The agent then implements one strategy again. At the end of the second period, two parties observe the realization of y_2. Compensation is finally paid.

Because the setting involves a signalling problem, the payment scheme will be either fully revealing under a separating Perfect Bayesian Equilibrium (PBE) or not under a pooling PBE. This setting therefore might have multiple equilibria. I use a belief-based refinement approach of Undefeated Equilibrium, introduced by Mailath, Okuno-Fujiwara, and Postlewaite (1993).[11] I will introduce this approach in Section 1.4 before the analysis of the equilibrium structure of information revelation.

[11] I also apply Cho and Kreps (1987)'s Intuitive Criterion, see Appendix 2.

1.3 Two Benchmark Cases

This section presents two benchmark cases. In the first benchmark case, the agent also knows the market condition. In the second benchmark case, I solve for a socially optimal contract under information asymmetry, assuming that the principal is constrained to offer a fully revealing contract.

1.3.1 Benchmark One – Symmetric Information

The time-line of the first benchmark case corresponds to the two period model in Section 1.2, but both parties know the true market conditions throughout. This benchmark case describes contracts offered by the principal and strategies taken by the agent when both parties have full knowledge of market changes.

The following proposition presents the equilibrium contract under this benchmark case. Superscripts UD, UU and DD represents the three types of market changes over the two periods. They indicate the compensation plan that is chosen by the principal in the respective market. Although both parties know the market condition, the principal's choice of which compensation plan to offer cannot be contracted upon due to the non-verifiability of the market condition.

Proposition 1 *Assume that both parties know the market condition.*

- *If $m_1 = U$, then the principal offers a contract $\{(w_{00}^{UU}, w_{01}^{UU}, w_{10}^{UU}, w_{11}^{UU}), (w_{00}^{UD}, w_{01}^{UD}, w_{10}^{UD}, w_{11}^{UD})\}$ which satisfies:*
  - *If $m_2 = U$, then the principal chooses the compensation plan $w_{01}^{UU} = \frac{C_h}{\theta_h}$, $w_{10}^{UU} = \frac{C_h}{\theta_h}$, $w_{11}^{UU} = \frac{C_h}{\theta_h} + w_{10}^{UU}$ and $w_{00}^{UU} = 0$. The agent always implements s_s.*
  - *If $m_2 = D$, then the principal chooses the compensation plan $w_{01}^{UD} = \frac{C_l}{\theta_l}$, $w_{10}^{UD} = \frac{C_h}{\theta_h}$, $w_{11}^{UD} = \frac{C_l}{\theta_l} + w_{10}^{UD}$ and $w_{00}^{UD} = 0$. The agent implements s_s in the first period and s_i in the second period.*
- *If $m_1 = D$, the principal offers a contract $w_{01}^{DD} = \frac{C_l}{\theta_l}$, $w_{10}^{DD} = \frac{C_l}{\theta_l}$, $w_{11}^{DD} = \frac{C_l}{\theta_l} + w_{10}^{DD}$ and $w_{00}^{DD} = 0$. The agent always implements s_i.*

The contract under symmetric information exhibits two interesting features. First, the agent is incentivised to implement the strategy that adapts to the current market condition. When the agent also knows the market condition and how it changes over time, the principal need not design a contract to inform or convince the agent. The only role of the contract is to motivate the agent to take the adaptive strategy. In the absence of information asymmetry, the firm achieves full adaptation.

Second, the contract in Proposition 1 requires no commitment, because it is sequentially efficient to the principal. In fact, as shown in Corollary 1, one could decompose the contract into two short-term contracts which consist of short-term incentive pay (bonus) based on the performance only in the current period. The principal in a good market in the first period offers a bonus $\frac{C_h}{\theta_h}$ following good performance and the same level if the market condition does not deteriorate, and offers $\frac{C_l}{\theta_l}$ if it deteriorates.

Corollary 1 *The contract in Proposition 1 can be implemented by two short-term contracts:*

- *If $m_1 = U$, then the principal, in the first period, offers $w_1^U = \frac{C_h}{\theta_h}$ and $w_0^U = 0$. In the second period,*
  - *If $m_2 = U$, then the principal continues to offer $w_1^U = \frac{C_h}{\theta_h}$ and $w_0^U = 0$;*
  - *If $m_2 = D$, then the principal offers $w_1^D = \frac{C_l}{\theta_l}$ and $w_0^D = 0$;*
- *If $m_1 = D$, then the principal offers $w_1^D = \frac{C_l}{\theta_l}$ and $w_0^D = 0$ in both periods.*

1.3.2 Benchmark Two – A Fully Revealing Contract

The second benchmark case characterizes a contract that fully reveals the principal's private information at any time t and following any performance levels. To be more precise, I impose truth-telling constraints onto the principal's maximization program in both periods and following both performance levels.

Lemma 1 *Conditional on the agent learning the market condition, the agent implements the strategy that adapts to the current market condition.*

Lemma 1 is very useful in simplifying the maximization program of this benchmark case. If the agent learns a good market condition, then she implements the market-sensitive strategy. If the agent learns a bad market condition, she implements the market-insensitive strategy. I could thus remove the agent's non-binding project choice constraints from the maximization program.

The intuition of Lemma 1 is simple. Conditional on the good market condition being revealed, if the principal offers the agent a compensation plan that just satisfies the agent's incentive constraint under the market-insensitive strategy, the agent will always prefer the market-sensitive strategy because of its high productivity. The principal therefore offers the agent an incentive pay that is just enough to compensate the agent's effort of taking the market-sensitive strategy. Similarly, conditional on the bad market condition being revealed, if the principal offers the agent a compensation plan that satisfies the agent's incentive constraint under the market-sensitive strategy, the agent will always prefer the market-insensitive strategy because of its low cost. The principal therefore offers the agent an incentive pay that is just enough to compensate the agent's effort of taking the market-insensitive strategy.

Below is the maximization program for a principal who privately knows a good market condition at date 0 and is forced to reveal information throughout.

$$\max_{\mathcal{M}} \quad \theta_h\{\beta(\theta_h(2-w_{11}^{UU})+(1-\theta_h)(1-w_{10}^{UU}))+(1-\beta)(\theta_l(2-w_{11}^{UD})+(1-\theta_l)(1-w_{10}^{UD}))\}$$
$$+(1-\theta_h)\{\beta(\theta_h(1-w_{01}^{UU})+(1-\theta_h)(0-w_{00}^{UU}))+(1-\beta)(\theta_l(1-w_{01}^{UD})+(1-\theta_l)(0-w_{00}^{UD}))\}$$

The maximization program is subject to several constraints. To save space, I do not list the limited liability constraints.[12]

$$s.t. \quad \theta_h w_{11}^{UU}+(1-\theta_h)w_{10}^{UU}-C_h \geq w_{10}^{UU} \tag{1.3.1}$$
$$\theta_h w_{01}^{UU}+(1-\theta_h)w_{00}^{UU}-C_h \geq w_{00}^{UU} \tag{1.3.2}$$
$$\theta_l w_{11}^{UD}+(1-\theta_l)w_{10}^{UD}-C_l \geq w_{10}^{UD} \tag{1.3.3}$$
$$\theta_l w_{01}^{UD}+(1-\theta_l)w_{00}^{UD}-C_l \geq w_{00}^{UD} \tag{1.3.4}$$
$$\theta_h\{\beta(\theta_h(w_{11}^{UU}-w_{10}^{UU})+w_{10}^{UU}-C_h)+(1-\beta)(\theta_l(w_{11}^{UD}-w_{10}^{UD})+w_{10}^{UD}-C_l)\} \tag{1.3.5}$$
$$+(1-\theta_h)\{\beta(\theta_h(w_{01}^{UU}-w_{00}^{UU})+w_{00}^{UU}-C_h)+(1-\beta)(\theta_l(w_{01}^{UD}-w_{00}^{UD})+w_{00}^{UD}-C_l)\}$$
$$-C_h \geq \beta(\theta_h(w_{01}^{UU}-w_{00}^{UU})+w_{00}^{UU}-C_h)+(1-\beta)(\theta_l(w_{01}^{UD}-w_{00}^{UD})+w_{00}^{UD}-C_l)$$
$$\theta_l(2-w_{11}^{UD})+(1-\theta_l)(1-w_{10}^{UD}) \geq \theta_l(2-w_{11}^{UU})+(1-\theta_l)(1-w_{10}^{UU}) \tag{1.3.6}$$
$$\theta_l(1-w_{01}^{UD})+(1-\theta_l)(0-w_{00}^{UD}) \geq \theta_l(1-w_{01}^{UU})+(1-\theta_l)(0-w_{00}^{UU}) \tag{1.3.7}$$
$$\theta_l^2(2-w_{11}^{DD})+\theta_l(1-\theta_l)(1-w_{10}^{DD})+\theta_l(1-\theta_l)(1-w_{01}^{DD})+(1-\theta_l)^2(0-w_{00}^{DD}) \tag{1.3.8}$$
$$\geq \theta_l^2(2-w_{11}^{UD})+\theta_l(1-\theta_l)(1-w_{10}^{UD})+\theta_l(1-\theta_l)(1-w_{01}^{UD})+(1-\theta_l)^2(0-w_{00}^{UD})$$

Constraints 3.1-3.5 are the agent's incentive constraints. One could easily verify that, due to zero reservation utility, the agent's participation constraints will be automatically satisfied if her incentive constraints are satisfied.[13] Constraints 3.6-3.8 are the principal's truth-telling constraints.

Proposition 2 ***A fully revealing contract.***

[12]All contingent payment must be greater than or equal to zero.

[13]In fact, under information symmetry, because the probability of success is assumed to be zero if the agent shirks and w_{00}^{UU} is equal to zero, the incentive constraint 3.2 is also the agent's participation constraint following good performance in a non-deteriorating market.

If $m_1 = U$, the principal commits to such a contract that restricts her to only two compensation plans to choose from in the second period:

$$
\begin{aligned}
w^{UD}_{00} =&\Delta; \quad & w^{UD}_{01} =&\frac{C_l}{\theta_l} + \Delta; \quad & w^{UD}_{10} =&\frac{C_h}{\theta_h} + \Delta; \quad & w^{UD}_{11} =&\frac{C_l}{\theta_l} + \frac{C_h}{\theta_h} + \Delta; \textit{or} \\
w^{UU}_{00} =&2\Delta; \quad & w^{UU}_{01} =&\frac{C_l}{\theta_l} + 2\Delta; \quad & w^{UU}_{10} =&\frac{C_h}{\theta_h} + 2\Delta; \quad & w^{UU}_{11} =&2\frac{C_h}{\theta_h} + 2\Delta.
\end{aligned}
$$

If $m_1 = D$, the principal offers the same contract as described in Proposition 1.

Proposition 3 indicates three interesting features of a fully separating contract. First, as in the benchmark case, when the agent is informed of the market changes, the firm achieves full adaptation. A fully revealing contract thus incentivises the agent to implement the strategy that adapts to the current market condition.

Second, neither of the two compensation plans contain long-term incentive pay that depends on both y_1 and y_2. The two plans can be decomposed into a fixed component (salary) and short-term performance-based pay (bonus). The equilibrium contracts, as shown in the next section, do not possess this feature, because the equilibrium structure of information revelation is path-dependent.

Third, although the two compensation plans can be replicated by short-term contracts, the contract offered in the first period cannot. Essentially, the principal commits to a long-term contract which promises increasing compensation in the second period.

Corollary 2 *The contract in Proposition 3 can be further characterized by a downward rigid contract. The principal commits to paying*

- *in $t = 1$, $w^U_1 = \frac{C_h}{\theta_h} + \frac{1}{2}\Delta$ and $w^U_0 = \frac{1}{2}\Delta$;*
- *in $t = 2$, either $w^D_1 = \frac{C_l}{\theta_l} + \frac{C_h}{\theta_h} + \frac{1}{2}\Delta$ and $w^D_0 = \frac{1}{2}\Delta$, or $w^U_1 = \frac{C_h}{\theta_h} + \frac{3}{2}\Delta$ and $w^U_0 = \frac{3}{2}\Delta$.*

Corollary 2 illustrates this feature. w^U_0 is the salary and $w^U_1 - w^U_0$ is the bonus for good performance in the first period. Salary and bonus in the second period can be decomposed in the same way. Figure 1.2 presents a graphical illustration of the downward rigid contract in Corollary 2. Red color indicates a constantly good market and blue a deteriorating market. Dashed lines represents salary and solid lines represent bonus following good performance. As shown in Figure 1.2, salary and bonus both exhibit downward rigidity which only allows upward but not downward adjustment. The principal raises either the salary in the second period if the market remains good or the bonus if the market condition deteriorates.

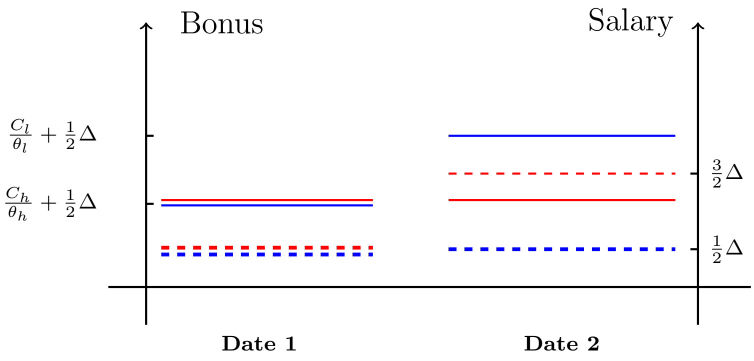


Figure 1.2: A fully separating contract

Note: Dashed – salary; Line–bonus; Red: UU; Blue: UD.

1.4 Equilibrium Structure of Information Revelation

This section analyzes the equilibrium structure of information revelation. Given that a number of equilibria can be supported by a variety of off-the-equilibrium path beliefs, I use a belief-based refinement approach of Undefeated Equilibrium introduced by Mailath et al. (1993). Intuitive Criterion and D1 (Cho and Kreps, 1987) eliminates all pooling equilibria, among which some are interesting and reasonable. This is due to its lack of a "global" consistency which neglects all the subsequent adjustments in strategies and beliefs that will take place after a disequilibrium message is sent.[14] In fact, in a monotonic signaling game or if the sorting condition is satisfied, any equilibrium where two or more types assign positive probability to the same action must fail Intuitive Criterion and D1 (Cho and Sobel, 1990).[15] A stronger concept of universal divinity proposed by Banks and Sobel (1987) selects the separating equilibrium with more than two types but coincides with D1 when there are only two types. None of them therefore have bite in this model.

The Undefeated Equilibrium approach selects among different pure strategy PBEs and selects a unique equilibrium outcome for a given set of parameters. In my setting, these equilibria are such that

1. Principal in each type of markets uses a pure strategy and maximizes profits given the agent's choices and the other principal's pure strategy;

2. The agent chooses either the market-insensitive or the market-sensitive strategy

[14] Cho and Kreps attribute this reasoning to Stiglitz.

[15] Cho and Kreps also observe that D1 picks out the separating equilibrium with three types.

conditional on the contract offered by the principal;

3. Beliefs are calculated using Bayes's rule for the contract offered by the principal used with positive probability.

Undefeated Equilibrium[16] is defined as follows. A PBE, G, defeats another PBE, G', if:

1. There is a message m sent only in G by a set of types K;

2. The set of types K who send m are all better off in G than in G', and at least one of them is strictly so;

3. Off-the-equilibrium beliefs under G' about at least one type in K conditional on sending m are not a posterior probability assuming: (i) only types in K send m with positive probability and (ii) those types in K that are strictly better off under G send m with probability one.[17]

A PBE G is said to be undefeated if there does not exist another PBE G' that defeats it. The undefeated approach is essentially a lexicographically maximum refinement concept and works by checking that no types in one equilibrium are better off in another equilibrium where they choose a different action/message.

Figure 1.3 and Figure 1.4 show all the possible structures of information revelation. Specifically, Figure 1.3 includes those that reveal the market condition in the first period, and Figure 1.4 includes those that do not. Information revelation in the second period may depend on the performance realization at date 1. The analysis of equilibrium contracts involves more complications than the second benchmark case for two reasons.

First, I impose truth-telling constraints in the analysis of a fully revealing contract, the structure of which is presented in Figure 1.3a. In equilibrium, the principal, however, may not be willing to offer a fully revealing contract. Which truth-telling constraint binds has to be endogenously determined by the principal. The principal may want to offer a partial revealing contract to save the cost of information revelation.

Second, one major difference between structures in Figure 1.3 and Figure 1.4 lies in how the agent learns the market condition. If the principal does not reveal information in the first period, the agent has to learn it from past performance. Upon observing

[16]This refinement approach is also used in several other papers, including Taylor (1999), Gomes (2000), Fishman and Hagerty (2003) and Josephson and Shapiro (2014).

[17]The third condition is imposed on the off-the-equilibrium belief of type K and ensures that there is a message m that is sent only in G.

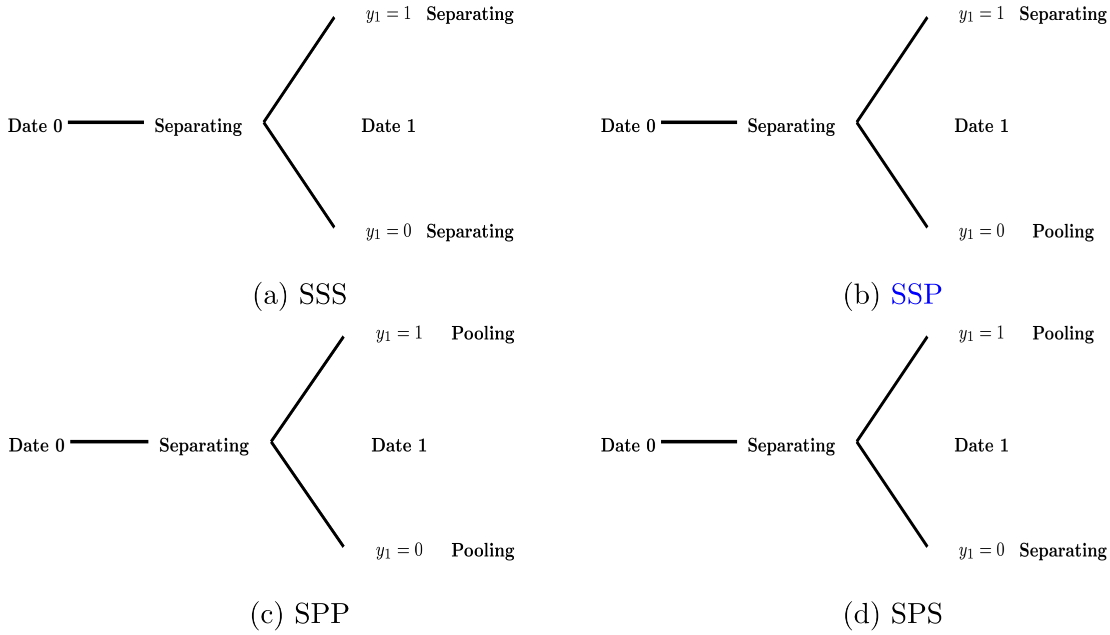


Figure 1.3: Possible Structures of Information Revelation I

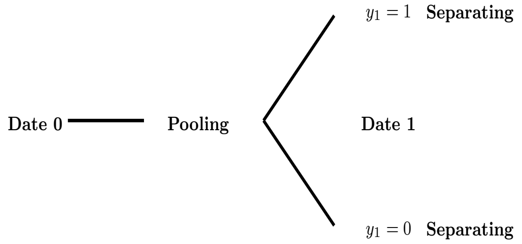
Note: S represents separating; P represents pooling.

good first period performance, the agent believes that the market is more likely to be good today as well as tomorrow due to the persistence in market changes. Conversely, upon observing bad first period performance, the agent believes in a bad market condition today and tomorrow. The principal therefore has to consider the inference made by the agent when designing the contract at date 0.

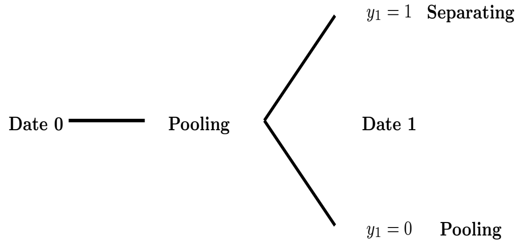
Before I characterize the equilibrium structures of information revelation, I first describe the strategies implemented by the agent in equilibrium in Lemma 2.

Lemma 2 *In any equilibrium in which the agent does not learn the current market condition, she implements the market-sensitive strategy. In any equilibrium in which the agent learns the market condition, she implements the strategy that adapts to the current market condition.*

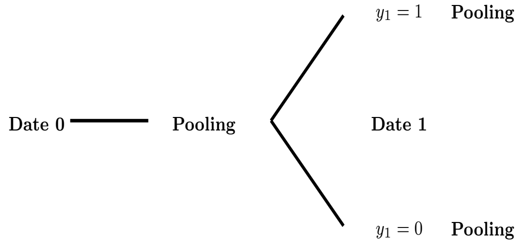
The first result of Lemma 2 can be intuitively explained as follows. Assume that the equilibrium contract is designed in a way that agents in both markets implement s_i under no information revelation. The principal in a good market is better off offering a contract that reveals the market condition to the agent. This is because she would not have to incur a loss in output that is greater than the cost of information revelation. The intuition of the second result of Lemma 2 is consistent with Lemma 1 in the second benchmark case. If the principal's private information of the market condition is indeed



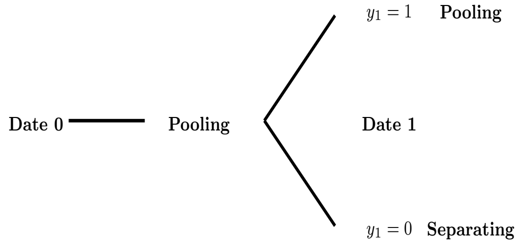
(a) PSS



(b) PSP



(c) PPP



(d) PPS

Figure 1.4: Possible Structures of Information Revelation II

Note: S represents separating; P represents pooling.

revealed through the contract, then the agent implements s_i in the bad market and s_s in the good market.

Proposition 3 ***The equilibrium structures of information revelation.***

- *If α (the probability of $m_1 = U$) and β (the probability of $m_2 = U$ conditional on $m_1 = U$) are sufficiently small, SSP is the equilibrium information structure.*
- *If α and β are sufficiently large, PPS is the equilibrium information structure.*
- *If α is sufficiently small and β is sufficiently large, PSS is the equilibrium information structure.*
- *If α is sufficiently large and β is sufficiently small, PPS is the equilibrium information structure if C_h is sufficiently low, and SSP is the equilibrium information structure if C_h is sufficiently high.*

Proposition 3 characterizes the equilibrium structures of information revelation. Figure 1.3b, Figure 1.4a and Figure 1.4d (with captions in blue) graphically represent the structures of information revelation of the three equilibrium contracts, each of which is unique given a set of parameter values. I choose to focus on SSP and PPS as shown in Figure 1.3b and 1.4d respectively. These two contracts possess an interesting feature

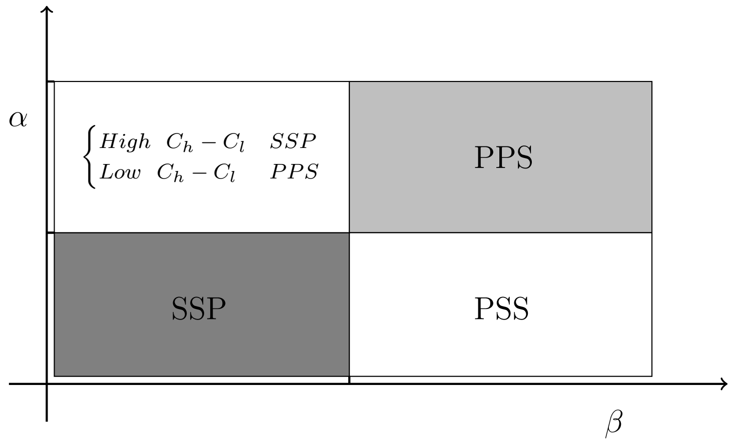


Figure 1.5: Equilibrium Contracts

of path-dependency in their structures of information revelation – whether information is revealed in the second period depends on the performance in the first period. For a full analysis of the three contracts, readers could refer to the Appendix 1.8. I call SSP *early adaptation* and PPS *late adaptation*.[18]

According to Proposition 3, although the degree and timing of insufficient adaptation differ among different equilibria, it is a robust and critical feature in every equilibrium. Full information revelation (SSS), which leads to full adaptation, however, is not an equilibrium. In Corollary 2, I show that the fully revealing contract can be characterized downward rigidly. But promising a non-decreasing contract is costly. Once the principal can choose when to reveal information, or once the imposed truth-telling constraints are lifted, insufficient adaptation arises in equilibrium. Intuitively, full information revelation is so costly that a principal in a good market chooses not to inform the agent of market changes.

Early adaptation. Early adaptation features full information revelation in the first period. Agents in two types of markets implement the strategy that adapts to the respective market condition. In the second period, the contract, however, reveals the market condition only following good performance. If the agent achieves good performance in the first period, she then, in the second period, implements the strategy that adapts to the new market condition. Otherwise she implements the market-sensitive strategy.

The intuition is as follows. A firm in a good market can maximize the benefit of information revelation in the first period if there is no information revelation following bad performance in the second period. It would be mistaken as in a bad market

[18]In fact, equilibrium PSS also involves late adaptation. To ease the exposition, late adaptation in the paper only refers to PPS.

following bad performance if it does not signal in the first period and incur a high incentive cost. Specifically, two forces work in the opposite directions under information asymmetry. I take an extreme scenario in which α and β approach zero for a simple illustration, since this is the environment in which full information revelation (SSS) is most likely to arise. But the principal chooses early adaptation (SSP) over full information revelation. Not revealing information following bad performance increases the *expected* second-period incentive cost of a firm in a good market by an amount of $(1-\theta_h)(C_h - C_l)$. However, a firm which operates in a good market in the first period is less likely than a firm in a bad market to achieve bad performance. As a result, the reduction in the cost of first-period information revelation, which is $(1-\theta_l)(C_h - C_l)$, outweighs the increase in incentive cost by an amount of $(\theta_h - \theta_l)(C_h - C_l)$. Early adaptation thus achieves the greater extent of information revelation than late adaptation.

Following bad performance, firms which operate in a deteriorating market could thus induce its agent to adopt the market-sensitive strategy at a lower cost in the second period due to pooling. But following good performance, those firms will be correctly identified. While firms which still operate in a good market keep implementing the market-sensitive strategy, firms in a deteriorating market have to act more conservatively, for instance, cutting costs, orienting businesses on local markets, becoming technologically moderate, and etc.. In other words, information revelation exhibits a form of procyclicality, that is, information environment in a hierarchical organization improves following good performance but worsens following bad performance.

Late adaptation. The contract offered in late adaptation reveals information only in the second period and only following bad first period performance. Although the agent is not able to learn the first period market condition from the contract, she tries to infer it from her past performance.

$$Pr(m_1 = U | y_1 = 0) = \frac{\alpha(1-\theta_h)}{\alpha(1-\theta_h) + (1-\alpha)(1-\theta_l)} < \alpha \tag{1.4.1}$$

As shown in Equation 1.4.1, bad past performance reinforces the agent's negative belief of the market condition. But such inference works to the disadvantage of a firm which operates in good market in the first period, as it is more likely for such a firm to stay in a good market than a firm in a bad market does. Because not revealing the market condition through contracts would cause a substantial increase in the cost of incentivising the agent to implement s_s, the principal, if she continues to operate in the good market, chooses to reveal the market condition following bad performance.

$$Pr(m_1 = U|y_1 = 1) = \frac{\alpha\theta_h}{\alpha\theta_h + (1-\alpha)\theta_l} > \alpha \tag{1.4.2}$$

On the contrary, good past performance reinforces the agent's positive belief of the market condition as shown in Equation 1.4.2. Such inference works to the advantage of a firm which operates in a good market. Because it is more likely for a firm in a good market to stay in a good market than a firm which operates in a bad market does, not revealing the market condition through contracts would not cause an increase in the cost of incentivising the agent to implement s_s that is greater than the cost of information revelation.

If information is not revealed following bad performance, the agent does not only require an incentive pay that is higher than if it is due to the negative inference in the second period but also in the first period. Because the agent knows that bad performance could lead to a high incentive pay being offered in the second period, the principal has to offer a even higher incentive pay $w_{11}^{UU} - w_{01}^{UU}$ in the first period to induce the agent's first-period effort. By committing to information revelation following bad performance, the principal in a good market is able to save the first-period incentive cost.

Following good performance, firms which operate in a deteriorating market could thus induce its agent to take s_s at a lower cost in the second period by pooling with firms in a good market. They are tempted to follow the market-sensitive strategy, as they could take advantage of the agent's wrong inference and therefore reduce the incentive cost. This result suggests that over-adoption of the market-sensitive strategy is more likely to arise when firms are performing well. One could also interpret the market-sensitive strategy as more risky, innovative, or expansive than the market-insensitive strategy. Under late adaptation, the economy overall exhibits excessive adoption of such strategies. But if the performance worsens, firms which operate in a deteriorating market can be correctly identified. While firms which still operate in a good market successfully encourage the agent to implement the market-sensitive strategy, firms in a deteriorating market have to act more conservatively. In contrast with early adaptation, information revelation in late adaptation, exhibits a form of countercyclicality, that is, information environment improves following bad performance but worsens following good performance.

Discussion. Which information revelation structure is more likely to arise in equilibrium? Figure 1.5 shows the different regions in which each equilibrium contract exists. If the ex-ante probability of a good market condition is high (α is high) and the probability of market deterioration is low (or β is high), then late adaptation occurs. If both are low, early adaptation occurs. In other words, if the distribution of market conditions reflects a dim prospect both today and tomorrow, early adaptation is the equilibrium path of adaptation. Intuitively, a pressing market condition makes a firm more willing to inform employees of the market changes and to motivate adaptation, as it increases the incentive cost of pooling relative to the cost of information revelation.

My model therefore predicts a cohort effect of initial market environment on shaping a firm's adaptation path. Other things being equal, firms which start its business under a more pressing market condition are more vibrant in adapting to external changes. This is also in reminiscent of Schumpeterian view that economic downturns play a positive role in promoting long-run productivity growth.

However, if the initial market condition is very likely to be good (high α) and the market does not change very persistently (low β), early adaptation can still survive as an equilibrium but only if the cost of implementing the market-sensitive strategy is very high relative to that of implementing the market-insensitive strategy. Intuitively, if α is high, only a high C_h relative to C_l can deter the firm in a bad market from pooling at the market-sensitive strategy. If C_h is not sufficiently high, the cost information revelation in the first period outweighs the reduction in incentive cost, which gives rise to late adaptation. If one interprets the difference between C_h and C_l as the drasticity of strategic changes, this model also implies that early adaptation is more likely to occur if the new market condition requries a drastic strategic change.

Both early and late adaptation exhibit inertia or stickiness in the adoption of new strategies, a consequence of the trade-off between information revelation and saving incentive cost. In early adaptation, following bad performance in the first period, a firm in a deteriorating market condition sticks to the strategy adopted in the first period and fails to implement the market-insensitive strategy, which has become more cost-effective in the changed market. In late adaptation, inertia occurs, however, following good performance. My model thus predicts that inertial implementation of strategies is more likely to occur a hierarchical organization in which employees are less informed of the market environment than the senior management. Moreover, such inertia takes different forms depending on a firm's past performance and the prospect of the market. Both failure and success could cause inertia.

To summarize, equilibrium contracts do not fully reveal market conditions. Full information revelation is so costly that a firm sometimes chooses not to inform the agent of market changes. As a result, the equilibrium path of adaptation, either in the form of early or late adaptation, entails adaptation failure. Moreover, information revelation, as well as adaptation, is path-dependent. Depending on underlying economic prospects, inertial implementation of an old strategy, a consequence of insufficient information revelation, may happen following either success or failure.

1.5 Equilibrium Compensation Structure

The previous section investigates the equilibrium structure of information revelation. In this section, I characterize equilibrium contract which supports the respective information revelation structure.

1.5.1 Contract under Early Adaptation

Proposition 4 describes the contract under early adaptation. The principal in a deteriorating market, following bad performance, offers the same compensation plan as the principal in a stable market, thus no information revelation following bad performance.

Proposition 4 ***Compensation structure under early adaptation.***

- *If $m_1 = U$, the principal commits to the following contract that restricts her to choosing from only two compensation plans in the second period:*
  1. *Compensation plan one: $w_{00}^{UD} = (1-\theta_l)(C_l - C_h\theta_l/\bar{\theta}) + \Delta + \beta\theta_l\Delta$, $w_{10}^{UD} = \frac{C_h}{\theta_h} - \beta\Delta + w_{00}^{UD}$, $w_{01}^{UD} = w_{00}^{UD} + C_h/\bar{\theta}$, and $w_{11}^{UD} = \frac{C_l}{\theta_l} + \frac{C_h}{\theta_h} - \beta\Delta + w_{00}^{UD}$;*
  2. *Compensation plan two: $w_{00}^{UU} = (1-\theta_l)(C_l - C_h\theta_l/\bar{\theta}) + \Delta + \beta\theta_l\Delta$, $w_{10}^{UU} = \frac{C_h}{\theta_h} + (1-\beta)\Delta + w_{00}^{UU}$, $w_{01}^{UU} = w_{00}^{UU} + C_h/\bar{\theta}$, and $w_{11}^{UU} = 2\frac{C_h}{\theta_h} + (1-\beta)\Delta + w_{00}^{UU}$.*
- *If $m_1 = D$, the principal offers the same contract as described in Proposition 1.*

The principal offers the first compensation plan if the market condition deteriorates and the second if it does not. Compensation structure in Proposition 4 exhibits two interesting features. First, performance-based pay is path-dependent and cannot be replicated by two short-term performance pay. To be clearer, following bad interim performance, the agent receives an extra amount of $w_{01}^{UD} - w_{00}^{UD} = w_{01}^{UU} - w_{00}^{UU} = C_h/\bar{\theta}$ if she achieves a high output. However, following good interim performance, the agent either receives an extra amount of $w_{11}^{UD} - w_{10}^{UD} = \frac{C_l}{\theta_l}$ or $w_{11}^{UU} - w_{10}^{UU} = \frac{C_h}{\theta_h}$ if she achieves a high output. Rewards for a high output in the second period following good and bad performance are not the same.

The use of long-term equity under information asymmetry is a direct implication of path-dependent information revelation. As argued in the previous section, revealing market condition only following good performance saves the principal's signalling cost in the first period more than the increase in the incentive cost following bad performance. Therefore, reward for high performance in the second period is higher following bad performance than following good performance ($C_h/\bar{\theta} = w_{01}^{UU} - w_{00}^{UU} > w_{11}^{UU} - w_{10}^{UU} = \frac{C_h}{\theta_h}$).

In contrast, the two benchmark cases show that rewards for a high output in the second period following good and bad performance are the same. In the first benchmark case, the principal in a good market does not need to inform as the agent is symmetrically informed of the market condition. In the second benchmark of information asymmetry, truth-telling constraints are imposed to ensure full information revelation and full adaptation. Neither of the two cases involve path-dependent information revelation, nor require the use of long-term incentive pay.

The second interesting feature of this contract is that, as shown in Corollary 1.5.1, it can always be implemented by a downward rigid contract.

Corollary 3 *Contract in Proposition 4 can be implemented by a downward rigid structure:*

- *In the first period,* $w_0^U = w_{00}^{UD}/2 - \beta\Delta/2$, $w_1^U = \frac{C_h}{\theta_h} + w_0^U$.
- *In the second period, following* $y_1 = 0$, $w_0^U = w_0^D = w_{00}^{UD}/2 + \beta\Delta/2$, $w_1^U = w_1^D = C_h/\bar{\theta} + w_0^U$.
- *Following* $y_1 = 1$, *the principal commits to paying either*
  - $w_0^U = \Delta + w_{00}^{UD}/2 - \beta\Delta/2$, $w_1^U = \frac{C_h}{\theta_h} + w_0^U$; *Or*
  - $w_0^D = w_{00}^{UD}/2 - \beta\Delta/2$, $w_1^D = \frac{C_l}{\theta_l} + w_0^D$.

Corollary suggests that contract under early adaptation is non-decreasing in both salary and performance-based pay. The principal offers a salary of $w_{00}^{UD}/2 - \beta\Delta/2$ in the first period. If the market stays good, the principal offers $w_{00}^{UD}/2 + \beta\Delta/2$ following bad performance and $\Delta + w_{00}^{UD}/2 - \beta\Delta/2$ following good performance. If the market worsens, the principal also offers $w_{00}^{UD}/2 + \beta\Delta/2$ following bad performance due to no information revelation and $w_{00}^{UD}/2 - \beta\Delta/2$ following good performance. The increase in salary is most pronounced following good performance and in a stable market, a situation in which the principal keeps revealing information.

Performance-based pay increases from $\frac{C_h}{\theta_h}$ in the first period to $C_h/\bar{\theta}$ in the second period following bad performance. This is to motivate the agent to implement the market-sensitive strategy under pooling. It stays the same in the second period following good performance in a non-deteriorating market but increases to $\frac{C_l}{\theta_l}$ in a deteriorating market, as information revelation following good performance allows the principal to offer an incentive pay to induce the adaptive strategy to be implemented.

1.5.2 Contract under Late Adaptation

Proposition 5 describes the contract under late adaptation. The principal in a good market does not reveal market condition in the first period or in the second period following good performance. First define $\alpha' = Pr(\theta_1 = \theta_h | y_1 = 1) = \alpha\theta_h/(\alpha\theta_h + (1 - \alpha)\theta_l)$ and $\bar{\theta}_\alpha = \alpha' q\theta_h + (1 - \alpha' q)\theta_l$.

Proposition 5 ***Compensation structure under late adaptation***.

- *If $m_1 = U$, the principal commits to the following contract that restricts her to choosing from only two compensation plans in the second period:*

  1. *Compensation plan one:* $w_{00}^{UD} = 0$, $w_{01}^{UD} = \frac{C_l}{\theta_l}$, $w_{10}^{UD} = \frac{1}{\theta_\alpha}\{(\theta_\alpha + 1)C_h + \alpha\beta\theta_h\Delta - ((\alpha\theta_h\bar{\theta} + (1-\alpha)\theta_l^2))C_h/\bar{\theta}_\alpha\}$, $w_{11}^{UD} = C_h/\bar{\theta}_\alpha + w_{10}^{UD}$;
  2. *Compensation plan two:* $w_{00}^{UU} = \Delta$, $w_{01}^{UU} = \frac{C_h}{\theta_h} + \Delta$, $w_{10}^{UU} = \frac{1}{\theta_\alpha}\{(\theta_\alpha + 1)C_h + \alpha\beta\theta_h\Delta - ((\alpha\theta_h\bar{\theta} + (1-\alpha)\theta_l^2))C_h/\bar{\theta}_\alpha\}$, $w_{11}^{UU} = C_h/\bar{\theta}_\alpha + w_{10}^{UU}$.

- *If $m_1 = D$, the principal offers the same contract.*

The principal in a good market in the first period offers the first compensation plan if the market condition deteriorates and the second if it does not. The principal in a bad market in the first period offers the same contract but chooses the first compensation plan. Compensation structure in Proposition 4 also exhibits two interesting features.

First, as in early adaptation, performance-based pay is also path-dependent and cannot be replicated by two short-term performance pay. To be clearer, following good performance, the agent receives an extra amount of $w_{11}^{UU} - w_{10}^{UU} = w_{11}^{UD} - w_{10}^{UD} = C_h/\bar{\theta}_\alpha$ if she achieves a high output in the second period. However, following bad performance, the agent either receives an extra amount of $w_{01}^{UD} - w_{00}^{UD} = \frac{C_l}{\theta_l}$ or $w_{01}^{UU} - w_{00}^{UU} = \frac{C_h}{\theta_h}$ if she achieves a high output in the second period. Reward for a high output in the second period following good and bad performance are not the same.

The intuition of this feature is similar to early adaptation. The use of long-term equity under information asymmetry is also a direct implication of path-dependent information revelation. As argued in the previous section, revealing market condition following bad performance saves the principal's incentive cost of pooling in the first period. Therefore, reward for high performance in the second period is higher following good performance than following bad performance($C_h/\bar{\theta}_\alpha = w_{11}^{UU} - w_{10}^{UU} > w_{01}^{UU} - w_{00}^{UU} = \frac{C_h}{\theta_h}$).

In contrast with early adaptation, Corollary 4 shows that a downward rigid contract cannot always be implemented in late adaptation.

Corollary 4 *Contract in Proposition 5 can be implemented in the following form:*

- *In the first period,* $w_1^U = w_1^D = w_{10}^{UD}$, $w_1^U = w_1^D = 0$.
- *In the second period, following* $y_1 = 1$, $w_1^U = w_1^D = C_h/\bar{\theta}_\alpha$, $w_1^U = w_1^D = 0$.
- *Following* $y_1 = 0$, *the principal commits to paying either*
  - $w_0^U = \Delta$, $w_1^U = \frac{C_h}{\theta_h} + w_0^U$; *Or*
  - $w_0^D = 0$, $w_1^D = \frac{C_l}{\theta_l}$.

If α and β are sufficiently large, under which late adaptation is mostly likely to be the equilibrium, performance pay in the first period is greater than that in the second period, $w_{10}^{UU} > w_{01}^{UU} - w_{00}^{UU}$ and $w_{10}^{UU} > w_{11}^{UU} - w_{10}^{UU}$.[19]

The contract in late adaptation does not reveal the market condition in the first period and only does so in the second period following good performance. Intuitively, to motivate the agent to implement the market-sensitive strategy under pooling in the first period, the principal in a good market has to provide a high performance-based pay, which explains $w_{10}^{UU} > w_{01}^{UU} - w_{00}^{UU}$. Because the agent could receive salary even following bad performance due to information revelation, the principal has to pay an even higher incentive reward to induce first period effort, which explains $w_{10}^{UU} > w_{11}^{UU} - w_{10}^{UU}$.

1.5.3 Long-term vs Short-term Contracts

So far my model suggests that a firm's incentive system is intimately interlinked with its adaptation path. Corollary 5 further suggests that long-term contracts impose a legacy problem that prevents full adaptation.

Corollary 5 *Contracts in early and late adaptation cannot be replicated by short-term contracts.*

In early adaptation, the principal in a good market saves the first period signalling cost by committing to information revelation only following good performance, at the expenses of an increase in the incentive cost following bad performance. In late adaptation, the principal in a good market saves the first period incentive cost by committing to information revelation only following bad performance, at the expenses of an increase in the signalling cost following bad performance. In neither of the two cases can short-term contracts replicate the long-term contracts.

[19] Or equivalently $w_{10}^{UD} > w_{11}^{UD} - w_{10}^{UD}$.

In early adaptation, while equilibrium long-term contracts improve ex-ante profit, they, ex post, create internal resistance by incumbent employees, which prevents compensation reconfiguration and full adaptation. The principal in a non-deteriorating market chooses not to reveal information following bad performance to save the first period signalling cost at the expense of an increase in the second period incentive pay, even in situations with low α and β. Can the principal still reveal the market condition following good performance by negotiating down the incentive pay and offering a salary? No. If the information was revealed, its employee, upon knowing the good market condition, would grow "vested interest", not agreeing to give up the overly-paid incentive pay.

The long-term commitment, therefore, limits the firm in a non-deteriorating market to restructure its incentive system cheaply and to induce adaptation in the second period. Consequently, a firm which faces a worsening market will not be willing to give up the contract in place to induce the adoption of the market-insensitive strategy, as the incentive cost if not pooling is too high. Internal resistance, therefore, endogenously arises in the firm's design of an optimal incentive system.

In late adaptation, the principal in a good market commits to choosing from only two compensation plans and information revelation following bad performance. Such a principal, therefore, cannot offer a new contract that is not informative of the market condition following bad performance, even in situations with high α and β. But still, informaiton asymmetry prevents the principal from revealing informaiton in the first period and in the second period following good performance.

1.6 Discussion

1.6.1 Government Intervention

Previous sections show that information asymmetry leads to a failure in information revelation and insufficient adaptation. This section discusses two policies that a government could either apply a subsidy rate or a tax rate to facilitate or direct adaptation. A firm is entitled to a subsidy if it achieves good performance and is charged a tax only if it achieves bad performance. Both the subsidy and tax are rate-based and are applied to firm's after-compensation earnings.

The *selective* policies create differential effects on a firm in a good market and a firm in a bad market. Specifically, the effective subsidy awarded to the firm in a good market is higher than the firm in a bad market, and the effective tax levied on the firm in a bad market is higher than that in a good market. The policies alleviate the truth-telling constraints in the sense that they make it more costly for the firm in a bad market to mimic. The signalling cost is reduced for the firm in a good market and it is thus more willing to choose the socially efficient contract. In a one-period model, the signalling cost is reduced from Δ to $\Delta/(\frac{1-t\theta_l}{1-t})$ if a subsidy rate t is applied and to $\Delta/(1+t(1-\theta_l))$ if a tax rate t is applied.

In the following analysis, I take the tax policy for a detailed illustration under early adaptation. Lemma 3 and Lemma 4 present the compensation structure under the second benchmark case in which truth-telling constraints are imposed and the compensation structure in early adaptation.

Lemma 3 ***A fully revealing contract under bad-performance tax rate*** t.
If $m_1 = U$, *the principal commits to such a contract that restricts her to only two compensation plans to choose from in the second period:*

1. *Compensation planone:* $w_{00}^{UD} = \Delta/(1+t(1-\theta_l)^2)$, $w_{01}^{UD} = \frac{C_l}{\theta_l} + w_{00}^{UD}$, $w_{10}^{UD} = \frac{C_h}{\theta_h} + w_{00}^{UD}$, *and* $w_{11}^{UD} = \frac{C_l}{\theta_l} + w_{10}^{UD}$;
2. *Compensation plan two:* $w_{00}^{UU} = \Delta/(1+t(1-\theta_l)^2) + \Delta/(1+t(1-\theta_l))$, $w_{01}^{UU} = \frac{C_l}{\theta_l} + w_{00}^{UU}$, $w_{10}^{UU} = \frac{C_h}{\theta_h} + w_{00}^{UU}$, *and* $w_{11}^{UU} = \frac{C_h}{\theta_h} + w_{10}^{UU}$.

If $m_1 = D$, *the principal offers the same contract as described in Proposition 1.*

Lemma 3 shows that both w_{00}^{UD} and w_{00}^{UU} are reduced compared to the level under zero tax. Bad performance tax alleviates the signalling problem imposed by information asymmetry.

Lemma 4 ***Early adaptation contract under bad-performance tax rate*** t

- *If* $m_1 = U$, *the principal commits to the following contract that restricts her to choosing from only two compensation plans in the second period:*

  1. *Compensation plan one:* $w_{00}^{UD} = ((1-\theta_l)(C_l - C_h\theta_l/\bar{\theta}) + \Delta + \beta\theta_l\Delta)/(1+t(1-\theta_l)^2)$, $w_{10}^{UD} = \frac{C_h}{\theta_h} - \beta\Delta/(1+t(1-\theta_l)^2) + w_{00}^{UD}$, $w_{01}^{UD} = C_h/\bar{\theta} + w_{00}^{UD}$, *and* $w_{11}^{UD} = \frac{C_l}{\theta_l} + w_{10}^{UD}$;
  2. *Compensation plan two:* $w_{00}^{UU} = ((1-\theta_l)(C_l - C_h\theta_l/\bar{\theta}) + \Delta + \beta\theta_l\Delta)/(1+t(1-\theta_l)^2)$, $w_{10}^{UU} = \frac{C_h}{\theta_h} + \Delta/(1+t(1-\theta_l)) - \beta\Delta/(1+t(1-\theta_l)^2) + w_{00}^{UU}$, $w_{01}^{UU} = C_h/\bar{\theta} + w_{00}^{UU}$, *and* $w_{11}^{UU} = \frac{C_h}{\theta_h} + w_{10}^{UU}$.

- *If* $m_1 = D$, *the principal offers the same contract as described in Proposition 1.*

Lemma 4 shows that both w_{00}^{UD} and $w_{10}^{UU} - w_{10}^{UD}$ are reduced compared to the level under zero tax. Bad performance tax alleviates the signalling problem imposed by information asymmetry.

Proposition 6 *If* $t \geq \frac{\theta_h - \theta_l}{(1-\theta_h)(1-\theta_l^2)}$*, the principal in a good market chooses the fully revealing contract over the early adaptation contract.*

In Proposition 6, I show that there exists a minimum level of tax rate above which the informed principal in a good market adopts the fully-revealing contract. The intuition is that the tax rate should be sufficiently high to make mimicking costly to the principal in a bad market. In addition, for government policies to facilitate adaptation, they need to be selective. One could easily show that taxing or subsidizing firms following both good and bad performance does not reduce the signalling cost. The discussion in this section therefore highlights the form and extent of public policies that the government should implement, especially in industries that are experiencing upgrade and restructuring.

1.6.2 Termination of Employment

In reality, the actual contract space includes other incentive tools in addition to compensation scheme. For instance, the principal could dismiss the agent following bad performance. In general, termination of employment reduces the rent that the agent could extract from the principal due to the protection of limited liability. However, since the model assumes that the probability of success is zero if the agent shirks, the rent reduction effect brought by termination threat does not exist.

One might argue that endogenizing termination gives the principal an additional signalling device. The principal in the good market could reveal her private information by committing *not* to dismiss the agent. This signalling device, however, is only useful if the principal in a bad market does not use it. Based on the argument in the previous paragraph, one could easily verify that firms that operate in the bad market will not use termination, which renders committing not to terminate employment useless as a signalling device.

Counter-intuitively, one might also argue that the principal in the good market could reveal her private information by committing to dismiss the agent following bad performance if the firm discontinues the business and finding a replacement is costly. This is theoretically sound as the principal in the good market is less likely to attain bad performance than the principal in the bad market. Committing to dismiss the agent following bad performance therefore incurs less profit loss.

Although the paper focuses on compensation contract as the sole signalling device, an informed principal could in practice employ multiple signalling devices including termination of employment to inform adaptation. A direct implication by allowing for more than one signalling device is a reduction in the signalling cost. Early adaptation may still survive as an equilibrium, because by committing to terminate the employment following bad performance could also help save the signalling cost in the first period. The basic mechanism underlying early adaptation is therefore not affected. However, late adaptation may not exist as an equilibrium, because termination following bad performance leads to no information revelation and thus no adaptation.

1.7 Implications and Conclusions

The results suggest that a firm may fail to adapt to market changes due to information asymmetry that widely exists in hierarchical organizations in which the senior management is more visionary of those changes. A firm needs to structure its employees' compensation contracts to both inform and motivate them to adapt. But it is costly to credibly inform employees of those changes and convince them of the efficacy of new strategies. A failure to overturn their belief about the changing market condition may lead to insufficient adaptation. Moreover, adaptation is path-dependent and inertial; depending on the distribution of market conditions, bad performance can either foster or suppress future adaptation. In fact, a more pressing market condition induces earlier adaptation and greater information revelation. Lastly, the contract that induces full adaptation is non-decreasing in both salary and performance-based pay, which, however, is too costly to offer in equilibrium. Equilibrium contracts impose a legacy problem that restrains the reconfiguration of the incentive system in place and hinders adaptation.

Those results give rise to a number of new empirical implications. First, firms that are more decentralized and are less subject to information asymmetry are less likely to encounter adaptation difficulty or inertial adoption of old strategies. Second, adaptation is more likely to occur under a more pressing market outlook. Third, given a pressing (resp., promising) market outlook, future adaptation is more likely to occur following good (resp., bad) performance. Fourth, firms that offer employees long-term contracts which promise non-decreasing compensation are more able to adapt to market changes.

Although most of the direct predictions of the model still need to be tested, there is some additional evidence in support of the forces underlying our model. Based on a new panel dataset on auto innovations, Aghion, Dechezleprêtre, Hemous, Martin, and Van Reenen (2014) find that a firm's propensity to innovate in clean technologies appears to be stimulated by its own past history of clean innovations. Tax-inclusive fuel prices (their proxy for a carbon tax) help overcome the inertia and induce firms to redirect technical change away from dirty innovation and toward clean innovation.

Finally, there are many directions in which the model can be extended. The model emphasizes the situation in which the principal is correct in her vision of market changes. However, one could also argue that the principal might only have superior information over the distribution of market changes, as they might not be fully sure how the market trend evolves in future. Another extension to consider is the case in which the principal is more informed of the change in macro-economic conditions and the agent is more informed of the change in local market conditions. The principal therefore is not only concerned with transmitting her private information to the agent but also soliciting private information from the agent. This is also a promising avenue for future theoretical and empirical explorations.

1.8 Appendix 1

Proof of Proposition 1.

Proof 1. First derive the contract if $\theta = \theta_l$ over time.

$$\max_{w\{..\}} \quad \theta_l^2(2 - w_{11}^{DD}) + \theta_l(1-\theta_l)(1 - w_{10}^{DD} + 1 - w_{01}^{DD}) + (1-\theta_l)^2(-w_{00}^{DD})$$

$$s.t. \quad \theta_l(w_{11}^{DD} - w_{10}^{DD}) \geq C_l$$

$$\theta_l(w_{01}^{DD} - w_{00}^{DD}) \geq C_l$$

$$\theta_l(w_{10}^{DD} - w_{00}^{DD} \geq C_l$$

I show that $w_{00}^{DD} = 0$. If $w_{00}^{DD} > 0$, reducing w_{00}^{DD} by ϵ could decrease w_{11}^{DD}, w_{10}^{DD} and w_{01}^{DD} by ϵ. The principal's profit could increase by $\theta_l^2\epsilon + \theta_l(1-\theta_l)\epsilon + (1-\theta_l)^2\epsilon = \epsilon$. Hence, $w_{01}^{DD} = \frac{C_l}{\theta_l}$, $w_{10}^{DD} = \frac{C_l}{\theta_l}$, $w_{11}^{DD} = 2\frac{C_l}{\theta_l}$ and $w_{00}^{DD} = 0$.

2. Assume that if agent knows the market condition θ_h, she takes the strategy that fits the market condition (verify later), derive the equilibrium contract.

$$\max_{w\{..\}} \quad \theta_h\{q(\theta_h(2 - w_{11}^{UU}) + (1-\theta_h)(1 - w_{10}^{UU})) + (1-q)(\theta_l(2 - w_{11}^{UD}) + (1-\theta_l)(1 - w_{10}^{UD}))\}$$

$$+(1-\theta_h)\{q(\theta_h(1 - w_{01}^{UU}) + (1-\theta_h)(0 - w_{00}^{UU})) + (1-q)(\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}))\}$$

$$s.t. \quad \theta_h(w_{11}^{UU} - w_{10}^{UU}) \geq C_h$$

$$\theta_h(w_{01}^{UU} - w_{00}^{UU}) \geq C_h$$

$$\theta_l(w_{11}^{UD} - w_{10}^{UD}) \geq C_l$$

$$\theta_l(w_{01}^{UD} - w_{00}^{UD}) \geq C_l$$

$$\theta_h(qw_{10}^{UU} + (1-q)w_{10}^{UD} \geq C_h$$

To save space, limited liability constraints are not listed.[20] Because Constraints A.6 and A.7 are binding, the principal who still operates in the good market in the second market will find it indifferent between revealing and not revealing information. For efficiency reason, I assume that the principal in such a situation chooses the basic compensation unit that reveals information.This argument applies to all the following analysis.

If $\theta = \theta_h$ at $t = 1$ and $\theta = \theta_h$ at $t = 2$, the principal will offer a contract $w_{01}^{UU} = \frac{C_h}{\theta_h}$, $w_{10}^{UU} \geq 0$, $w_{11}^{UU} = \frac{C_h}{\theta_h} + w_{10}^{UU}$ and $w_{00}^{UU} = 0$. If $\theta = \theta_h$ at $t = 1$ and $\theta = \theta_l$ at $t = 2$, the principal will offer a contract $w_{01}^{UD} = \frac{C_l}{\theta_l}$, $w_{10}^{UD} \geq 0$, $w_{11}^{UD} = \frac{C_l}{\theta_l} + w_{10}^{UD}$ and $w_{00}^{UD} = 0$.

It is easy to show that if $\theta = \theta_l$ at $t = 1$ and $\theta = \theta_l$ at $t = 2$, the principal will offer a contract $w_{01}^{DD} = \frac{C_l}{\theta_l}$, $w_{10}^{DD} = \frac{C_l}{\theta_l}$, $w_{11}^{DD} = \frac{2C_l}{\theta_l}$ and $w_{00}^{DD} = 0$.

[20] All contingent payment must be greater than or equal to zero.

3. Verify that under the above contract, if agent knows the market condition, she indeed takes the strategy that fits the market condition. Please refer to the proof of Lemma 1.

Set $w_{10}^{UD} = \frac{C_h}{\theta_h}$. Because $qw_{10}^{UU}+(1-q)w_{10}^{UD} = \frac{C_h}{\theta_h}$, $w_{10}^{UU} = \frac{C_h}{\theta_h}$. $w_{11}^{UD} = \frac{C_l}{\theta_l}+w_{10}^{UD} = \frac{C_l}{\theta_l}+\frac{C_h}{\theta_h}$. $w_{11}^{UU} = \frac{C_h}{\theta_h} + w_{10}^{UU} = \frac{C_l}{\theta_l} + \frac{C_h}{\theta_h}$.

This implementation does not rely on the market condition α and q.

Q.E.D.

Proof of Lemma 1.

Proof 1. If the agent knows $\theta = \theta_h$, she will not take the conservative strategy. Because $\frac{C_h}{\theta_h} < \frac{C_l}{\theta_l}$, principal will offer $\frac{C_h}{\theta_h}$ if $y = 1$. If the agent takes the conservative strategy, she gets $\theta_l \frac{C_h}{\theta_h} - C_l < 0$. If she takes the innovative strategy, she gets 0.

2. If the agent knows $\theta = \theta_l$, she will not take the conservative strategy. The principal offers $\frac{C_l}{\theta_l}$ if $y = 1$. If the agent takes the conservative strategy, she gets 0. If she takes the innovative strategy, she gets $\theta_l \frac{C_l}{\theta_l} - C_h < 0$.

Q.E.D.

Proof of Proposition 3.

Proof

$$\max_{w\{..\}} \quad \theta_h\{q(\theta_h(2-w_{11}^{UU}) + (1-\theta_h)(1-w_{10}^{UU})) + (1-q)(\theta_l(2-w_{11}^{UD}) + (1-\theta_l)(1-w_{10}^{UD}))\}$$
$$+(1-\theta_h)\{q(\theta_h(1-w_{01}^{UU}) + (1-\theta_h)(0-w_{00}^{UU})) + (1-q)(\theta_l(1-w_{01}^{UD}) + (1-\theta_l)(0-w_{00}^{UD}))\}$$

To save space, limited liability constraints are not listed. Constraint A.1-A.5 are the agent's incentive constraints.

$$s.t. \;\; \theta_h(w_{11}^{UU} - w_{10}^{UU}) \geq C_h \tag{1.8.1}$$
$$\theta_h(w_{01}^{UU} - w_{00}^{UU}) \geq C_h \tag{1.8.2}$$
$$\theta_l(w_{11}^{UD} - w_{10}^{UD}) \geq C_l \tag{1.8.3}$$
$$\theta_l(w_{01}^{UD} - w_{00}^{UD}) \geq C_l \tag{1.8.4}$$
$$\theta_h\{q(\theta_h(w_{11}^{UU} - w_{10}^{UU}) + w_{10}^{UU} - C_h) + (1-q)(\theta_l(w_{11}^{UD} - w_{10}^{UD}) + w_{10}^{UD} - C_l)\} \tag{1.8.5}$$
$$+(1-\theta_h)\{q(\theta_h(w_{01}^{UU} - w_{00}^{UU}) + w_{00}^{UU} - C_h) + (1-q)(\theta_l(w_{01}^{UD} - w_{00}^{UD}) + w_{00}^{UD} - C_l)\}$$
$$-C_h \geq q(\theta_h(w_{01}^{UU} - w_{00}^{UU}) + w_{00}^{UU} - C_h) + (1-q)(\theta_l(w_{01}^{UD} - w_{00}^{UD}) + w_{00}^{UD} - C_l$$

Constraint A.6-A.8 are the principal's truth-telling constraints.

$$\theta_l(2-w_{11}^{UD})+(1-\theta_l)(1-w_{10}^{UD}) \geq \theta_l(2-w_{11}^{UU})+(1-\theta_l)(1-w_{10}^{UU}) \quad (1.8.6)$$
$$\theta_l(1-w_{01}^{UD})+(1-\theta_l)(0-w_{00}^{UD}) \geq \theta_l(1-w_{01}^{UU})+(1-\theta_l)(0-w_{00}^{UU}) \quad (1.8.7)$$
$$\theta_l^2(2-w_{11}^{DD})+\theta_l(1-\theta_l)(1-w_{10}^{DD})+\theta_l(1-\theta_l)(1-w_{01}^{DD})+(1-\theta_l)^2(0-w_{00}^{DD}) \quad (1.8.8)$$
$$\geq \theta_l^2(2-w_{11}^{UD})+\theta_l(1-\theta_l)(1-w_{10}^{UD})+\theta_l(1-\theta_l)(1-w_{01}^{UD})+(1-\theta_l)^2(0-w_{00}^{UD})$$

Interim incentive constraints of the agent guarantee that the contract is interim renegotiation-proof.

$$(A.5)\to(A.5') \qquad q\theta_h(w_{10}^{UU}-w_{00}^{UU})+(1-q)\theta_h(w_{10}^{UD}-w_{00}^{UD})\geq C_h$$
$$(A.6)\to(A.6') \qquad \theta_l(w_{11}^{UU}-w_{10}^{UU})-\theta_l(w_{11}^{UD}-w_{10}^{UD})\geq w_{10}^{UD}-w_{10}^{UU}$$
$$(A.7)\to(A.7') \qquad \theta_l(w_{01}^{UU}-w_{00}^{UU})-\theta_l(w_{01}^{UD}-w_{00}^{UD})\geq w_{00}^{UD}-w_{00}^{UU}$$
$$(A.8)\to(A.8') \qquad \theta_l^2 w_{11}^{UD}+\theta_l(1-\theta_l)(w_{10}^{UD}+w_{01}^{UD})+(1-\theta_l)^2 w_{00}^{UD}\geq 2C_l$$

Rearrange the above equations, I get:

$$(A.1),(A.6') \qquad w_{10}^{UU}\geq w_{10}^{UD}+\Delta$$
$$(A.2),(A.4),(A.7') \qquad w_{00}^{UU}\geq w_{00}^{UD}+\Delta$$
$$(A.5)',(A.6') \qquad w_{10}^{UD}\geq w_{00}^{UD}+\frac{C_h}{\theta_h}$$

Substitute all into $(A.8')$, one could solve for w_{00}^{UD}, and all other variables.

$$w_{00}^{UD}=\Delta; \quad w_{01}^{UD}=\frac{C_l}{\theta_l}+\Delta; \quad w_{10}^{UD}=\frac{C_h}{\theta_h}+\Delta; \quad w_{11}^{UD}=\frac{C_l}{\theta_l}+\frac{C_h}{\theta_h}+\Delta; or$$
$$w_{00}^{UU}=2\Delta; \quad w_{01}^{UU}=\frac{C_l}{\theta_l}+2\Delta; \quad w_{10}^{UU}=\frac{C_h}{\theta_h}+2\Delta; \quad w_{11}^{UU}=2\frac{C_h}{\theta_h}+2\Delta.$$

Q.E.D.

Proof of Corollary 2.

Proof

1. If $\theta_1=\theta_h$ in the first period, because $w_{11}^{UD}-w_{01}^{UD}=w_{10}^{UD}-w_{00}^{UD}=w_{11}^{UU}-w_{01}^{UU}=w_{10}^{UU}-w_{00}^{UU}=\frac{C_h}{\theta_h}$, the first period incentive pay is $\frac{C_h}{\theta_h}$ if $y_1=1$.
2. If $\theta_2=\theta_h$ in the second period, because $w_{11}^{UU}-w_{10}^{UU}=w_{01}^{UU}-w_{00}^{UU}=\frac{C_h}{\theta_h}$, the second period incentive pay is $\frac{C_h}{\theta_h}$ if $y_2=1$.
3. If $\theta_2=\theta_l$ in the second period, because $w_{11}^{UD}-w_{10}^{UD}=w_{01}^{UD}-w_{00}^{UD}=\frac{C_l}{\theta_l}$, the first period incentive pay is $\frac{C_l}{\theta_l}$ if $y_2=1$.
4. Because $w_{11}^{UU}-w_{11}^{UD}=w_{10}^{UU}-w_{10}^{UD}=w_{01}^{UU}-w_{01}^{UD}=w_{00}^{UU}-w_{00}^{UD}=\Delta$, if $\theta_2=\theta_h$, the second period salary should increase by Δ from the first period. Because the hl type

has a salary Δ since $w_{00}^{UD} = \Delta$. I set first period salary to $\frac{1}{2}\Delta$ and second period to $\frac{1}{2}\Delta$ for hl and $\frac{3}{2}\Delta$ for hh.

Q.E.D.

Proof of Proposition 3.

Proof Step 1 is to characterize the contract under SSP.

1. Define $\bar{\theta} = q\theta_h + (1-q)\theta_l$. From Lemma 1, if the agent knows information, she takes the adaptive strategy which fits the external market. If the agent does not know the information, I first characterize the contract under which the agent takes the innovative strategy under pooling.

$$\max_{w\{..\}} \theta_h\{q(\theta_h(2-w_{11}^{UU}) + (1-\theta_h)(1-w_{10}^{UU})) + (1-q)(\theta_l(2-w_{11}^{UD}) + (1-\theta_l)(1-w_{10}^{UD}))\}$$
$$+(1-\theta_h)\{q(\theta_h(1-w_{01}^{UX}) + (1-\theta_h)(0-w_{00}^{UX})) + (1-q)(\theta_l(1-w_{01}^{UX}) + (1-\theta_l)(0-w_{00}^{UX}))\}$$

To save space, limited liability constraints are not listed. Constraints A.9-A.12 are the agent's incentive constraints.

$$s.t.\ \theta_h(w_{11}^{UU} - w_{10}^{UU}) \geq C_h \quad (1.8.9)$$
$$\theta_l(w_{11}^{UD} - w_{10}^{UD}) \geq C_l \quad (1.8.10)$$
$$\bar{\theta}(w_{01}^{UX} - w_{00}^{UX}) \geq C_h \quad (1.8.11)$$
$$\theta_h\{q(\theta_h(w_{11}^{UU} - w_{10}^{UU}) + w_{10}^{UU} - C_h) + (1-q)(\theta_l(w_{11}^{UD} - w_{10}^{UD}) + w_{10}^{UD} - C_l)\} \quad (1.8.12)$$
$$+(1-\theta_h)\{\bar{\theta}(w_{01}^{UX} - w_{00}^{UX}) + w_{00}^{UX} - C_h)\} - C_h \geq \bar{\theta}(w_{01}^{UX} - w_{00}^{UX}) + w_{00}^{UX} - C_h$$

Constraint A.13 and A.14 are the principal's truth-telling constraints.

$$\theta_l(2-w_{11}^{UD}) + (1-\theta_l)(1-w_{10}^{UD}) \geq \theta_l(2-w_{11}^{UU}) + (1-\theta_l)(1-w_{10}^{UU}) \quad (1.8.13)$$
$$\theta_l^2(2-w_{11}^{DD}) + \theta_l(1-\theta_l)(1-w_{10}^{DD}) + \theta_l(1-\theta_l)(1-w_{01}^{DD}) + (1-\theta_l)^2(0-w_{00}^{DD}) \quad (1.8.14)$$
$$\geq \theta_l^2(2-w_{11}^{UD}) + \theta_l(1-\theta_l)(1-w_{10}^{UD}) + \theta_l(1-\theta_l)(1-w_{01}^{UX}) + (1-\theta_l)^2(0-w_{00}^{UX})$$

Rearrange the above, I get the following:

$$(A.13) \to (A.13') \qquad q\theta_h(w_{10}^{UU} - w_{00}^{UU}) + (1-q)\theta_h(w_{10}^{UD} - w_{00}^{UD}) \geq C_h$$
$$(A.14) \to (A.14') \qquad \theta_l^2 w_{11}^{UD} + \theta_l(1-\theta_l)(w_{10}^{UD} + w_{01}^{UX}) + (1-\theta_l)^2 w_{00}^{UX} \geq 2C_l$$

From above, we obtain the following compensation component expressed in w_{00}^{UX}:

$$\begin{aligned}
w_{10}^{UU} &\geq w_{00}^{UX} + \frac{C_h}{\theta_h} + (1-q)\Delta \\
w_{01}^{UX} &\geq w_{00}^{UX} + \frac{C_h}{\bar{\theta}} \\
w_{10}^{UD} &\geq w_{00}^{UX} + \frac{C_h}{\theta_h} - q\Delta \\
w_{11}^{UD} &\geq w_{00}^{UX} + \frac{C_h}{\theta_h} + \frac{C_l}{\theta_l} - q\Delta \\
w_{11}^{UU} &\geq w_{00}^{UX} + 2\frac{C_h}{\theta_h} + (1-q)\Delta
\end{aligned}$$

Substitute all into $A.14'$, one could solve for w_{00}^{UX}, and all other variables.

$$w_{00}^{UX} = 2C_l - \frac{\bar{\theta}+\theta_h}{\bar{\theta}\theta_h}\theta_l C_h - \theta_l(C_l - \frac{\theta_l}{\bar{\theta}}C_h) + q\theta_l\Delta$$

2. Verify if the principal offers a contract which induces the agent to take the conservative strategy under pooling $(w_{00}^{UX}, w_{01}^{UX})$, the principal will deviate to separate and offer a different contract which induces the agent to take the innovative strategy. In other words, there exists a contract which induces pooling at the innovative strategy which defeats the contract inducing pooling at the conservative strategy.

   Following bad performance $y_1 = 0$, if the principal in the good market deviates and offers $w_{01}^{UU} = \frac{C_h}{\theta_h}$ and $w_{00}^{UU} = \Delta + w_{00}^{UX}$, the agent could get $\theta_h w_{01}^{UU} + (1-\theta_h)w_{00}^{UU} - C_h = \Delta + w_{00}^{UX}$, higher than $\theta_l w_{01}^{UX} + (1-\theta_l)w_{00}^{UX} - C_l = w_{00}^{UX}$ under w_{01}^{UX} and w_{00}^{UX}. The principal in the good market gets $\theta_h(1 - w_{01}^{UU}) - (1-\theta_h)w_{00}^{UU}$, which is higher than $\theta_l(1 - w_{01}^{UX}) - (1-\theta_l)w_{00}^{UX}$ under w_{01}^{UX} and w_{00}^{UX}, the difference is $(\theta_h - \theta_l)(1 - \frac{C_h}{\theta_h})$. Given that $w_{00}^{UU} = \Delta + w_{00}^{UX}$, the principal in the bad market will not want to mimic. Based on the concept of Undefeated Equilibrium, if the principal in the bad market does not follow, she will be considered as the bad type and offer a contract that fully separates himself. This contract corresponds to an annual bonus of $\frac{C_l}{\theta_l}$ and zero salary. The principal gets no less profit from this contract than (w_{01}^{UX}, w_{00}^{UX}) and is thus better off than (w_{01}^{UU}, w_{00}^{UU}).

Step 2 is to characterize the contract under PPS.

1. If the principal does not reveal the market condition in the first period, the agent will have to infer it from past performance following good performance. Following bad performance, the principal reveals the market condition, thus the agent does not have to infer it on her own.

$$\alpha' = Pr(\theta_1 = \theta_h | y_1 = 1) = \frac{\alpha\theta_h}{\alpha\theta_h + (1-\alpha)\theta_l} > \alpha \qquad (1.8.15)$$

Define $\bar{\theta}_\alpha = \alpha' q\theta_h + (1-\alpha' q)\theta_l$. From Lemma 1, if the agent knows information, she takes the adaptive strategy which fits the external market. If the agent does not know the information, I first characterize the contract under which the agent takes the innovative strategy under pooling.

$$\max_{w\{..\}} \ \theta_h\{q(\theta_h(2-w_{11}^{XX})+(1-\theta_h)(1-w_{10}^{XX}))+(1-q)(\theta_l(2-w_{11}^{XX})+(1-\theta_l)(1-w_{10}^{XX}))\}$$
$$+(1-\theta_h)\{q(\theta_h(1-w_{01}^{XU})+(1-\theta_h)(0-w_{00}^{XU}))+(1-q)(\theta_l(1-w_{01}^{XD})+(1-\theta_l)(0-w_{00}^{XD}))\}$$

To save space, limited liability constraints are not listed. Constraint A.16-A.19 are the agent's incentive constraints, Constraint A.20 is the agent's project choice constraint.

$$s.t.\ \bar{\theta}_\alpha(w_{11}^{XX}-w_{10}^{XX}) \geq C_h \tag{1.8.16}$$
$$\theta_h(w_{01}^{XU}-w_{00}^{XU}) \geq C_h \tag{1.8.17}$$
$$\theta_l(w_{01}^{XD}-w_{00}^{XD}) \geq C_l \tag{1.8.18}$$

$$\begin{aligned}
&\alpha\{\theta_h q(\theta_h(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h)+(1-q)\theta_h(\theta_l(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h) \\
&+(1-\theta_h)q(\theta_h(w_{01}^{XU}-w_{00}^{XU})+w_{00}^{XU}-C_h)+(1-\theta_h)(1-q)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)\} \\
&(1-\alpha)\{\theta_l(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h)+(1-\theta_l)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)\}-C_h \\
&\geq \alpha q(\theta_h(w_{01}^{XU}-w_{10}^{XX})+w_{00}^{XU}-C_h)+(1-\alpha q)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)
\end{aligned} \tag{1.8.19}$$

$$\begin{aligned}
&\alpha\{\theta_h q(\theta_h(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h)+(1-q)\theta_h(\theta_l(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h) \\
&+(1-\theta_h)q(\theta_h(w_{01}^{XU}-w_{00}^{XU})+w_{00}^{XU}-C_h)+(1-\theta_h)(1-q)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)\} \\
&(1-\alpha)\{\theta_l(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h)+(1-\theta_l)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)\}-C_h \\
&\geq \theta_l(\theta_l(w_{11}^{XX}-w_{10}^{XX})+w_{10}^{XX}-C_h)+(1-\theta_l)(\theta_l(w_{01}^{XD}-w_{00}^{XD})+w_{00}^{XD}-C_l)-C_l
\end{aligned} \tag{1.8.20}$$

I verify in the next step, Constraint A.20 is redundant. Constraint A.21 is the principal's truth-telling constraint.

$$w_{00}^{XU}-w_{00}^{XD} \geq \Delta \tag{1.8.21}$$

Constraint A.22 is the mimicking constraint of the principal in the bad market. One could verify that in parameter ranges in which PPS is the equilibrium, A.22 will be automatically satisfied.

$$\begin{aligned}
&\theta_l(\theta_l(2-w_{11}^{XX})+(1-\theta_l)(1-w_{10}^{XX}))+(1-\theta_l)(\theta_l(1-w_{01}^{XD})+(1-\theta_l)(0-w_{00}^{XD})) \\
&\geq \theta_l(\theta_l(2-\frac{C_l}{\theta_l})+(1-\theta_l)(1-\frac{C_l}{\theta_l}))+(1-\theta_l)(\theta_l(1-\frac{C_l}{\theta_l})+(1-\theta_l)(0-0))
\end{aligned} \tag{1.8.22}$$

To solve the above maximization problem, I get the following contract:

$$w_{00}^{XD} = 0, w_{01}^{XD} = \frac{C_l}{\theta_l}, w_{00}^{XU} = \Delta, w_{01}^{XU} = \frac{C_h}{\theta_h} + \Delta, w_{11}^{XX} = \frac{C_h}{\bar{\theta}_\alpha} + w_{10}^{XX}$$
$$w_{10}^{XX} = \frac{1}{\theta_\alpha}\{(\theta_\alpha + 1)C_h + q\alpha\theta_h\Delta - (\alpha\theta_h\bar{\theta} + (1-\alpha)\theta_l^2)\frac{C_h}{\bar{\theta}_\alpha}\}$$

2. Verify that if the principal offers a contract which induces the agent to take the conservative strategy under pooling, the principal will deviate and offer a different contract which induces the agent to take the innovative strategy. In other words, there exists a contract which induces pooling at the innovative strategy which defeats the contract inducing pooling at the conservative strategy.

   Following the argument in the previous step of SSP, one could easily verify that pooling at conservative strategy in the second period is not renegotiation proof. The principal in the good market will always want to induce the agent to undertake the innovative strategy. I next verify the principal in the good market will also not pool at the conservative strategy in the first period. Assume that the principal offer w_{11}^{UX}, w_{10}^{UX}, w_{01}^{UU}, w_{00}^{UU}, w_{01}^{UD} and w_{00}^{UD}. To prevent the principal in the bad market from mimicking, the following equation must hold:

$$\begin{aligned}&\theta_l(\theta_l(2 - w_{11}^{UX}) + (1-\theta_l)(1 - w_{10}^{UX})) + (1-\theta_l)(\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}) \quad (1.8.23)\\ &\quad\le \theta_l(\theta_l(2 - w_{11}^{XX}) + (1-\theta_l)(1 - w_{10}^{XX})) + (1-\theta_l)(\theta_l(1 - w_{01}^{XD}) + (1-\theta_l)(0 - w_{00}^{XD})\end{aligned}$$

   If Constraint A.24 and A.25 hold, then Constraint A.23 holds.

$$\theta_l(2 - w_{11}^{UX}) + (1-\theta_l)(1 - w_{10}^{UX})) \le \theta_l(2 - w_{11}^{XX}) + (1-\theta_l)(1 - w_{10}^{XX}) \quad (1.8.24)$$
$$\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}) \le \theta_l(1 - w_{01}^{XD}) + (1-\theta_l)(0 - w_{00}^{XD}) \quad (1.8.25)$$

   Constraint A.24 and A.25 imply that $w_{10}^{UX} \ge w_{10}^{XX} + \Delta$ and $w_{00}^{UD} \ge w_{00}^{XD} + \Delta$. If the principal in the good market deviates, the change in profit is:

$$\begin{aligned}&\pi^D - \pi^{ND}\\ =&\theta_h\{q(\theta_h(2 - w_{11}^{UX}) + (1-\theta_h)(1 - w_{10}^{UX})) + (1-q)(\theta_l(2 - w_{11}^{UX}) + (1-\theta_l)(1 - w_{10}^{UX}))\}\\ &+(1-\theta_h)\{q(\theta_h(1 - w_{01}^{UU}) + (1-\theta_h)(0 - w_{00}^{UU})) + (1-q)(\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}))\}\\ &-\theta_l(\theta_l(2 - w_{11}^{XX}) + (1-\theta_l)(1 - w_{10}^{XX})) - (1-\theta_l)(\theta_l(1 - w_{01}^{XD}) + (1-\theta_l)(0 - w_{00}^{XD}))\\ \ge&\theta_h\{q(\theta_h(2 - w_{11}^{UX}) + (1-\theta_h)(1 - w_{10}^{UX})) + (1-q)(\theta_l(2 - w_{11}^{UX}) + (1-\theta_l)(1 - w_{10}^{UX}))\}\\ &+(1-\theta_h)\{q(\theta_h(1 - w_{01}^{UU}) + (1-\theta_h)(0 - w_{00}^{UU})) + (1-q)(\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}))\}\\ &-\theta_h(\theta_l(2 - w_{11}^{XX}) + (1-\theta_l)(1 - w_{10}^{XX})) - (1-\theta_h)(\theta_l(1 - w_{01}^{XD}) + (1-\theta_l)(0 - w_{00}^{XD}))\end{aligned}$$

Substitute w_{10}^{UX} and w_{10}^{UX} into the above equation,

$$\pi^D - \pi^{ND} \geq q(\theta_h - \theta_l)(1 - \frac{C_h}{\theta_h}) + \theta_h(1-q)\Delta \geq 0$$

Based on the concept of Undefeated Equilibrium, if the principal in the bad market does not follow, she will be considered as the bad type and offer a contract that fully separates himself. This contract corresponds to an annual bonus of $\frac{C_l}{\theta_l}$ and zero salary. The principal gets no less profit from this contract than the pooling contract inducing conservative strategy, and is thus better off than (w_{10}^{UX}, w_{11}^{UX}, w_{01}^{UD}, w_{00}^{UD}).

Step 3 Proof of contract under other information revelation structures – SPS.

1. Characterize the contract under SPS.

$$\max_{w\{..\}} \; \theta_h\{q(\theta_h(2 - w_{11}^{UX}) + (1-\theta_h)(1 - w_{10}^{UX})) + (1-q)(\theta_l(2 - w_{11}^{UX}) + (1-\theta_l)(1 - w_{10}^{UX}))\}$$
$$+(1-\theta_h)\{q(\theta_h(1 - w_{01}^{UU}) + (1-\theta_h)(0 - w_{00}^{UU})) + (1-q)(\theta_l(1 - w_{01}^{UD}) + (1-\theta_l)(0 - w_{00}^{UD}))\}$$

To save space, limited liability constraints are not listed. Constraint A.26-A.29 are the agent's incentive constraints.

$$s.t. \;\; \bar{\theta}(w_{11}^{UX} - w_{10}^{UX}) \geq C_h \tag{1.8.26}$$
$$\theta_h(w_{01}^{UU} - w_{00}^{UU}) \geq C_h \tag{1.8.27}$$
$$\theta_l(w_{01}^{UD} - w_{00}^{UD}) \geq C_l \tag{1.8.28}$$
$$\begin{aligned}\theta_h\{\bar{\theta}(w_{11}^{UX} - w_{10}^{UX}) + w_{10}^{UX} - C_h)\} + (1-\theta_h)\{q(\theta_h(w_{01}^{UU} - w_{00}^{UU}) + w_{00}^{UU} - C_h) \\ +(1-q)(\theta_l(w_{01}^{UD} - w_{00}^{UD}) + w_{00}^{UD} - C_l)\} - C_h \geq q(\theta_h(w_{01}^{UU} - w_{00}^{UU}) + w_{00}^{UU} - C_h) \\ +(1-q)(\theta_l(w_{01}^{UD} - w_{00}^{UD}) + w_{00}^{UD} - C_l)\end{aligned} \tag{1.8.29}$$

Constraint A.30 and A.31 are the principal's truth-telling constraints.

$$w_{00}^{UU} \geq w_{00}^{UD} + \Delta \tag{1.8.30}$$
$$\theta_l^2 w_{11}^{UX} + \theta_l(1-\theta_l)w_{10}^{UX} + \theta_l(1-\theta_l)(1 - w_{01}^{UD}) + (1-\theta_l)^2 w_{00}^{UD} \geq 2C_l \tag{1.8.31}$$

To solve the above maximization problem, I get the following contract:

$$w_{00}^{UD} = (1 - q\theta_l)\Delta, w_{01}^{UD} = \frac{C_l}{\theta_l} + w_{00}^{UD}, w_{00}^{UU} = w_{00}^{UD} + \Delta$$
$$w_{01}^{UU} = \frac{C_h}{\theta_h} + \Delta + w_{00}^{UD}, w_{10}^{UX} = \frac{C_h}{\theta_h} + q\Delta + w_{00}^{UD}, w_{11}^{UX} = \frac{C_h}{\bar{\theta}} + \frac{C_h}{\theta_h} + q\Delta + w_{00}^{UD}$$

As in the derivation of SSP and PPS, one could easily verify that if the principal offers a contract which induces the agent to take the conservative strategy under pooling, the principal will deviate and offer a different contract which induces the agent to take the innovative strategy.

2. Prove that the principal in the good market is strictly better off in the SSS than in SPS.

$$
\begin{aligned}
\Delta\pi =& \mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{SPS}\right] \\
=& \theta_h\{q(\theta_h(\frac{C_h}{\theta_h} + w_{00}^{UD} + \Delta - \frac{C_h}{\theta_h} - 2\Delta) + (1-\theta_h)(w_{00}^{UD} + \Delta - 2\Delta)) \\
&+ (1-q)(\theta_l(\frac{C_l}{\theta_l} + w_{00}^{UD} - \frac{C_l}{\theta_l} - \Delta) + (1-\theta_l)(w_{00}^{UD} - \Delta))\} \\
&+ (1-\theta_h)\{q(\theta_h(\frac{C_h}{\bar{\theta}} + w_{00}^{UD} + q\Delta - \frac{C_h}{\theta_h} - 2\Delta) + (1-\theta_h)(w_{00}^{UD} + q\Delta - 2\Delta)) \\
&+ (1-q)(\theta_l(\frac{C_l}{\bar{\theta}} + w_{00}^{UD} + q\Delta - \frac{C_l}{\theta_l} - \Delta) + (1-\theta_l)(w_{00}^{UD} + q\Delta - \Delta))\} \\
=& - q\theta_l\Delta + \theta_l(C_l - \frac{\theta_l}{\bar{\theta}}C_h) + \theta_h C_h - \theta_h(qC_h + (1-q)C_l)
\end{aligned}
$$

I then verify the monotonicity of $\Delta\pi$ over $q \in [0,1]$.

$$
\begin{aligned}
\frac{\partial \Delta\pi}{\partial q} &= -\theta_l\Delta + \theta_h(C_h - C_l) + \frac{\theta_l^2}{\bar{\theta}^2}C_h(\theta_h - \theta_l) \\
(\text{let } q \to 0) &\to -\theta_l\Delta + \theta_h(C_h - C_l) + C_h(\theta_h - \theta_l) \\
&= (\theta_h - \theta_l)\Delta \\
&> 0
\end{aligned}
$$

In addition, at $q = 0$ and $q = 1$ $\Delta\pi$ is non-negative:

$$
\begin{aligned}
\Delta\pi(q=0) &= (\theta_h - \theta_l)(C_h - C_l) > 0 \\
\Delta\pi(q=1) &= -\theta_l\Delta + \theta_l\Delta + \theta_h C_h - \theta_h C_h = 0
\end{aligned}
$$

Then $\Delta\pi$ is increasing over $q \in [0,1]$. In other words, the principal in the good market is strictly better off in the SSS than in SPS. Following the proof here, one could verify that SPP will be dominated by SSP because the principal in the good market will not want to pool following good performance.

Step 4 Proof of contract under other information revelation structures – PSS.

1. Characterize the contract under PSS.

$$
\begin{aligned}
\max_{w\{..\}} \; & \theta_h\{q(\theta_h(2 - w_{11}^{XU}) + (1-\theta_h)(1 - w_{10}^{XU})) + (1-q)(\theta_l(2 - w_{11}^{XD}) + (1-\theta_l)(1 - w_{10}^{XD}))\} \\
+ & (1-\theta_h)\{q(\theta_h(1 - w_{01}^{XU}) + (1-\theta_h)(0 - w_{00}^{XU})) + (1-q)(\theta_l(1 - w_{01}^{XD}) + (1-\theta_l)(0 - w_{00}^{XD}))\}
\end{aligned}
$$

To save space, limited liability constraints are not listed. Constraint A.32-A.36 are the agent's incentive constraints, Constraint A.37 is the agent's project choice constraint.

$$s.t.\ \theta_h(w_{11}^{XU} - w_{10}^{XU}) \geq C_h \tag{1.8.32}$$
$$\theta_h(w_{01}^{XU} - w_{00}^{XU}) \geq C_h \tag{1.8.33}$$
$$\theta_l(w_{11}^{XD} - w_{10}^{XD}) \geq C_l \tag{1.8.34}$$
$$\theta_l(w_{01}^{XD} - w_{00}^{XD}) \geq C_l \tag{1.8.35}$$

$$\begin{aligned}&\alpha\{q(\theta_h^2 w_{11}^{XU} + \theta_h(1-\theta_h)(w_{10}^{XU} + w_{01}^{XU}) + (1-\theta_h)^2 w_{00}^{XU} - C_h) + (1-q)(\theta_h\theta_l w_{11}^{XD} \\ &\quad +\theta_h(1-\theta_l)w_{10}^{XD} + \theta_l(1-\theta_h)w_{01}^{XD} + (1-\theta_h)(1-\theta_l)w_{00}^{XD}) - C_l)\} \\ &\quad +(1-\alpha)\{\theta_l^2 w_{11}^{XD} + \theta_l(1-\theta_l)(w_{10}^{XD} + w_{01}^{XD}) + (1-\theta_l)^2 w_{00}^{XD} - C_l\} - C_h \\ &\quad \geq \alpha\{q(\theta_h w_{01}^{XU} + (1-\theta_h)w_{00}^{XU} - C_h) + (1-q)(\theta_l w_{01}^{XD} + (1-\theta_l)w_{00}^{XD} - C_l)\} \\ &\quad +(1-\alpha)(\theta_l w_{01}^{XD} + (1-\theta_l)w_{00}^{XD} - C_l)\end{aligned} \tag{1.8.36}$$

$$\begin{aligned}&\alpha\{q(\theta_h^2 w_{11}^{XU} + \theta_h(1-\theta_h)(w_{10}^{XU} + w_{01}^{XU}) + (1-\theta_h)^2 w_{00}^{XU} - C_h) + (1-q)(\theta_h\theta_l w_{11}^{XD} \\ &\quad +\theta_h(1-\theta_l)w_{10}^{XD} + \theta_l(1-\theta_h)w_{01}^{XD} + (1-\theta_h)(1-\theta_l)w_{00}^{XD} - C_l)\} \\ &\quad +(1-\alpha)\{\theta_l^2 w_{11}^{XD} + \theta_l(1-\theta_l)(w_{10}^{XD} + w_{01}^{XD}) + (1-\theta_l)^2 w_{00}^{XD} - C_l\} - C_h \\ &\quad \geq \theta_l^2 w_{11}^{XD} + \theta_l(1-\theta_l)(w_{10}^{XD} + w_{01}^{XD}) + (1-\theta_l)^2 w_{00}^{XD} - 2C_l\end{aligned} \tag{1.8.37}$$

Constraint A.37 and A.38 are the principal's truth-telling constraints.

$$w_{10}^{XU} - w_{10}^{XD} \geq \Delta \tag{1.8.38}$$
$$w_{00}^{XU} - w_{00}^{XD} \geq \Delta \tag{1.8.39}$$

If $\theta_\alpha \geq \frac{C_h}{C_l}\theta_l$, the principal in the bad market will want to mimic. Rearrange $A.35$ and $A.36$:

$$(A.35) \rightarrow (A.35)'\ \ w_{10}^{XD} - w_{00}^{XD} \geq \frac{C_h}{\theta_\alpha}$$
$$(A.36) \rightarrow (A.36)'\ \ w_{10}^{XD} - w_{00}^{XD} \geq \frac{C_h - C_l - q\alpha\Delta}{\theta_\alpha - \theta_l}\frac{C_h - C_l - q\alpha\Delta}{\theta_\alpha - \theta_l} - \frac{C_h}{\theta_\alpha} \quad = \frac{\theta_l C_h - \theta_\alpha(C_l + q\alpha\Delta)}{(\theta_\alpha - \theta_l)\theta_\alpha}$$

As in the previous proof, one could easily verify that pooling at the conservative strategy is not an equilibrium. As a result, Constraint $A.37$ will be redundant.

To solve the above maximization problem, I get the following contract if $\theta_\alpha \geq \frac{C_h}{C_l}\theta_l$:

$$w_{00}^{XD} = 0, w_{10}^{XD} = \frac{C_h}{\theta_\alpha}, w_{01}^{XD} = \frac{C_l}{\theta_l}, w_{11}^{XD} = \frac{C_l}{\theta_l} + \frac{C_h}{\theta_\alpha};$$
$$w_{00}^{XU} = \Delta, w_{01}^{XU} = \frac{C_h}{\theta_h} + \Delta, w_{10}^{XU} = \frac{C_h}{\theta_\alpha} + \Delta, w_{11}^{XU} = \frac{C_h}{\theta_\alpha} + \frac{C_h}{\theta_h} + \Delta.$$

Step 4 Proof of contract under other information revelation structures – PSP dominated by PSS.

1. If the principal does not reveal the market condition in the first period, the agent will have to infer it from past performance following bad performance. Following good performance, the principal reveals the market condition, thus the agent does not have to infer it on her own.

$$\alpha' = Pr(\theta_1 = \theta_h | y_1 = 0) = \frac{\alpha(1-\theta_h)}{\alpha(1-\theta_h) + (1-\alpha)(1-\theta_l)} < \alpha \tag{1.8.40}$$

Define $\bar{\theta}_\alpha = \alpha' q\theta_h + (1-\alpha' q)\theta_l$. From Lemma 1, if the agent knows information, she takes the adaptive strategy which fits the external market. If the agent does not know the information, I first characterize the contract under which the agent takes the innovative strategy under pooling.

$$\max_{w\{..\}} \theta_h\{q(\theta_h(2-w_{11}^{XU}) + (1-\theta_h)(1-w_{10}^{XU})) + (1-q)(\theta_l(2-w_{11}^{XD}) + (1-\theta_l)(1-w_{10}^{XD}))\}$$
$$+(1-\theta_h)\{q(\theta_h(1-w_{01}^{XX}) + (1-\theta_h)(0-w_{00}^{XX})) + (1-q)(\theta_l(1-w_{01}^{XX}) + (1-\theta_l)(0-w_{00}^{XX}))\}$$

To save space, limited liability constraints are not listed. Constraint A.40-A.43 are the agent's incentive constraints, constraint A.44 is the agent's project choice constraint.

$$s.t.\ \bar{\theta}_\alpha(w_{11}^{XU} - w_{10}^{XU}) \geq C_h \tag{1.8.41}$$
$$\theta_l(w_{11}^{XD} - w_{10}^{XD}) \geq C_l \tag{1.8.42}$$
$$\bar{\theta}_\alpha(w_{01}^{XX} - w_{00}^{XX}) \geq C_h \tag{1.8.43}$$

$$\begin{aligned}
&\alpha\{\theta_h q(\theta_h(w_{11}^{XU} - w_{10}^{XU}) + w_{10}^{XU} - C_h) + (1-q)\theta_h(\theta_l(w_{11}^{XD} - w_{10}^{XD}) + w_{10}^{XD} - C_l) \\
&+(1-\theta_h)q(\theta_h(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h) + (1-\theta_h)(1-q)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h)\} \\
&(1-\alpha)\{\theta_l(\theta_l(w_{11}^{XD} - w_{10}^{XD}) + w_{10}^{XD} - C_l) + (1-\theta_l)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h)\} - C_h \\
&\geq \alpha q(\theta_h(w_{01}^{XX} - w_{10}^{XX}) + w_{00}^{XX} - C_h) + (1-\alpha q)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h)
\end{aligned} \tag{1.8.44}$$

$$\begin{aligned}
&\alpha\{\theta_h q(\theta_h(w_{11}^{XU} - w_{10}^{XU}) + w_{10}^{XU} - C_h) + (1-q)\theta_h(\theta_l(w_{11}^{XD} - w_{10}^{XD}) + w_{10}^{XD} - C_l) \\
&+(1-\theta_h)q(\theta_h(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h) + (1-\theta_h)(1-q)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h)\} \\
&(1-\alpha)\{\theta_l(\theta_l(w_{11}^{XD} - w_{10}^{XD}) + w_{10}^{XD} - C_l) + (1-\theta_l)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h)\} - C_h \\
&\geq \theta_l(\theta_l(w_{11}^{XD} - w_{10}^{XD}) + w_{10}^{XD} - C_l) + (1-\theta_l)(\theta_l(w_{01}^{XX} - w_{00}^{XX}) + w_{00}^{XX} - C_h) - C_l
\end{aligned} \tag{1.8.45}$$

Constraint A.46 is the principal's truth-telling constraint.

$$w_{10}^{XU} - w_{10}^{XD} \geq \Delta \tag{1.8.46}$$

Constraint A.47 is the mimicking constraint of the principal in the bad market.

$$\theta_l(\theta_l(2-w_{11}^{XD})+(1-\theta_l)(1-w_{10}^{XD}))+(1-\theta_l)(\theta_l(1-w_{01}^{XX})+(1-\theta_l)(0-w_{00}^{XX})) \quad (1.8.47)$$
$$\geq \theta_l(\theta_l(2-\frac{C_l}{\theta_l})+(1-\theta_l)(1-\frac{C_l}{\theta_l}))+(1-\theta_l)(\theta_l(1-\frac{C_l}{\theta_l})+(1-\theta_l)(0-0))$$

Rearrange A.43 and A.44, I get:

$$(A.43)\rightarrow(A.43'):$$
$$w_{10}^{XD}\geq\frac{1}{\theta_\alpha}\{C_h+\frac{C_h}{\bar{\theta}_\alpha}(q\alpha\theta_h+(1-q\alpha)\theta_l)-\alpha(1-\theta_h)\frac{\bar{\theta}}{\bar{\theta}_\alpha}C_h-(1-\alpha)(1-\theta_l)\frac{\theta_l}{\bar{\theta}_\alpha}C_h-q\alpha\theta_h\Delta\}$$
$$(A.44)\rightarrow(A.44'):$$
$$w_{10}^{XD}\geq\frac{1}{\theta_\alpha-\theta_l}\{(2-\theta_\alpha)C_h-(2-\theta_l)C_l+\alpha\frac{C_h}{\bar{\theta}_\alpha}((1-\theta_l)\theta_l-(1-\theta_h)\bar{\theta})-q\alpha\theta_h\Delta\}$$

Deduct the left hand side of $A.43'$ from that of $A.44'$, I get:

Define $f(q)=A.44'-A.43'$
$$=\frac{1}{\theta_\alpha-\theta_l}\{(2-\theta_\alpha)C_h-(2-\theta_l)C_l+\alpha\frac{C_h}{\bar{\theta}_\alpha}((1-\theta_l)\theta_l-(1-\theta_h)\bar{\theta})-q\alpha\theta_h\Delta\}$$
$$-\frac{1}{\theta_\alpha}\{C_h+\frac{C_h}{\bar{\theta}_\alpha}(q\alpha\theta_h+(1-q\alpha)\theta_l)-\alpha(1-\theta_h)\frac{\bar{\theta}}{\bar{\theta}_\alpha}C_h-(1-\alpha)(1-\theta_l)\frac{\theta_l}{\bar{\theta}_\alpha}C_h-q\alpha\theta_h\Delta\}$$

One could easily verify that $f(q)$ is decreasing in q. $\exists\ \bar{q}$ such that if $q>\bar{q}$, A.43 binds. I later verify that the contract if A.43 binds will be strictly dominated by PSS, which means the contract if A.44 binds will also be strictly dominated by PSS. This is because A.43 is no longer a binding constraint if $q\leq\bar{q}$ and the contract subject to only A.43 offers the principal a higher payoff than A.44.

To solve the above maximization problem, I get the following contract:

$$w_{10}^{XD}=\frac{1}{\theta_\alpha}\{C_h+\frac{C_h}{\bar{\theta}_\alpha}(q\alpha\theta_h+(1-q\alpha)\theta_l)-\alpha(1-\theta_h)\frac{\bar{\theta}}{\bar{\theta}_\alpha}C_h-(1-\alpha)(1-\theta_l)\frac{\theta_l}{\bar{\theta}_\alpha}C_h-q\alpha\theta_h\Delta\},$$
$$w_{11}^{XD}=\frac{C_l}{\theta_l}+w_{10}^{XD},w_{10}^{XU}=w_{10}^{XD}+\Delta,w_{11}^{XU}=\frac{C_h}{\theta_h}+w_{10}^{XD}+\Delta.$$

2. I then verify that the contract if A.43 binds will be strictly dominated by PSS.

$$\Delta\pi(q)=\mathbb{E}\left[\pi^{PSS}\right]-\mathbb{E}\left[\pi^{PSP}\right]$$
$$=\theta_h(w_{10}^{XD}-\frac{C_h}{\theta_\alpha})+(1-\theta_h)(1-q)\theta_l(\frac{C_h}{\theta_\alpha}-\frac{C_l}{\theta_l})+(1-\theta_h)q((\frac{C_h}{\bar{\theta}_\alpha}-\frac{C_h}{\theta_h})\theta_h-\Delta)$$
$$\Delta\pi(q=0)=\frac{\theta_h}{\theta_\alpha}(\theta_lC_h+\alpha(\theta_h-\theta_l)C_h)+(1-\theta_h)(C_h-C_l)$$
$$>0$$

Define the following $M(q)$ and $N(q)$:

$$\begin{aligned} M(q) =& (1-\theta_h)(1-q)\theta_l(\frac{C_h}{\theta_\alpha} - \frac{C_l}{\theta_l}) + (1-\theta_h)q((\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h})\theta_h - \Delta) \\ N(q) =& \theta_h w_{10}^{XD} \end{aligned}$$

Take the derivative of $M(q)$ and $N(q)$ w.r.t. q:

$$\begin{aligned} \frac{dM(q)}{dq} =& (\theta_h - \theta_l)C_h(\underbrace{\frac{1}{\bar{\theta}_\alpha} - \frac{1-\theta_h}{\theta_h}}_{>0} + \frac{\bar{\theta}}{\bar{\theta}_\alpha^2}) \\ \frac{dN(q)}{dq} =& \frac{\alpha C_h \theta_h}{\theta_\alpha \bar{\theta}_\alpha}(\theta_h - \theta_l)\{\theta_h - \frac{\alpha'}{\alpha\bar{\theta}_\alpha}(\theta_l^2 - \alpha\theta_l^2 + \alpha\theta_l\theta_h + q\alpha\theta_h(\theta_h - \theta_l))\} \end{aligned}$$

One could verify that:

$$(\theta_h - \theta_l)C_h\frac{\bar{\theta}}{\bar{\theta}_\alpha^2} - \frac{\alpha C_h\theta_h}{\theta_\alpha\bar{\theta}_\alpha}(\theta_h - \theta_l)\frac{\alpha'}{\alpha\bar{\theta}_\alpha}(\theta_l^2 - \alpha\theta_l^2 + \alpha\theta_l\theta_h + q\alpha\theta_h(\theta_h - \theta_l)) > 0$$

Given that $\mathbb{E}\left[\pi^{PSS}\right] - \mathbb{E}\left[\pi^{PSP}\right]$ is increasing in q and $\Delta\pi(q=0) = 0$, the principal in the good market will be better off in the PSS than in PSP. Following the proof here, one could verify that PPP will be dominated by PPS because the principal in the good market will not want to pool following bad performance

Step 5 Compare SSS to PPS, PSS and SSP.

1. Verify that the principal in the good market is better off in the PPS than in SSS if α is sufficiently small and q is sufficiently large.

$$\begin{aligned} \Delta\pi(\alpha, q) =& \mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{PPS}\right] \\ =& \theta_h(w_{10}^{XX} - \frac{C_h}{\theta_\alpha}) + \theta_h\{q(\theta_h(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h}) - 2\Delta) + (1-q)(\theta_h(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l}) - \Delta)\} \\ &+ (1-\theta_h)\{q(\theta_h(\frac{C_h}{\theta_h} - \frac{C_h}{\theta_h}) - \Delta) + (1-q)(\theta_h(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l}) - \Delta)\} \end{aligned}$$

$$\Delta\pi(\alpha = 0, q = 1) = 0 - (1+\theta_h)(C_h + \Delta) + \frac{\theta_h^2}{\theta_l}C_h < 0$$

As in the proof of PSP, one could easily show that $\mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{PPS}\right]$ is increasing in q and decreasing in α, the principal in the good market will be better off in the PPS than in SSS if α is sufficiently small and q is sufficiently large.

2. Verify that the principal in the good market is better off in the PSS than in SSS if α

is sufficiently large.

$$\begin{aligned}
\Delta\pi(\alpha) =&\mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{PSS}\right] \\
=&\theta_h(\frac{C_h}{\theta_\alpha} - \frac{C_h}{\theta_h}) + q(-\theta_h\Delta - (1-\theta_h)\Delta) + (1-q)(-\theta_l\Delta - (1-\theta_l)\Delta) \\
=&\frac{\theta_h}{\theta_\alpha}C_h - C_h - \Delta \\
\Delta\pi(\alpha=0) =& \frac{\theta_h}{\theta_l}C_h - C_h - C_l + \frac{\theta_l}{\theta_h}C_h \geq C_h - C_l > 0 \\
\Delta\pi(\alpha=1) =& -\Delta < 0
\end{aligned}$$

One could easily show that $\mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{PPS}\right]$ is decreasing in α. The principal in the good market is better off in the PSS than in SSS if α is sufficiently large.

3. Verify that the principal in the good market is better off in the SSP than in SSS if α is q is sufficiently small.

$$\begin{aligned}
\Delta\pi(q) =&\mathbb{E}\left[\pi^{SSS}\right] - \mathbb{E}\left[\pi^{SSP}\right] \\
=&(1-\theta_h)\{q(\theta_h(\frac{C_h}{\bar{\theta}} + w_{00}^{UX} - \frac{C_h}{\theta_h} - 2\Delta) + (1-\theta_h)(w_{00}^{UX} - 2\Delta)) \\
&+(1-q)(\theta_l(\frac{C_h}{\bar{\theta}} + w_{00}^{UX} - \frac{C_l}{\theta_l} - \Delta) + (1-\theta_l)(w_{00}^{UX} - \Delta))\} \\
&+\theta_h\{q(\theta_h(\frac{C_h}{\theta_h} + w_{00}^{UX} - q\Delta + \Delta - \frac{C_h}{\theta_h} - 2\Delta) + (1-\theta_h)(w_{00}^{UX} - q\Delta + \Delta - 2\Delta)) \\
&+(1-q)(\theta_l(\frac{C_l}{\theta_l} + w_{00}^{UX} - q\Delta - \frac{C_l}{\theta_l} - \Delta) + (1-\theta_l)(w_{00}^{UX} - q\Delta - \Delta))\} \\
=&C_h - q(C_h - C_l) - (1-\theta_h)\frac{\theta_l}{\bar{\theta}}C_h - C_l\theta_l - (1-\theta_l)q\Delta - \theta_h(1-q)(C_h - C_l)
\end{aligned}$$

Assume $G(q) = \bar{\theta}\Delta\pi(q)$. The value of $G(q)$ at $q=1$ and $q=0$ is:

$$\begin{aligned}
G(q=1) =&0 \\
G(q=0) =&-\theta_l(C_h - C_l)(\theta_h - \theta_l) < 0
\end{aligned}$$

I then check whether $G(q)$ is greater then zero or not.

$$\begin{aligned}
G(q) =&aq^2 + bq + c \\
a =&(\theta_h - \theta_l)^2(C_h - \frac{C_h}{\theta_h} - \Delta) < 0 \\
b =&(\theta_h - \theta_l)(C_h(1-\theta_h) + (\theta_h - \theta_l)C_l + \theta_l(C_h - C_l) - \frac{C_h}{\theta_h}\theta_l(1-\theta_l)) \\
G'(q=1) =&2a + b = (\theta_h - \theta_l)(C_h\theta_h - C_h - \Delta\theta_h - \frac{C_h}{\theta_h}\theta_l(1+\theta_l)) < 0
\end{aligned}$$

Based on the above inequalities, it is obvious that $\exists\ \bar{q}$, if $q < \bar{q}$ then $\mathbb{E}\left[\pi^{SSS}\right] -$

$\mathbb{E}\left[\pi^{SSP}\right] < 0$. The principal in the good market is better off in the SSP than in SSS if α is q is sufficiently small.

Step 6 If α is sufficiently small and q is sufficiently large, PSS is the equilibrium information structure.

$$\begin{aligned}
\Delta\pi(q,\alpha) =& \mathbb{E}\left[\pi^{PSS}\right] - \mathbb{E}\left[\pi^{PPS}\right] \\
=& \theta_h(w_{10}^{XX} - \frac{C_h}{\theta_\alpha}) + \theta_h q((\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h})\theta_h - \Delta) + \theta_h(1-q)\theta_l(\frac{C_h}{\theta_\alpha} - \frac{C_l}{\theta_l})
\end{aligned}$$

$$\begin{aligned}
\Delta\pi(q=1,\alpha=1) =& 0 \\
\Delta\pi(q=1,\alpha=0) >& 0
\end{aligned}$$

Following the proof in PSP, I show that $\mathbb{E}\left[\pi^{PSS}\right] - \mathbb{E}\left[\pi^{PPS}\right]$ is increasing in q: Define the following $M(q)$ and $N(q)$:

$$\begin{aligned}
M(q,\alpha) =& \theta_h q((\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h})\theta_h - \Delta) + \theta_h(1-q)\theta_l(\frac{C_h}{\theta_\alpha} - \frac{C_l}{\theta_l}) \\
N(q,\alpha) =& \theta_h w_{10}^{XX}
\end{aligned}$$

Take the derivative of $M(q)$ and $N(q)$ w.r.t. q:

$$\begin{aligned}
\frac{\partial M(q,\alpha)}{\partial q} =& \frac{\theta_h}{1-\theta_h}(\theta_h - \theta_l)C_h(\underbrace{\frac{1}{\bar{\theta}_\alpha} - \frac{1-\theta_h}{\theta_h}}_{>0} + \frac{\bar{\theta}}{\bar{\theta}_\alpha^2}) \\
\frac{\partial N(q,\alpha)}{\partial q} =& \frac{\theta_h}{\theta_\alpha}\{\theta_h\alpha\Delta + C_h\frac{\theta_h - \theta_l}{\bar{\theta}_\alpha^2}(\alpha\theta_h(\alpha-1)\theta_l + (1-\alpha)\theta_l^2\alpha')\}
\end{aligned}$$

One could verify that:

$$\frac{\theta_h}{1-\theta_h}(\theta_h - \theta_l)C_h(\frac{\bar{\theta}}{\bar{\theta}_\alpha^2}) - \frac{\theta_h}{\theta_\alpha}C_h\frac{\theta_h - \theta_l}{\bar{\theta}_\alpha^2}\alpha\theta_h(\alpha-1)\theta_l > 0$$

One could also prove that

$$\begin{aligned}
\frac{\partial M(q=1,\alpha=1)}{\partial \alpha} >& 0 \\
\frac{\partial M(q=1,\alpha=0)}{\partial \alpha} >& 0
\end{aligned}$$

Given that $\mathbb{E}\left[\pi^{PSS}\right] - \mathbb{E}\left[\pi^{PPS}\right]$ is increasing in q, $\Delta\pi(q=1,\alpha=1) = 0$ and $\Delta\pi(q=1,\alpha=0) > 0$, the principal in the good market will be better off in the PSS than in PPS if α is sufficiently small and q is sufficiently large. If Δ is sufficiently large, PSS is dominated by PPS, because:

$$\begin{aligned}\Delta\pi(q=0) =& -\frac{\theta_h}{\theta_\alpha}\Delta + \theta_h(C_h - C_l)\\ \leq& -\Delta + \theta_h(C_h - C_l)\end{aligned}$$

Step 7 If α and q are sufficiently small, adaptive innovation (SSP) is the equilibrium information structure. If α and q are sufficiently large, innovation inertia (PPS) is the equilibrium information structure.

$$\begin{aligned}\Delta\pi(\alpha,q) =&\mathbb{E}\left[\pi^{SSP}\right] - \mathbb{E}\left[\pi^{PPS}\right]\\ =&\theta_h(w_{10}^{XX} - \frac{C_h}{\theta_\alpha}) + \theta_h\{q(\theta_h(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h}) - (w_{00}^{UX} + \Delta - q\Delta)) + (1-q)(\theta_l(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l})\\ &-(w_{00}^{UX} - q\Delta))\} + (1-\theta_h)\{q(\theta_h(\frac{C_h}{\theta_h} - \frac{C_h}{\bar{\theta}}) + \Delta - w_{00}^{UX})\\ &+(1-q)(\theta_l(\frac{C_l}{\theta_l} - \frac{C_h}{\bar{\theta}}) - w_{00}^{UX})\}\end{aligned}$$

$$\begin{aligned}\Delta\pi(\alpha=0,q=1) =& -(1+\theta_h)(C_h + \frac{\theta_h^2}{\theta_l}C_h < 0\\ \Delta\pi(\alpha=1,q=1) =& -C_h - \Delta < 0\\ \Delta\pi(\alpha=0,q=0) =&\theta_h(\frac{C_h}{\theta_l} - \frac{C_h}{\theta_h}) - \Delta + (2\theta_h - \theta_l)(C_h - C_l) > 0\end{aligned}$$

In addition, I show the following derivatives:

$$\begin{aligned}\frac{\partial\Delta\pi(\alpha=0,q=0)}{\partial q} >&0 & \frac{\partial\Delta\pi(\alpha=1,q=1)}{\partial q} >&0\\ \frac{\partial\Delta\pi(\alpha=0,q=1)}{\partial q} >&0 & \frac{\partial\Delta\pi(\alpha=0,q=0)}{\partial q} >&0\\ \frac{\partial\Delta\pi(\alpha=1,q=0)}{\partial \alpha} >&0 & \frac{\partial\Delta\pi(\alpha=0,q=1)}{\partial \alpha} >&0\\ \frac{\partial\Delta\pi(\alpha=0,q=0)}{\partial \alpha} >&0 & \frac{\partial\Delta\pi(\alpha=1,q=1)}{\partial \alpha} >&0\end{aligned}$$

Therefore, there $\exists\ \bar{q},\bar{\alpha}$ if $q \geq \bar{q}$ and $\alpha \geq \bar{\alpha}$,$\mathbb{E}\left[\pi^{SSP}\right] - \mathbb{E}\left[\pi^{PPS}\right] \leq 0$. There $\exists\ \underline{q},\underline{\alpha}$ if $q \leq \underline{q}$ and $\alpha \leq \underline{\alpha}$,$\mathbb{E}\left[\pi^{SSP}\right] - \mathbb{E}\left[\pi^{PPS}\right] \geq 0$.

Q.E.D.

1.9 Appendix 2 – Intuitive Criterion Refinement

In this appendix, I implement Cho and Kreps (1987)'s Intuitive Criterion to refine the equilibria as a robustness check.

Proposition 7 ***Information Revelation Structures under the Refinement of Intuitive Criterion.*** *Only SSP survives Intuitive Criterion.*

Proof

1. I first verify that PSS does not satisfy the Intuitive Criterion. The principal in a good market gets $\pi_h^{pooling} = \theta_h(1-\frac{C_h}{\theta_\alpha}) + q((1-\frac{C_h}{\theta_h})\theta_h - \Delta) + (1-q)(1-\frac{C_l}{\theta_l})\theta_l$ under pooling. If she deviates and sends out a signal via paying ϵ, $\pi_h^{'} = \theta_h(1-\frac{C_h}{\theta_h}) - \epsilon + q((1-\frac{C_h}{\theta_h})\theta_h - \Delta) + (1-q)(1-\frac{C_l}{\theta_l})\theta_l$, $\epsilon \leq \pi_h^{'} - \pi_h^{pooling}$. The principal in a bad market gets $\pi_l^{pooling} = \theta_l(1-\frac{C_h}{\theta_\alpha}) + (1-\frac{C_l}{\theta_l})\theta_l$ under pooling. If she follows the deviation, $\pi_l^{'} = \theta_l(1-\frac{C_h}{\theta_h}) - \epsilon + (1-\frac{C_l}{\theta_l})\theta_l$. One could show that $\pi_h^{'} - \pi_h^{pooling} > \pi_l^{'} - \pi_l^{pooling}$.

$$\begin{aligned}
\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) &= \frac{\theta_h}{\theta_\alpha}C_h - C_h - (\frac{\theta_l}{\theta_\alpha}C_h - \frac{\theta_l}{\theta_h}C_h) \\
&= \frac{\theta_h - \theta_l}{\theta_\alpha}C_h - \frac{\theta_h - \theta_l}{\theta_h}C_h \\
&> 0
\end{aligned}$$

2. I now verify that SSP satisfies the Intuitive Criterion.

$$\begin{aligned}
\theta_l(1-w_{01}^{UU}) + (1-\theta_l)(0-w_{00}^{UU}) &\leq \theta_l(1-w_{01}^{UX}) + (1-\theta_l)(0-w_{00}^{UX}) \\
\Rightarrow -w_{00}^{UU} + \theta_l - \theta_l(w_{01}^{UU} - w_{00}^{UU}) &\leq -w_{00}^{UX} + \theta_l - \theta_l(w_{01}^{UX} - w_{00}^{UX}) \\
\Rightarrow w_{00}^{UU} &\geq w_{00}^{UX} + \Delta
\end{aligned}$$

If the principal could set w_{00}^{UU} to $w_{00}^{UX} + \Delta$, the principal in the bad market won't follow. Following performance $y_1 = 0$, the increase in profit if principal deviates from pooling is negative if the agent's belief is not affected by the deviation:

$$\begin{aligned}
\pi^D - \pi^{ND} &= \theta_h(1-w_{01}^{UU}) + (1-\theta_h)(0-w_{00}^{UU}) - [\theta_h(1-w_{01}^{UX}) + (1-\theta_h)(0-w_{00}^{UX})] \\
&= \theta_h - (C_h + w_{00}^{UU}) - [\theta_h - (C_h + w_{00}^{UX})] \\
&= (\theta_h - \theta_l)(\frac{1}{\bar{\theta}} - \frac{1}{\theta_h}) \\
&> 0
\end{aligned}$$

However, once the agent knows the principal's type, she will demand for a contract which at least yields the same utility given by w_{01}^{UX} and w_{00}^{UX}. This is different from the one period model. Since for the principal $\pi^D - \pi^{ND} > 0$, the agent, once knowing her own type, will not agree to the deviation. In the two period model, the contract in place interferes with the refinement of Intuitive Criterion.

3. I then verify that PPS satisfies the Intuitive Criterion if the principal in the good market deviates to SSS but not if she deviates to SPS.

   - Verify that if the agent is better off in the efficient pooling equilibrium following bad performance, then the Intuitive Criterion is satisfied.

     The argument here is the same as in SSP. Define W the old contract following bad performance under pooling and assume there exists a contract W' which makes the principal in the good market better off and prevents the principal in the bad market from mimicking. The agent who works for the principal in a good market learns the principal's type from W' and accepts W' only if it gives her higher payoffs than W. However, since the agent has learnt the market condition, it contradicts the assumption that W' makes the principal in the good market better off.

   - If α is close to zero and q is close to one, PPS could survive the Intuitive Criterion. If the principal in the good market deviates and sends out a signal via paying ϵ, she could get an increase in profit:

     $$\begin{aligned}\pi_h^{'} - \pi_h^{pooling} &= \theta_h(w_{10}^{XX} - \frac{C_h}{\theta_h}) - \epsilon + q(\theta_h(\frac{C_h}{\bar{\theta}_\alpha} \\ &- \frac{C_h}{\theta_h})\theta_h - \epsilon) + (1-q)\theta_h\theta_l(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l})\end{aligned}$$

     If the principal in the bad market follows, she could get an increase in profit:

     $$\pi_l^{'} - \pi_l^{pooling} = \theta_l(w_{10}^{XX} - \frac{C_h}{\theta_h}) - \epsilon + \theta_l^2(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l})$$

     To make $\pi_l^{'} - \pi_l^{pooling} = 0$, ϵ needs to be set at:

     $$\epsilon = \theta_l(w_{10}^{XX} - \frac{C_h}{\theta_h}) + \theta_l^2(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l})$$

     The principal in the good market chooses to deviate if the following is positive:

     $$\begin{aligned}\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) &= (\theta_h - \theta_l - q\theta_l)(w_{10}^{XX} - \frac{C_h}{\theta_h}) + q\theta_h^2(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\theta_h}) \\ &- (\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_l}{\theta_l})(q\theta_l^2(1+\theta_h) - (1-q)\theta_l(\theta_h - \theta_l))\end{aligned}$$

   - If α is close to zero and q is close to one, PPS could survive the Intuitive Criterion.

     $$\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) = (\theta_h - \theta_l)(\theta_h^2 + \theta_h - 2\theta_l)\frac{C_h}{\theta_h\theta_l} - (C_h - C_l)\theta_l(\theta_h + 1)$$

The principal will thus deviate as long as:

$$\frac{C_h - C_l}{C_h} = \frac{(\theta_h - \theta_l)(\theta_h^2 + \theta_h - 2\theta_l)}{\theta_l^2 \theta_h(\theta_h + 1)}$$

- If α is close to one and q is close to one, PPS does not survive the Intuitive Criterion.

$$\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) = (\theta_h - \theta_l + \theta_h + \theta_l)\Delta > 0$$

- If α is close to one and q is close to zero, PPS does not survive the Intuitive Criterion.

$$\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) = (C_h - C_l)(\theta_h - \theta_l) > 0$$

4. I then verify that PPS does not satisfy the Intuitive Criterion if the principal in the good market deviates to SPS. If the principal in the good market deviates and sends a signal via paying ϵ, she could get an increase in profit:

$$\begin{aligned}\pi_h^{'} - \pi_h^{pooling} &= \theta_h(w_{10}^{XX} - \frac{C_h}{\theta_h}) + \theta_h^2(\frac{C_h}{\bar{\theta}_\alpha} \\ &- (\frac{C_h}{\bar{\theta}} + w_{00}^{UD} + q\epsilon)) + \theta_h(1-\theta_h)(0 - (w_{00}^{UD} + q\epsilon)) + (1-\theta_h)(-w_{00}^{UD})\end{aligned}$$

If the principal in the bad market follows, she could get an increase in profit:

$$\begin{aligned}\pi_l^{'} - \pi_l^{pooling} &= \theta_l(w_{10}^{XX} - \frac{C_h}{\theta_h}) + \theta_l^2(\frac{C_h}{\bar{\theta}_\alpha} \\ &- (\frac{C_h}{\bar{\theta}} + w_{00}^{UD} + q\epsilon)) + \theta_l(1-\theta_l)(0 - (w_{00}^{UD} + q\epsilon)) + (1-\theta_l)(-w_{00}^{UD})\end{aligned}$$

To make $\pi_l^{'} - \pi_l^{pooling} = 0$, ϵ needs to be set at:

$$\epsilon = \theta_l(w_{10}^{XX} - \frac{C_h}{\theta_h}) + \theta_l^2(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\bar{\theta}})$$

The principal in the good market chooses to deviate if the following is positive:

$$\begin{aligned}\pi_h^{'} - \pi_h^{pooling} - (\pi_l^{'} - \pi_l^{pooling}) &= (w_{10}^{XX} - \frac{C_h}{\theta_h})(\theta_h - \theta_l)(1 - q\theta_l) + (\theta_h - \theta_l)(\frac{C_h}{\bar{\theta}_\alpha} - \frac{C_h}{\bar{\theta}}) \\ &> 0\end{aligned}$$

Q.E.D.

Chapter 2

Asymmetric Contractual Revision and Compensation Structure

2.1 Introduction

We rarely observe executive compensation cuts. This is puzzling, as compensation may go up or down in an environment where a firm's productive efficiency or its manager's matching quality with the firm changes over time. However, using CEO compensation data, Shue and Townsend (2014) show that not only do salary and bonus exhibit downward rigidity but option and stock grants also do. Moreover, Huang, Lü, and Xu (2015) find that many employment contracts of S&P 500 CEOs allow for discretionary compensation rewards by specifying a minimum level of salary and incentive pay and even explicitly prevent compensation cuts.[1,2]

Morale-based theory (Bewley, 2007) and pay-for-luck view (Bertrand and Mullainathan, 2001) offer a behavioral rational for asymmetric compensation adjustments. Information-based theories of managerial compensation also attempt to explain the empirical findings based on the assumption that managers have private information about either their own skills or the actions they take. For instance, increasing explicit incentives are to provide insurance for a risk-averse agent in moral hazard models with a risk-averse agent (Lambert, 1983; Sannikov, 2008) and to substitute declining implicit incentives in career concern models (Harris and Holmstrom, 1982; Gibbons and Murphy, 1992).

[1] Huang et al. (2015) show that 58.2% of S&P 500 CEOs' contracts explicitly allow for salary increases, 57.1% explicitly rule out salary cuts.

[2] For instance, in 2006, Mr. Ludwig, the CEO of Becton Dickinson &Co., got an increase in annual incentive payment target for fiscal year 2007 from 110% to 115% of base salary. In an another example, Mr. Flax signed a contract in 2005 with California Pizza Kitchen Inc., which offers him at least 30% of base salary for attainment of the performance based threshold amount to a maximum of 200% for exceptional performance.

In contrast, the general insight to be drawn from this paper is that compensation downward rigidity is an inherent feature in an economic setting in which the principal needs to correctly inform the agent of changing productive efficiency. This paper develops a two-period model in which an informed principal has private information about the changing productive efficiency and offers a compensation contract to the agent. The agent repeatedly makes a private effort to produce an output over two periods. The firm's productive efficiency is complementary with the agent's effort. In addition to the traditional role of the contract in providing incentives, it serves another role of credibly communicating the principal's private information to the agent. The model explores how managerial contracts deal with moral hazard and signalling problems at the same time, and more importantly, how an interaction of these two problems affect the dynamics of compensation structure.

The starting point of this paper is to observe that a combination of two characteristics seems especially important for senior managers. A McKinsey report (Casal and Caspar, 2014) indicates that board directors may have better knowledge than managers about industrial trends.[3] A recent survey conducted by Larcker et al. (2014) suggests that board directors may know managers' abilities very well.[4] Another essential characteristic of managerial compensation plans observed in many organizations is that they are not hard-wired, but instead leave a considerable amount of discretion to the board Huang et al. (2015). Unlike non-management employees who can be rewarded through rank-order tournaments, tools that can be employed to incentivise top executives are more limited. The principal's private information is therefore often just as important a friction in organizations as the agent's private information, because it affects how the principal executes her discretionary power in offering compensation contracts.

We show that the agent works harder if the principal reveals good private information via contract, as she realizes her labor productivity is higher than she initially perceived it to be. The principal with high productive efficiency in the first period commits to a compensation schedule, which consists of two basic compensation units. If the private information continues to be good, the principal chooses the compensation unit with a long-term performance-based pay and increasing salary. If it deteriorates, the principal instead offers the basic unit with increasing short-term performance-based pay and salary. The principal with low productive efficiency in the first period does not

[3]According to McKinsey Quarterly in February 2014, the right directors are knowledgeable about their roles and able to commit sufficient time to analyzing what drives value. They also actively engage in strategic planning and look for potential development areas.

[4]In the survey by Larcker et al. (2014), over half (55.1%) of directors report understanding the strengths and weaknesses of senior executives "extremely well" or "very well". A third (33.5%) understand these strengths and weaknesses "moderately well", and only the remainder (11.4%) understand them "slightly well" or "not at all well".

commit to such a costly downward-rigid contract but a long-term performance-based pay. The equilibrium contract, therefore, implies asymmetric contractual adjustments over time, only upward but not downward revision of each compensation component.

In my model, the effect of the principal's private information on the firm's output is twofold. First, private information regarding the productive efficiency directly enters into the firm's production function – higher productive efficiency leads to a higher output. Second, an agent who receives a good signal works harder, indirectly leading to a higher output. To prevent the principal with low productive efficiency from mimicking, the principal with high productive efficiency allocates the profit coming from the second indirect channel to the agent.

I begin by analyzing the one-period mode. I show that if the output function exhibits zero log-supermodularity in productive efficiency and effort, the principal fully relies on the salary (or the fixed component) to communicate her private information, while the bonus (the variable component) is paid at a level as if there was no information asymmetry. In this case, private information does not cause the bonus to be type-dependent, and the role of the bonus is only to provide incentives. The principal will not want to use a higher bonus to signal, as sharing profit with the agent is more costly than just offering a salary. If the condition is not satisfied, the firm either uses more profit sharing or under-effort-provision to communicate her private information. The bonus will differ from the level under symmetric information. In other words, the bonus is not only a reward for effort but also a means of signalling, regardless of the eventual performance.

For the two-period model, I choose a specific production technology with zero log-supermodularity in productive efficiency and effort. Such a technology allows me to assign the signalling role to the salary and the incentive role to the bonus in a one-period model and ensures that, in a two-period setting, the change in bonus over time is not caused by a change in the productive efficiency. The result that bonus in the one-period model is not affected by the content of the private information does not hold in a two-period setting.

I first consider the case in which commitment is disenabled as a benchmark case. If the principal cannot commit to long-term contracts, the agent knows that the principal will make a new take-it-or-leave-it offer when new information arrives. Anticipating this, the agent will not agree to an arrangement which promises a higher bonus in future but a lower salary today. The equilibrium contract is stationary in the sense that the second period contract does not depend on the private information of the first period.

I then allow commitment. But I first forbid long-term performance-based pay to be offered so that I could compare the two-period equilibrium contract directly to the one-period contract without the complication of equity payment. In equilibrium, the principal commits to a compensation schedule which consists of two basic compensation units. The principal pays more salary or bonus in the second period in exchange for less paid to the agent in the first period as a way of providing the first period signal. If the productive efficiency of the firm continues to be high, the principal will provide a higher salary rather than a higher bonus, as the salary is a less costly way of signalling than sharing profit. If the productive efficiency declines, the principal chooses a higher bonus based on the second period performance measure as a way of providing signals in the first period. Such an arrangement achieves greater efficiency by giving the agent more profit sharing opportunities and inducing greater effort.

Now consider the case in which the principal can offer long-term performance-based pay. In equilibrium, the principal commits to a compensation schedule which again consists of two basic compensation units. However, if the productive efficiency of the firm continues to be high, the principal provides the basic unit with a long-term performance-based pay and increasing salary, both of which are higher than the levels under symmetric information. If firm's productive efficiency declines, the principal chooses the basic unit with a performance-based pay heavily loaded on the second-period performance measure as a way of providing the signal for the first period. Such an arrangement also achieves greater efficiency by giving the agent more profit sharing opportunities and reduces the rent extracted by the agent due to the agent's limited liability constraint.

In brief, the principal allocates the signalling cost over time and even uses pay-performance-sensitivity to signal her private information in a dynamic setting, which differs from the one-period model in which only salary provides signal. As a result, the equilibrium contract exhibits downward rigidity in a whole spectrum of compensation structure.

My model generates a rich set of empirical predictions. First, discretionary rewards and long-term contracts are more likely to be offered to a high-skilled worker with greater likelihood of better performance in future. Moreover, long-term incentive pay sends out a stronger signal about the agent's skills than a discretionary bonus award. Second, back-loaded long-term contracts with high pay-performance sensitivities are more likely to be observed in positions with a high variation in skills, which can be R&D-oriented jobs and require great leadership. In a similar vein, start-up firms and rising industries are also more willing to offer such contracts. Third, substituting a salary raise with a bonus raise is more likely to happen if the productive efficiency or matching quality deteriorates. Lastly, if an agent has a strong bargaining power, for instance, high outside options, firms are more likely to offer a contract which gives the agent more profit sharing opportunities.

This paper also sheds light on the attempts of recent regulations to curb managerial bonuses.[5] My paper offers a new angle to evaluate possible effects of this policy on executive compensation, mostly applicable to firms which need to hire new executives or rely on subjective evaluation in providing compensation rewards. In the case of under-effort-provision in the one-period model, I show a surprising result that the principal will not cut the bonus but rather increase it under the current rule of bonus caps. The overall efficiency is improved, as the principal offers more profit sharing with the agent and thus induces the agent to make more efforts. However, this policy also creates its own distortion that the principal may not want to reveal her private information to the agent any more due to an increase in the signalling cost.

The extension considers the case in which the agent possesses transferable skills or has type-dependent outside options. Over the past three decades, the relative importance of general versus specific managerial skills has changed dramatically (Murphy and Zábojník, 2007). Technological innovation helps executives acquire firm-specific information about a company's operation more easily and requires managers to work under a more diverse environment due to an expansion of the product market. The principal chooses to offer a higher bonus instead of a even higher salary to meet the participation constraint of an agent with transferable skills and to retain her. However, if the agent's outside option value is too high, the principal would rather not to signal and pool with the principal who hires a low-skilled agent. I argue that a shift in the relative importance of general skills versus firm-specific skills leads to higher pay-performance sensitivity, because the principal needs to provide strong signal to motivate and retain the executive whose skills best match the firm's new operating environment.

Related Literature. A large body of theoretical research has investigated magnitudes and determinants of pay-performance sensitivity of executive compensation. One strand of literature is based on moral hazard models (Baker, Gibbons, and Murphy, 1994; Bull, 1987; MacLeod and Malcomson, 1989). The other strand of literature MacLeod (2003); Levin (2003); Fuchs (2015); Zábojník (2014) incorporates hidden information into their models in which only the principal observes the performance measure.

The first strand of the literature studies optimal contracts under moral hazard games in which both the principal and the agent could observe non-verifiable perfor-

[5] EU regulators have decided to institute bonus caps in order to rein in executive compensation. It came into effect at the beginning of 2014. Under this policy, certain bankers can only be paid a bonus equal to their annual salary or twice as much if their firm gets approval from shareholders.

mance measures. It focuses primarily on how repeated interaction between the principal and the agent allows firms to overcome the reneging problem wherein supervisors are tempted to underpay workers in order to save on labour costs (Baker et al., 1994; Bull, 1987; MacLeod and Malcomson, 1989). Performance incentives can then be sustained only through the threat of terminating cooperation in future periods if the principal were to behave opportunistically in any given period and deny the agent the bonus promised under the implicit agreement. Baker et al. (1994) further introduce a verifiable performance measure and study the interaction between implicit bonus, based on a non-verifiable performance measure, and explicit bonus, contractible upon a verifiable performance measure. They show that depending on the value of the fall-back position after reneging on an implicit contract, implicit bonus and explicit bonus can be substitutes or complements.

This strand of literature also studies long-term contracts under moral hazard models with a risk-averse agent (Lambert, 1983). Increasing explicit incentives are to provide insurance to the agent and reduce incentive cost. Relatedly, Acharya, John, and Sundaram (2000) study option resetting and show that companies do not penalize over-incentivized employees by resetting their options downwards following good performance. Combining moral hazard and learning, career concern models (Harris and Holmstrom, 1982; Gibbons and Murphy, 1992) in which both parties need to learn the agent's skills show that increasing explicit incentives are to substitute declining implicit incentives.

The second strand of literature (Fuchs, 2015; Levin, 2003; MacLeod, 2003; Zábojník, 2014) studies an optimal contracting problem with a privately informed principal. In their short-term contracting arrangements based only on subjective performance indicators, the principal faces a more severe problem in making any incentive provisions credible. MacLeod (2003) has generalized the logic of repeated game models by demonstrating that subjective schemes can be feasible even without infinite interaction if workers can punish a deviation from the implicit contract by imposing on the employer some type of socially wasteful cost. This model was further developed by Fuchs (2007), who extended it to a more dynamic environment, and by Ederhof, Rajan, and Reichelstein (2011), who introduced objective measures of performance.

In contrast with other papers in the second strand of literature, Fuchs (2015) and Zábojník (2014) consider the private information regarding the the production technology or the agent's skills. Fuchs (2015) studies a contract consisting of only fixed compensation and leaves aside the moral hazard problem. The paper shows that discretionary salary can be used as a signalling device. Zábojník (2014) further introduces moral hazard problem. In particular, for the subjective evaluations to provide any incentives, the second period performance-based pay must necessarily be distorted away from what would be optimal in the absence of subjective evaluation.

This paper belongs to the second branch of literature in the sense that my model also studies an optimal contracting problem in which the principal has private information regarding the production technology. It departs from the existing literature in two noticeable ways. First, I focus on the dual roles of the variable pay rather than the fixed pay in providing signals and incentives. Second, I study the role of long-term contracts in alleviating information asymmetry by incorporating a dynamic evolvement of private information into the model.

My paper is also related to the literature which studies general managerial skills (Dutta, 2008; Murphy and Zábojník, 2007). Dutta (2008) studies how general skills or the value of the agent's outside option affects the pay performance sensitivity if the agent has private information. Without any hidden information, Murphy and Zábojník (2007) study how managerial skills affect a firm's promotion decisions under a general equilibrium framework. They show that as firm-specific capital becomes relatively less important, the benefit of better matching increases relative to the cost of (lost) specific capital, and the prevalence of outside hires will increase.

Finally, my paper contributes to wage rigidity literature. Ederer (2010) shows that firms assess ordinary workers and provide feedback to them toward achieving a specific goal, for example, a promotion. Although firms could use other tools, such as promotions, to provide feedback to ordinary workers, wage still constitutes a large part of their rewards. My paper suggests a reason why worker wage is rigid in addition to efficiency wage theory (Akerlof and Yellen, 1990; Shapiro and Stiglitz, 1984; Weiss, 1980) and implicit contract theory (Stiglitz, 1986). Information asymmetry in my model creates such a friction that firms use long-term contracts to provide feedback and thus willingly constrain the flexibility in adjusting their workers' wage in future.

The structure of the paper is as follows. Section 2 describes the model. The baseline model is presented in section 3. Section 4 considers the case in which the principal's private information dynamically evolves. Section 5 discusses the recent regulation of bonus caps. Section 6 considers transferable managerial skills and disclosure policies. The last section concludes.

2.2 The Model

2.2.1 Subjective Measure

The model consists of two periods, period 1 and 2 ($t = 1, 2$). At the beginning of the employment relationship, the agent does not know the firm's productive efficiency θ precisely but only the public information that θ can take two values $\theta_i \in \{\theta_l, \theta_h\}$, with $0 < \theta_l < \theta_h \leq 1$ and $i \in \{i, h\}$. The probabilities of these types are α and $1 - \alpha$ respectively. The principal, however, has a private signal η_1 for the firm's productive efficiency, $\eta_1 \in \{\theta_l, \theta_h\}$[6].

In the second period, a principal of type θ_l in the first period continues to be θ_l, while a principal of type θ_h may either remain to be type θ_h with probability q or become type θ_l with probability $1 - q$. This assumption means that the firm's productivity may decline over time, which can be caused by a change in the product market or a sudden discontinuation of an investment project. Only the principal receives a private signal $\eta_2 \in \{\theta_h, \theta_l\}$ about the exact type of the productive efficiency at the beginning of the second period. To ease the exposition, signals are assumed to be perfect. I also provide an analysis of continuous type in the Appendix 2.

The principal will decide whether to convey her private information to the agent at date 0[7]. Due to the non-observability of the signal to the agent, it is impossible to write a contract contingent on it.

As argued in the introduction, such a signal can be interpreted in many ways. By virtue of monitoring many inputs, a supervisor gains superior information about the worker's productive talents (Alchian and Demsetz, 1972). It can be the principal's evaluation of the agent's leadership[8]. Or the principal might have superior information regarding the matching quality between the firm and the agent (Fuchs, 2015) . It can also be more general information beyond the agent's skills. For instance, the may have more accurate estimation of its own total factor productivity and have better

[6] One may prefer to interpret the private information as the matching quality between the two parties. Under this interpretation, one needs to assume that the principal will hire the agent even when she is known as a low type. This assumption can be motivated in two ways. First, the searching cost of finding a replacement is extremely high and the firm needs to find a stop-gap agent. Second, the principal could also make positive profit by hiring a low skilled agent.

[7] Unlike the agent in MacLeod (2003) who also receives a different private signal, the agent in this model does not receive any private signal. By avoiding opinion clashes between the principal and the agent, I could thus focus on the trade-off between signalling and moral hazard problems. Otherwise, there might be evaluation inflation, as the agent will impose costs upon the principal whenever there is a disagreement regarding performance.

[8] This type of behavior is related to the study of leadership by Hermalin (1998): He argues that leaders have superior information and a temptation to mislead their followers. In order for the leader to credibly signal his private information he must then either sacrifice or set an example (a costly action).

knowledge about industrial trends and macro conditions[9].

2.2.2 Production Technology

The principal supervises the agent over two periods. The agent's output in period t is denoted as $y_t \in \{0, 1\}$. It is verifiable to both parties. The probability of achieving an output y_t equal to 1 is $p = P(\theta, e_t)$. e_t is the effort that the agent makes in period t, and $e_t \in [0, 1]$. Those imperfect measure aggregates the agent's individual effort in a manner that differs from her contribution to the firm value. To ease the exposition, it will be assumed that y_1 and y_2 only become observable at the end of the end of period 2.[10]

The production technology has the following features: (1) $P(\theta, 0) = 0$; (2)$P(\theta, e_t)$ is differentiable in θ and e_t, $\frac{\partial P}{\partial e_t} > 0$ and $\frac{\partial P}{\partial \theta} > 0$; (3) $\frac{\partial^2 P}{\partial \theta \partial e_t} > 0$; (4) $P(\theta_h, 1) \leq 1$. Feature (1) means zero effort leads to zero output. Feature (2) says the probability of achieving a high output increases with skills and the amount efforts. Feature (3) suggests that super-modularity exists between efforts and skills. In other words, the marginal productivity per unit of agent's effort increases with the firm's productive efficiency. The last assumption ensures that the maximum value of the probability of achieving a high output is no greater than 1.

2.2.3 Contract

I characterize long-term contracts in equilibrium. To be more precise, the principal offers a compensation schedule $\mathcal{M}$ at date 0. The schedule is a subset of $\mathbb{R}^4_+$, $\mathcal{M} \subseteq \mathbb{R}^4_+$. $\mathbb{R}^4_+$ consists of four coordinates or four possible combinations of y_1 and y_2: w_{00} if $(y_1, y_2) = (0, 0)$, w_{10} if $(y_1, y_2) = (1, 0)$, w_{01} if $(y_1, y_2) = (0, 1)$, and w_{11} if $(y_1, y_2) = (1, 1)$. Any contingent payment is greater than or equal to zero due the limited liability constraint. For ease of exposition, define $w_{y_1 y_2} = \{w_{00}, w_{10}, w_{01}, w_{11}\}$ as the basic unit of a compensation schedule. $w_{y_1 y_2}$ is an element of $\mathbb{R}^4_+$, $w_{y_1 y_2} \in \mathbb{R}^4_+$.

While performance is contractible, the principal's private information is not. Based on the private information θ, the principal offers the agent a compensation schedule $\mathcal{M}$ at date 0. The compensation schedule $\mathcal{M}$ may contain more than one basic unit, and the principal commits to choosing one unit only from this schedule in future. After observing the private information θ_2 in the second period, the principal, at her sole

[9] According to a recent McKinsey report Casal and Caspar (2014), "Boards need to look further out than anyone else in the company," commented the chairman of a leading energy company. "There are times when CEOs are the last ones to see changes coming."

[10] This assumption is to shut down renegotiation caused by a change in continuation value after the principal observes the performance.

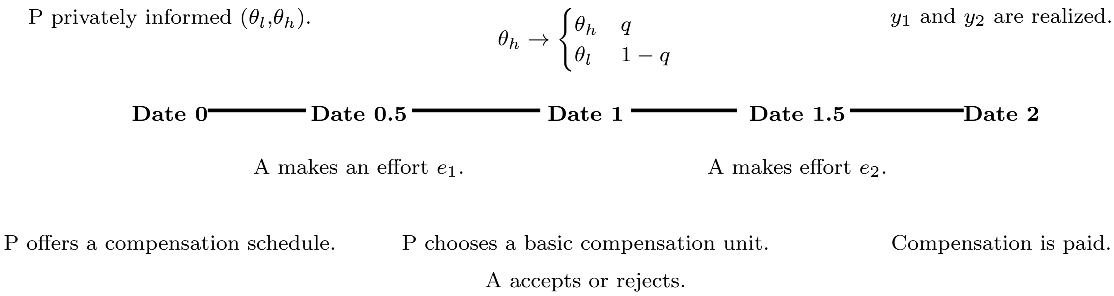


Figure 2.1: The Timeline of a Two-Period Model

Note: A represents the agent; P represents the principal.

discretion, decides on the the exact contract form $w_{y_1y_2}$, and $w_{y_1y_2} \subseteq \mathcal{M}$.

Because the setting involves a signalling problem, the payment scheme will be fully revealing of the principal's private information under a separating Perfect Bayesian Equilibrium (PBE). This setting of PBE might have multiple equilibria. For the purpose of this analysis, the most interesting among them is a separating PBE. I apply Cho and Kreps (1987)'s Intuitive Criterion to refine equilibria.

2.2.4 Preferences

The principal and the agent are risk neutral. For simplicity, I assume that the discount rate for future payoffs is zero. The principal's goal is to maximize the firm's profit. The agent's effort cost function is $\psi(e_t)$ in period $t = 1, 2$. It is twice differentiable in e_t. Assume that $\psi(0) = 0$, and $\psi'(e) > 0$. The agent maximizes the compensation after deducting the disutility from effort. Further assume that $\frac{\partial^2 P}{\partial^2 e} - \psi''(e) < 0$. This assumption ensures that the second order condition of the agent's utility is satisfied for any form of compensation scheme. The agent has zero initial wealth and is protected by limited liability. The agent's reservation utility is assumed to be zero over the entire time horizon, for all θ.[11]

2.2.5 Timing

At date 0, the principal has a private signal regarding the firm's productive efficiency θ and offers a contract $w\{..\}$ to the agent.[12] The agent could then choose to leave or stay. If she accepts the contract, she then makes an effort e_1 at an interim date 0.5. If

[11] In extensions, I analyze the case in which the agent's reservation utility is type dependent.

[12] This is a signalling setting, where the informed party offers a contract. If the uninformed party offers a contract, the agent will set $b_1(1) = 1$ and zero salary, the principal will then be indifferent between lying and truthfully reporting. As such, compensation design is irrelevant.

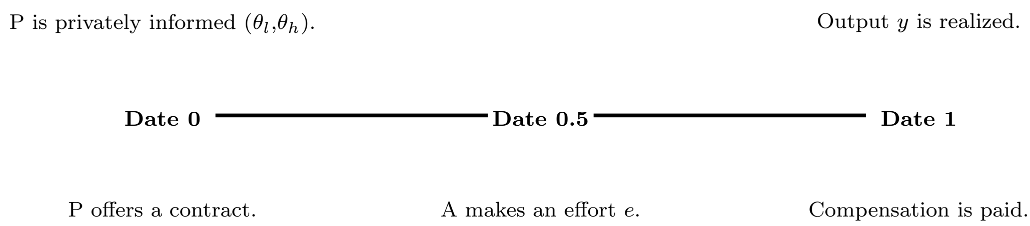


Figure 2.2: The Timeline of a One-Period Model

Note: A represents the agent; P represents the principal.

she leaves, she gets reservation utility equal to zero. At the beginning of period 2, after receiving a new signal, the principal decides whether to offer a new contract. However, two parties will abide by the old contract if the agent refuses the new offer. The agent makes further effort $e_2 \in [0, 1]$. At the end of period 2, y_1 and y_2 are observed and compensation is paid. Figure 2.1 is the timeline of the two-period model.

2.3 A One-Period Model

Before I proceed to characterize the optimal contract under asymmetric information, it is helpful to start with an analysis of a benchmark case in which the agent receives the the same signal as the principal and thus knows her own type. I then solve for the optimal contract under asymmetric information and compare it with the benchmark contract. One could verify that the compensation can be implemented through a fixed salary f_1 and a bonus $b_1(y_1)$ which is contingent on the output measure y_1.[13] Figure 2.2 represents the timeline of a one-period model.

2.3.1 Symmetric Information

In this benchmark case, the principal and the agent are both informed of the firm's productivity θ, thus there is no need for the principal to use compensation as a means of providing signal to the agent. Consequently, paying f_1 is not necessary, as it has neither incentive value nor signalling value.

Because the agent is protected by limited liability, she cannot be punished when the performance is bad. The principal thus chooses to pay the minimum to the agent in case of a low output, that is, $b_1(0) = 0$ if $y = 0$. Define $b_1(1) = b_1$ if $y_1 = 1$. In the benchmark case, both parties receive the signal η. As a result, superscript b_1 as b_1^i ($i \in \{l, h\}$) for the low and high type respectively.

[13]For proof, please refer to Lemma 6.

I first analyze the agent's problem. Given b_1, an agent of type θ_i chooses effort e to maximize her utility:

$$\max_{e} P(\theta_i, e) b_1^i - \psi(e)$$

From the first order condition, the optimal level of effort as a function of θ and b_1 is $e^* = e(\theta, b_1)$. The effort level depends on the productivity θ and the bonus b_1. Given the optimal effort level of the agent, the maximization problem P_0 of a principal who hires an agent of type θ_i is as follows:

$$\begin{aligned} &\max_{b_1} P(\theta_i, e)(1 - b_1^i) && \\ s.t. \quad & e^* = e(\theta_i, b_1) && IC_a \\ & P(\theta_i, e) b_1 - \psi(e) \geq 0 && IR_a \end{aligned}$$

Constraint IC_a is the agent's incentive constraint obtained from her own maximization program. Constraint IR_a is the participation constraint of the agent. The limited liability constraint if the output is low ($y = 0$) will be binding. The objective function of Program P_0 already takes this into account.

The optimal level of bonus under symmetric information is denoted as b_1^s. It is given by the first order condition of the principal's maximization program:

$$\frac{\partial P(\theta, e)}{\partial e} \frac{\partial e(\theta, b_1^s)}{\partial b_1^s}(1 - b_1^s) = P(\theta, e) \tag{2.3.1}$$

The LHS of Equation 2.3.1 measures the marginal benefit that results from a unit increase in b_1 through an increase in the agent's effort, deducting the compensation. The RHS of Equation 2.3.1 represents the marginal cost, the direct effect of an increase in b_1 on the marginal cost. As argued at the beginning of this section, salary is not necessary under symmetric information. This means that the bonus solely serves the role of incentivising the agent. The following lemma characterizes the conditions under which the incentive effect becomes weaker or stronger as the type varies.

Proposition 8 *Objective incentive compensation under symmetric information:*

- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} = 0$, *then* $b_1^{s,h} = b_1^{s,l}$;
- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} > 0$, *then* $b_1^{s,h} > b_1^{s,l}$;
- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} < 0$, *then* $b_1^{s,h} < b_1^{s,l}$.

This proposition shows that linear super-modularity between type and effort is not sufficient to give rise to an increasing bonus with respect to the type θ. To suffice a positive relationship, stronger super-modularity, more specifically, positive log super-modularity is required. Under information symmetry, the variation of the bonus with respect to θ is derived from the increasing absolute value of Marginal Rate of Substitution (MRS) between effort and compensation as θ improves. If the log super-modularity of the output function between the productivity and effort is zero, the MRS between effort and compensation does not vary with the type. If the log super-modularity is positive, the absolute value of MRS between effort and compensation increases with the agent's ability. The principal with higher productive efficiency offers higher bonus to an agent. Proposition 8 will have other important implications for later analysis.

2.3.2 Asymmetric Information with Informative Bonus

This section characterizes the one-period optimal contract under asymmetric information and the full spectrum of the contract is set free to provide feedback. The model is formally a signalling game and as such can have multiple equilibria. For the purpose of this analysis, the most interesting among them is a separating Perfect Bayesian Equilibrium (PBE).[14] The separating PBE in this section is defined as follows:

Definition A separating equilibrium is a Perfect Bayesian Equilibrium which satisfies:

1. The principal offers a contract $[f_1^i, b_1^i(y)]$ that maximizes the firm's profit.
2. The agent's belief of the principal's evaluation is $\beta(\eta = \theta_i | f_1^i, b_1^i(y)) = 1$.
3. Given the contract $[f_1^i, b_1^i(y)]$ and the belief, the agent chooses an effort level which maximizes her own utility.

I first analyze the agent's problem. Given b_1, the agent chooses effort e to maximize her utility:

$$\max_e P(m, e)b_1 - \psi(e)$$

From the first order condition, the optimal level of effort as a function of m and b_1 is $e^* = e(m, b_1)$. The effort level depends on the message m and the bonus b_1. Unlike b_1 in the previous section, the bonus b_1 under asymmetric information depends on the message the principal wants to convey to the agent.

[14]Pooling equilibrium does not survive the Intuitive Criterion.

Given the optimal effort level of the agent, the principal's problem ($P1$) with a type θ_h is as follows:

$$
\begin{aligned}
&\max_{f_1^h, b_1^h} P(\theta_h, e)(1-b_1^h) - f_1^h \\
s.t.\quad & e^* = e(\theta_h, b_1^h) && IC_a \\
& P(\theta_h, e^*)b_1^h - \psi(e^*) + f_1^h \geq 0 && IR_a \\
& P(\theta_h, e(\theta_h, b_1^h))(1-b_1^h) - f_1^h \geq P(\theta_h, e(\theta_l, b_1^l))(1-b_1^l) - f_1^l \quad for\ \theta_h && IC_p \\
& P(\theta_l, e(\theta_l, b_1^l))(1-b_1^l) - f_1^l \geq P(\theta_l, e(\theta_h, b_1^h))(1-b_1^h) - f_1^h \quad for\ \theta_l && IC_p
\end{aligned}
$$

Constraint IC_a is the agent's incentive constraint obtained from her maximization program. Constraint IC_p for θ_h and θ_l are the principal's truth-telling constraints.[15] If they are satisfied, the principal will truthfully report the private signal. If Constraint IC_p for θ_l is satisfied, then Constraint IC_p for θ_h will be satisfied. Apply Cho and Kreps (1987)'s Intuitive Criterion, the least costly separating equilibrium is the one under which $f_1^l = 0$ and Constraint IC_p for θ_l is binding.

Lemma 5 *The principal's problem $P1$ is equivalent to the following maximization problem $P1'$ for each type of θ:*

$$\max_{f_1^h, b_1^h} P(\theta_h, e(\theta_h, b_1^h))(1-b_1^h) - P(\theta_l, e(\theta_h, b_1^h))(1-b_1^h)$$

Substituting Constraint IC_p for θ_l and Constraint IC_a into the objective function, problem $P1$ with two constraints is simplified to problem $P1'$. The optimal level of bonus under asymmetric information is given by the following equation:

$$\left(\frac{\partial P(\theta_h, e(\theta_h, b_1^h))}{\partial e} - \frac{\partial P(\theta_l, e(\theta_h, b_1^h))}{\partial e}\right)\frac{\partial e(\theta_h, b_1^h)}{\partial b_1^h}(1-b_1^h) = P(\theta_h, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_h, b_1^h)) \quad (2.3.2)$$

Similar to Equation 2.3.1, the LHS of Equation 2.3.2 measures the marginal benefit due to a unit increase in b_1 through an increase in the agent's effort, deducting compensation. The RHS of Equation 2.3.2 represents the marginal cost. In contrast with Equation 2.3.1, it is not the output $P(\theta_h, e)$ but the output sensitivity to private information that matters for the characterization of the optimal level of bonus.

On the one hand, higher θ directly results in a higher output. On the other hand, an agent who receives a higher signal will work harder, leading indirectly to a higher output. When deciding the optimal bonus, the principal maximizes the part of profit that directly comes from the private information θ and only considers the effect of b_1

[15]Since the production technology exhibits super-modularity, the concavity of the principal's truth-telling constraint can be ensured.

on this part of profit. Profit coming indirectly from the private information, which is through an increase in effort after the agent observing the contract, will be allocated to the agent as rent for signalling in order to prevent the low type principal from mimicking.

The following proposition shows the condition under which the bonus does not provide feedback to the agent:

Proposition 9 ***Bonus not providing feedback***
If the following information invariant condition holds, $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} = 0$*, then the bonus is information insensitive, or* $b_1^h = b_1^{s,h}$*. The salary is* $f_1^h = \{P(\theta_l, e(\theta_h, b_1^h) - P(\theta_l, e(\theta_l, b_1^l))\}(1 - b_1^l)$*, and* $f_1^h > f_1^l = 0$*.*

In other words, the production technology can be expressed as a product of two separating functions of θ and e respectively. Under information asymmetry, the principal only considers how the bonus affects the part of profit that directly comes from private information θ. If the condition is satisfied, the MRS between effort and bonus is the same as that under symmetric information. In other words, the information invariant condition mutes any effects of information asymmetry on the bonus. The principal finds that maximizing the profit coming directly from θ is the same as maximizing the total profit. Its only role is to provide incentives. The principal fully relies on the salary to provide feedback, while the bonus is paid at a level as if the agent knew her own type (recall Proposition 8).

By offering f_1^h, the principal credibly communicates its private information to the agent, which changes the agent's belief and motivates her to make more efforts. This channel is different from the incentive effect provided by bonus b_1. The salary affects the agent's effort through convincing the agent of her ability to achieve a higher output, while the pay per unit of effort is held constant. The incentive channel affects the agent's effort level through raising the pay b_1, while the agent's belief of productivity is held constant.

Salary is increasing in the signal that the principal receives. This is to say, to prevent the principal of a low type from mimicking, the principal of a high type needs to pay a higher salary to signal her private information. The principal in this case will not want to offer a higher bonus to substitute the salary, because it would imply giving away too much profit.

Proposition 10 ***Bonus Providing Feedback I***

$$If \quad \frac{\partial^2 lnP(\theta, e)}{\partial\theta\partial e} > 0, \quad b_1^h > b_1^{s,h}.$$

- *The first period salary is*

$$f_1^h = \underbrace{\{P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l))\}(1 - b_1^h)}_{>0} - \underbrace{(P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)}_{>0},$$

If the information invariant condition is not satisfied, Proposition 10 shows that the bonus could provide feedback as well. A higher evaluation outcome (higher θ) improves the marginal productivity of effort in terms of log likelihood of high output. This is equivalent to saying that the MRS between effort and bonus, which determines the level of b_1, is higher than that under symmetric information. This leads to higher bonus under asymmetric information, that is, $b_1^a > b_1^s$ (recall Proposition8).

As a result, to prevent a principal with low productive efficiency from mimicking, a principal with high productive efficiency will increase the bonus beyond the level under symmetric information. This is because when the total factor productivity contributes a lot to the firm's output function, signalling through bonus becomes cheaper, as this is very costly for the low type to mimic. Similar to the case under the information invariant condition, the salary could provide feedback to the agent, as the first term of f_1^h is positive. However, the last term of f_1^h is negative, which implies that the importance of salary in signalling is undermined, because the bonus takes over the role of signalling.

Proposition 11 ***bonus Providing Feedback II***

$$If \ \ \frac{\partial^2 lnP(\theta, e)}{\partial\theta\partial e} < 0, \quad b_1^h < b_1^{s,h}.$$

- *The first period salary is*

$$f_1^h = \underbrace{\{P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l))\}(1 - b_1^h)}_{>0} - \underbrace{(P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)}_{<0},$$

Proposition 11 shows that if log super-modularity is negative, bonus is even lower than the level under symmetric information. This result characterizes the condition under which there is under-effort-provision compared to the case of symmetric information. In this case, the contribution of the agent's skills to output is so insignificant that the firm finds it less costly to use under-effort-provision to signal the agent's type, compared to more profit sharing when the log super-modularity is positive.

2.4 A Two-period Model

I choose a specific form of production technology of which the the log-supermodularity between productive efficiency and labour effort is zero, $P(\theta_i, e_t) = \theta_i e_t$, and a quadratic disutility function $\psi e_t = \frac{1}{2}e_t^2$. As shown in the baseline model, the optimal contract offered under this technology includes a bonus which is not information sensitive. Any subsequent changes that lead to a bonus different from the level under symmetric information will thus not be a result of a change in type but rather a consequence of principal's ability to commit to a long term contract. I will come back to this point in more details.

Define a different contract specification that a principal of type θ_h offers at the date 0: $\{b\{.\}, f\{.\}\} = \{b_1^{hh}, b_2^{hh}, b_3^{hh}, f^{hh}; b_1^{hl}, b_2^{hl}, b_3^{hl}, f^{hl}\}$. $b\{.\}$ are bonuses, $b_1^i > 0$, if $y_1 = 1$ otherwise zero, $b_2^i > 0$, if $y_2 = 1$ otherwise zero, $b_3^i > 0$, if $y_1 = y_2 = 1$ otherwise zero, f^i is the fixed compensation.

Lemma 6 *Contract* $\{w_{00}^{hh}, w_{10}^{hh}, w_{01}^{hh}, w_{11}^{hh}; w_{00}^{hl}, w_{10}^{hl}, w_{01}^{hl}, w_{11}^{hl}\}$ *can be equivalently implemented by contract* $\{b\{.\}, f\{.\}\} = \{b_1^{hh}, b_2^{hh}, b_3^{hh}, f^{hh}; b_1^{hl}, b_2^{hl}, b_3^{hl}, f^{hl}\}$.

Lemma 6 shows that a general contract can be implemented by a different specification which consists of fixed salaries and variable bonuses. The fixed component will be paid to the agent under any realizations of the output. The variable component is contingent on the realization of the output measures. This specification offers us a convenient interpretation of the compensation structure.

Long term contracts will benefit the firm in two ways: First, principal could use cross-pledging to alleviate the incentive problem; Second, the principal could signal her private information by using a bonus based on second period output measure. In order to see the two effects clearly, I proceed by first considering the case in which cross-pledging of two periods' payoff is not allowed, from which I obtain the basic mechanism. I then characterize the optimal contract which allows cross-pledging.

2.4.1 Without Cross-pledging

In this subsection, I first disallow the principal to offer long-term contracts. This corresponds to the situation where committing to a long term contract is impossible. The following lemma shows that optimal contracts are stationary in the sense that it does not depend on the private information in the prior period.

Lemma 7 ***Optimal one-period contracts***. *If committing to a long term contract is impossible,*

- *The two optimal one-period contracts for θ_h and θ_l in the first period are:*
  For θ_h, $\{f_1 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_1(1) = \frac{1}{2}\}$.
  For θ_l, $\{f_1 = 0, b_1(1) = \frac{1}{2}\}$.

- *The two optimal one-period contracts for θ_h and θ_l in the second period are:*
  For θ_h, $\{f_2 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_2(1) = \frac{1}{2}\}$.
  For θ_l, $\{f_2 = 0, b_2(1) = \frac{1}{2}\}$.

According to Lemma 7, when commitment is impossible, private information in the first period does not affect the equilibrium contract in the second period. In the second period, the contract for a low type who was low in the first period is the same as the one for a low type who was once a high-skilled worker in the first period. When new information arrives in the second period, the principal wants to renegotiate and make a take-it-or-leave-it offer. The principal operates as if she was in two separate one-period models.

I then remove the restriction on being able to commit and allow the principal to offer long-term contracts. However, I disallow cross-pledging. In other words, the principal cannot use equity compensation that can only be vested at the end of the second period.

Such contracts depart away from the short term contracts in the sense that commitment allows for a reallocation of the agent's first period rent to the second period. However, not all long-term contracts are renegotiation-proof. One has to bear in mind that renegotiation might happen both when the productive efficiency continues to be good and deteriorates. On the one hand, the principal wants to signal and separate again in the second period when good private information arrives. On the other hand, when bad information arrives, a long-term contract with $f_l > 0$ may be subject to renegotiation, as salary has no signalling value anymore. The following proposition characterizes the renegotiation-proof contract.

Lemma 8 *In the equivalent contract specification $\{b\{.\}, f\{.\}\}$, if b_3^{hh}, b_3^{hl} and b_3^{ll} are set to 0, to induce effort, the following components are greater than zero: $b_1^h > 0$, $b_2^{hh} > 0$, $b_2^{hl} > 0$, $b_1^{ll} > 0$ and $b_2^{ll} > 0$. In order to satisfy the limited liability, $f^1 \geq 0$ and $f^h \geq 0$.*

The logic behind Lemma 2.10 is that when cross-pledging is not allowed, the agent's incentive problems in the two periods are tied only through the principal's truthtelling constraint, otherwise they are independent of each other.

Proposition 12 ***Low separating profit*** *($\theta_h < 2\theta_l$). Define $\pi_{2l} = \frac{1}{4}\theta_l^2(2 - \frac{\theta_h}{\theta_l})$.*

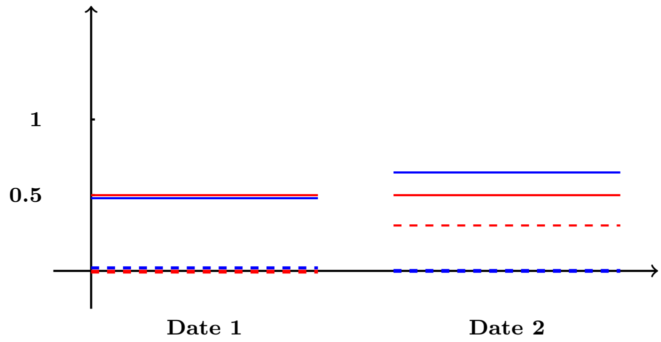


Figure 2.3: Equilibrium Contracts under Low Separating Profit

Note: Dashed – salary; Line–bonus; Red: $\theta_h \to \theta_h$; Blue: $\theta_h \to \theta_l$.

- *The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_1 = \frac{1}{2}$; $b_2^{hh} = \frac{1}{2}$, $f^h = \frac{1}{4}\theta_l\theta_h - \pi_{2l}$; $b_2^{hl} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1})$, $f^l = 0$.*
- *The principal will offer two short-term contracts to type θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.*

Proposition 13 ***High separating profit*** *($\theta_h \geq 2\theta_l$).*

- *The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_1 = \frac{1}{2}$; $b_2^{hh} = \frac{1}{2}$, $f^h = \frac{1}{2}\theta_l(\theta_h - \theta_l)$; $b_2^{hl} = 1$, $f^l = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)$.*
- *The principal will offer two short-term contracts to type θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.*

Two propositions show that the aggregate welfare differs between these two cases depending on the value of b_{2l}. The principal can send out her positive signal at date 0 in two ways. She could either pay a high salary or promise greater profit sharing even if the agent becomes low skilled in the next period. The greater profit sharing (a higher bonus) the principal offers, the greater aggregate welfare the contract could achieve. However, in the first case when $\theta_h < 2\theta_l$, leaving positive amount of salary paid at the end of date 2 to a low-skilled agent is not renegotiation-proof. Because salary does not have either incentive or signalling value. The principal after receiving bad information will want to renegotiate the contract and substitute the salary with a higher bonus. Figure 2.3 depicts the equilibrium contracts.

However, when $\theta_h \geq 2\theta_l$, the principal will want to pay a salary to the agent after receiving bad private information in the second period. Because the separating profit and mimicking profit are both very high, if the principal receives good private information in the first period, she does not want to renegotiate the contract by using

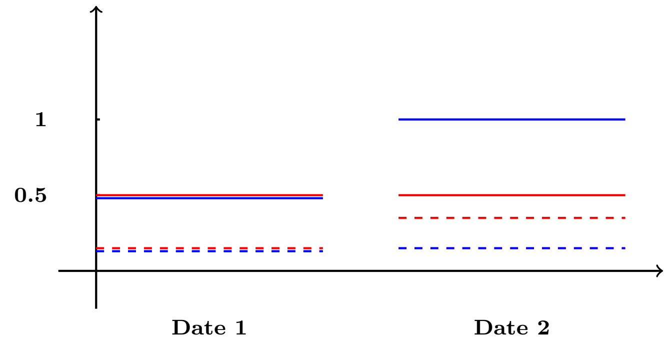


Figure 2.4: Equilibrium Contracts under High Separating Profit

Note: Dashed – salary; Line–bonus; Red: $\theta_h \to \theta_h$; Blue: $\theta_h \to \theta_l$.

more bonus to signal, once the bonus reaches the optimal level 1. In both cases, the principal pays more rent to the agent in the form of a higher bonus based on y_2. With long term contracts, the principal is able to reallocate the cost of signalling for the first period to be based on the second period measure, which allows greater profit sharing with the agent and induces more effort from the agent.

If $\theta_h \geq 2\theta_l$, the renegotiation-proof contracts identified in Proposition 2.10 that offered by a high type principal can be implemented by paying the agent at the beginning $f = \frac{1}{8}\theta_l(\theta_h - 2\theta_l)$ at date 1 and same amount of salary at the date 2 if the productive efficiency deteriorates. The principal will not want to renegotiate the positive salary away by offering a higher bonus. This long term contract gives the agent the right to obtain at least what is offered in the contract even the situation worsens in the next period. A downward-rigid contract with a positive salary paid in each period can only be implemented when there is enough variation in the levels of productive efficiency. In such a contract, both the salary and the bonus could be made downward rigid. Figure depicts this implementation.

To summarize, my result does not only explain downward rigidity in the total compensation but in the salary and the bonus as well. In addition, a discretionary salary award sends out a stronger signal than a discretionary bonus award. The intuition is that a salary award is more costly for a firm as it has zero incentive value. The agent will still have to make effort in order to obtain the bonus even the contractual level of pay-performance sensitivity is raised, which allows the firm to recoup at least some profit.

2.4.2 With Cross-pledging

When cross-pledging is allowed, the principal will use b_{3h}, b_{3l} and b_{3ll} to alleviate the incentive problem. In other words, the principal could use equity compensation that can only be vested at the end of the second period. By shirking in one period, the agent reduces the probability of full success and reward for the effort made in the other period.

Proposition 14 *If information is asymmetric,*

- *The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_3^{hh} = \sqrt{\frac{1-(1-q)\theta_l^2}{q\theta_h^4}}$, $b_2^{hl} = \frac{3q\theta_l^2-2q^2\theta_l^2-q\theta_h+q^2\theta_h-\theta_l^2+1}{2q(1-q)\theta_l^2}$, and $b_3^{hl} = \frac{1}{2\theta_l^2} - \frac{\theta_h^2\beta^{hh}b_3^{hh}}{2(1-q)\theta_l^2}$. $f^h = f^l + \theta_l\theta_h^2 b_3^{hh}(1-\theta_h b_3^{hh})$, and $f^l = 1-\theta_l+(\theta_h-\theta_l)\theta_l^2 b_3^{hl}$.*
- *The principal will offer a long-term contract to type θ_l at date 0, $b_3^{ll} = \frac{1}{\theta_l^2}$.*

Compared to the optimal contract without cross-pledging, the bonus in this case based on the first period measure is zero. It leads to less rent extraction due to limited liability. The agent will make more effort in the first and second period in order to obtain a higher bonus based on two measure. The principal of high productive efficiency only uses b_{3h} and f_h to induce effort and signal her private information.

Corollary 6 *Under information asymmetry, The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_3^{hh} > \frac{1}{\theta_h^2}$, $b_2^{hl} > 1$, and $b_3^{hl} < 0$. $f^h > 0$, $f^l > 0$.*

According to Corollary 6, there are several interesting features about the optimal contract.

First, due to information asymmetry, the principal with deteriorating efficiency will offer the agent more bonus based on the second period measure. This can be seen easily with the Lemma 2.10 below. The intuition is the same as in the case without cross-pledging. If $b_{2l} = 0$ and the principal only uses b_{3l} to induce effort and provide feedback for the first period, the principal will want to renegotiate the contract in the second period when she is privately informed. Since the effort in the first period is already sunk, the principal in the second period will want to renegotiate b_{3l} down and increase b_{2l} to provide feedback for the first period. Consequently, the compensation offered by a principal with decreasing efficiency pays more compensation based on the second performance measure. Unlike in the case of symmetric information, here, b_{3l} does not enter into the the principal's maximization function in a linearly fashion. On the one hand, like in the case of symmetric information, high b_{3l} leads to less rent extracted by the agent. On the other hand, it increases the mimicking profit of the principal with low efficiency in the first period. Consequently, b_3^{hl} is smaller than the level under symmetric information.

Lemma 9 *$b_.^s$ denote the bonus under symmetric information. If information is symmetric:*

- *Any contracts $\{b\{.\}, f\{.\}\}$ for type hh, hl and ll under symmetric information can be replaced by contracts that only consist of b_3^{hh}, b_3^{hl} or b_3^{ll} respectively.*
- $b_1^{hh} = b_2^{hh} = 0$, $b_3^{hh} = \frac{1}{\theta_h^2}$; $b_1^{hl} = b_2^{hl} = 0$, $b_3^{ll} = \frac{1}{\theta_l \theta_h}$; $b_1^{ll} = b_2^{ll} = 0$, $b_1^{ll} = \frac{1}{\theta_l^2}$.

Lemma 2.10 shows that when cross-pledging is allowed, the principal will optimally use a bonus based on two performance measures to induce effort. In this way, the principal minimizes the rent the agent extracts due to the limited liability.

Second, $b_3^{hh} > \frac{1}{\theta_h^2}$. In order to induce sufficient effort in the first period, the principal with constantly high efficiency will offer a higher long-term equity compensation to induce first period effort, because the agent knows that if the productive efficiency declines, the principal will offer a bonus based on the second period output measure. Expecting this, the agent's first period incentive would be lowered if the principal did not raise b_3^{hh}. Long-term equity compensation that could only vested at the end of the second period is also used to provide signal.

I also verify that pooling equilibrium does not survive the Intuitive Criterion.

Lemma 10 *With the cross-pledging effect, pooling equilibrium does not survive the Intuitive Criterion.*

2.5 Bonus Caps and Efficiency Implications

A banker bonus cap was passed by the EU Parliament in April 2013 and was set to go into effect in January 2014. The cap will limit bonuses for employee's 2014 performance year to the level of the employee's salary, or to twice the employee's salary if shareholder approval is obtained. On February 25th 2014, the European Parliament (EP) and the European Council (Council), agreed to restrict the bonuses of retail asset managers.[16]

In this section, I study the impact of bonus caps on the efficiency based on the baseline model. The timeline is the same as the baseline model. After observing the

[16]The Council also agreed not to include a bonus cap for managers and advisors of UCITS funds (UCITS funds are similar to US-registered mutual funds). In place of the cap, the Council and EP resolved that at least 50% of bonus amounts must be paid in shares of the fund under management, and at least 40% of bonus amounts must be deferred for three years.

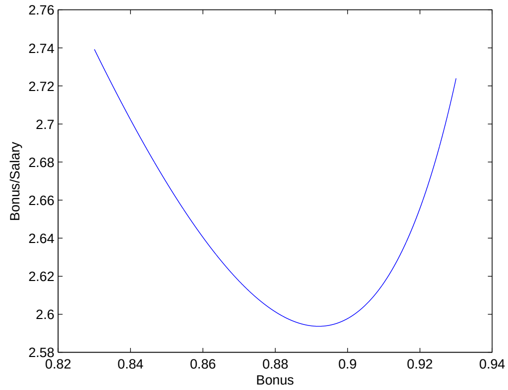


Figure 2.5: Bonus over Salary

Note: Parameter values are $k = 1$, $\theta_h = 0.9$, $\theta_l = 0.4$.

contract, the agent makes effort $e \in [0, 1]$ with disutility $\psi(e) = \frac{1}{2}e^2$. An output y is realized at the end of date 1, and $y \in \{0, 1\}$. To make the efficiency comparison tractable, the production technology takes a parametric form that, at the end of date 1, the probability of getting $y = 1$ is $P(\theta, e) = \theta_i e(\theta_i + \frac{1}{2}ke)$.

Salary equals the profit of the low type principal if she mimics minus the profit if she does not. As shown in Figure 2.5, the ratio of bonus over salary is not monotonic in bonus. When bonus is low, the deduction of non-mimicking profit is very sensitive to salary changes, and an increase in bonus means a decrease in the ratio because of the high sensitivity of the deduction of the non-mimicking profit to bonus changes. When the bonus is high, an increase in bonus implies too much profit sharing relative to more effort provision thus low mimicking profit, thus leading to an increasing bonus to salary ratio.

Table 2.1 provides a simple analysis of the welfare of the board and the CEO under three cases. The first column is the contract the firm chooses without a bonus cap. The second and third column represent the contracts under the cap on bonus to salary ratio, which is set at 2.5939. The second column shows the contract with the highest bonus possible under the cap. The third column is an alternative contract with the same bonus as under no cap. It shows that the contract in the second column yields greater efficiency at the expenses of the principal. The board has to pay a greater signalling cost in order to abide by the rule. The agent benefits from bonus caps. However, the board may consider to increase only the salary. As is shown in the third column, such an approach may lead to greater profit destruction to the principal compared to the second approach, as there is no greater effort produced thus no efficiency gain.

	Without a Cap	With a Cap	Change salary only
Bonus over salary	2.5955	2.5939	2.5939
Salary	0.3415	0.3431	0.3417
Bonus	0.8863	0.89	0.8863
Profit of the firm	0.6294	0.6293	0.6292
Profit of the CEO	1.6151	1.6489	1.6153
Total profit	2.2445	2.2782	2.2445

Table 2.1: A Comparison of Efficiency

Note: Parameter values are $k = 1$, $\theta_h = 0.9$, $\theta_l = 0.4$.
The first and the second column show the optimal contracts without and with a cap on bonus to salary ratio, respectively. The third column characterises, under bonus caps, the optimal contract, by allowing for only salary adjustments.

The intuition is as follows. In this example, the bonus under symmetric information is 0.8969. An increase in the bonus implies a faster increase in the salary if there is under provision of effort, because the mimicking profit increases faster than the bonus. By imposing a limit on the ratio of bonus to salary, the regulator enforces the board to adjust the bonus and consequently the salary higher, achieving a lower bonus to salary ratio. In other words, the principal has to pay a higher signaling cost and consequently induces more effort from the agent. As a result, the overall efficiency is enhanced.

If the bonus is already more than the level under symmetric information, an increase in the bonus implies a slower increase in the salary. By imposing a limit on the ratio of bonus to salary, regulators can only induce the principal to give away more profit to the agent through an increase in the salary. In this case, the board just reallocates the profit without any efficiency enhancement.

It would be unfair to criticize and guide the policy simply based on my result, considering other important intentions of this policy, such as to curb managers' risk-taking incentives in highly leveraged banks. However, my model offers a potential reason for the moves taken by some banks since last year and predicts possible consequences of this policy going forward, especially for banks which operate in a volatile environment and consider hiring new executives. Some banks have restructured their CEOs' compensation by increasing the base salary, resulting in a higher estimated total pay[17].

However, this policy also creates its own distortion. Imposing a high signalling cost may dampen information sharing. In other words, the principal of high productive efficiency may find it better not to provide the signal and pool with a low type. If

[17] For instance, HSBC's chief executive, Stuart Gulliver, will see his basic pay jump from £1.2 million to £2.9 million thanks to a £32,000 weekly shares of "fixed pay allowance". In 2012 the top HSBC earner (presumably Gulliver) earned £7 million: £650,000 in salary and £6.35 million bonus. Now Gulliver is certain to earn £4.2 million, potentially £11.4 million if he actually does a good job.

so, bonus caps will exacerbate the information problem by making truth-telling more costly or even impossible.

In order to achieve efficiency improvement, finding the appropriate ratio of bonus over salary is important. If it is too low, firms will find it difficult to motivate and retain the talented. The private sector expresses such concerns. Penny Hughes, the non-executive director who chairs the remuneration committee of RBS, said in the annual report:

> "I know it is not always easy to accept, but if RBS is to thrive we must do what it takes to attract and keep the people who will help us achieve our goals. While we are sensitive to public opinion, particularly given our ownership structure, the ability to pay competitively is fundamental to getting RBS to where we need it to be."

Another implication of this extension is that the heterogeneous effects of bonus caps on firms with different technologies need to be taken into account. An overall effect of bonus caps on the societal welfare depends on the distribution of the different types technology. In some firms, managerial talent contributes a lot to the output function, while in others firms it does not. Imposing bonus caps, however, only improves the efficiency of the latter. Hence, bonus caps improve the societal efficiency only if there are more firms with production technologies that rely less on managerial skills.

2.6 Transferable Skills and Disclosure Policies

Researchers in the field of executive compensation have long been interested in pay-performance sensitivity and its relation with human capital. Frydman and Jenter (2010) find that the evolution of managerial compensation since World War II can be broadly divided into two distinct periods. Prior to the 1970s, they observe low levels of pay and moderate pay-performance sensitivities. From the mid-1970s to the early 2000s, compensation levels grew dramatically, and equity incentives tied managers' wealth closer to firm performance. If the principal's private information is about the matching quality between two parties, this extension aims to provide new interpretations of the above empirical finding from the the perspective of transferable managerial skills under an informed principal framework.

In the previous analysis, the agent's reservation utility does not vary with her type. This implies that the agent's skill is non-transferable or other firms perceive the agent skills to be firm-specific. Nevertheless, as documented in Murphy and Zábojník (2007) and Dutta (2008), executives do not only possess firm-specific skills but also transferable skills. This section extends the model by assuming type-dependent reservation utility. When the agent's reservation utility becomes type contingent, her participation constraint may be binding. This means that the principal needs to provide higher compensation in order to retain the agent.

This case becomes particularly interesting in light of the mandatory compensation disclosure policy. The current executive compensation disclosure requirements applicable to most US domestic issuers, and to those non-US companies that do not qualify as foreign private issuers, were adopted by the US Securities and Exchange Commission (SEC) in 1992. In the setting of this paper, mandatory compensation disclosure will affect the compensation level and structure. The value of the outside option of the agent relies on whether the market believes the agent has better skills or not. If firms are not required to disclose the compensation, the market will then not know or find it very costly to assess the agent's skills. Constant reservation utility represents an extreme case in which the agent has only firm-specific skills or the market has no way to infer the agent's general skills. Once the agent skills become transferable, the second-best contract that can be implemented if information invariant condition is satisfied is not feasible due to a binding participation constraint.

The timeline in this extension is the same as in the one-period baseline model. However, the reservation utility of the high type is R and the low type is 0. At the interim date 0.5, as previously, the agent makes effort $e \in [0, 1]$ with disutility $\psi(e) = \frac{1}{2}e^2$. At the end of date 1, the probability of getting $y = 1$ is $p = P(\theta, e) = \theta e$. It can be easily verify that the agent's effort of type θ_i ($i \in \{l, h\}$) given a contract $\{f_1^i, b_1^i(1)\}$ is $e^{i*} = \theta_i b_1^i$. Thus the maximization program for a principal who receives a high signal is:

$$
\begin{aligned}
&\max_{f_1^h, b_1^h} \theta_h e^{h*}(1 - b_1^h) - f_1^h \\
s.t. \quad & \theta_l e^{l*}(1 - b_1^l) - f_1^l \geq \theta_l e^{h*}(1 - b_1^h) - f_1^h && IC_p \\
& \theta_h e_h^* b_1^h + f_1^h - \frac{1}{2} e_h^{*2} \geq R && IR_a
\end{aligned}
$$

When $R = 0$ as in the baseline analysis of the previous section, the IR_a constraint is not strictly binding because the agent is protected by limited liability. However, when R is sufficiently large, the surplus the agent extracts due to the limited liability may not be large enough. Assume λ to be the Lagrangian multiplier of the IR_a constraint. I consider the case in which IR_a is strictly binding ($\lambda > 0$).

Regarding the IC_p constraint, under the least costly separating equilibrium the principal pays $f_1^l = 0$ to a low skilled agent, since there is no signalling gain to motivate a low skilled agent. It can be easily verified that $b_1^l = \frac{1}{2}$. A principal who receives a high signal pays the agent only at a level that just makes the IC_p constraint binding. IC_p constraint can thus be absorbed by substituting b_1^h into the objective function and the IR_a constraint.

Proposition 15 ***Pay Performance Sensitivity and Managerial Skills***
Assume $b_1^{o,h}$ to be the bonus paid to a high skilled agent with zero reservation utility, and f_1^h and b_1^h with positive reservation utility.

- *If $0 \leq R \leq \underline{R}$, IR_a constraint is not binding ($\lambda = 0$). $b_1^h = b_1^{o,h} = \frac{1}{2}$. Only separating equilibrium exists.*

- *If $\underline{R} < R \leq \bar{R}$, IR_a constraint is binding ($\lambda > 0$).*

$$b_1^h = \frac{\Delta\theta + \lambda\theta_l}{2(\Delta\theta + \lambda\theta_l) - \lambda\theta_h}$$

  And $b_1^h > b_1^{o,h} = \frac{1}{2}$. Only separating equilibrium exists.

- *If $R > \bar{R}$, only pooling equilibrium exists.*

The above proposition indicates that when the agent possesses general skills and her compensation is subject to mandatory disclosure, the agent receives a greater bonus. When the reservation utility for the high type is zero or sufficiently small, the contract could still induce the second best effort ($b_1^{o,h} = \frac{1}{2}$) under the information invariant condition. This is because the rent that the agent extracts due to limited liability is greater than the value of her outside option. When the reservation utility is too high, the principal no longer finds it profitable to credible signal her private information. Instead, it chooses to pool with the board which receives a low signal. As a result, no separating equilibrium exists.

When the reservation utility is at an intermediate level, the IR_a constraint binds. The agent's general skills imply higher performance sensitivity. One might propose to set bonus at $\frac{1}{2}$ level and to increase the salary so that the IR_a binds. However, this is not optimal. To see the intuition more clearly, the agent's utility is $\frac{1}{2}\theta_h^2(b_1^h)^2$. A binding IR_a constraint (or a positive shadow price of the constraint) implies that the marginal benefit relative to the marginal cost of setting bonus at $\frac{1}{2}$ increases. Hence, to increase the bonus makes the IR_a constraint more easily bind. In other words, when the agent's skills become sufficiently transferable, compensation disclosure may result in high powered incentives.

Similar to Oyer (2004), a binding participation constraint in my model also results in a high incentive pay. However, the participation constraint only becomes binding if the principal's truth-telling constraint is satisfied. That is, a high incentive pay exists only if the principal provides a credible signal for the agent's managerial skills. Otherwise, the participation constraint only needs to bind on average. Thus a shift in the relative importance of general skills versus firm-specific skills leads to higher pay-performance sensitivity, because the principal needs to provide strong feedback to motivate and retain the high skilled agent.

This extension further sheds light on the disclosure policy. The US Securities and Exchange Commission (SEC) in 1992 has adopted executive compensation disclosure requirements applicable to most US domestic issuers, and to those non-US companies that do not qualify as foreign private issuers. The recent Dodd-Frank Wall Street Reform and Consumer Protection Act contains new disclosure policies which affect the governance of issuers.[18]

Compensation disclosure policies are intended to improve corporate governance by, for example, curbing managerial power and facilitating investor monitoring. In the setting of this paper, mandatory disclosure helps transmit the principal's private information of the agent's managerial skills to the market. While those policies reduce information asymmetry and lead to more competitive pay, they, as indicated in this extension, also increase a firm's cost of providing feedback when executives possess transferable skills. In the extreme case, the signalling cost may become so large that the principal chooses not to provide feedback, which exacerbates information asymmetry and discourages effort.

2.7 Conclusions

This paper characterizes the optimal contract which deals with the moral hazard and signalling problems at the same time in a dynamic environment where the principal's private information changes over time. Contracts thus have two roles of providing feedback and incentives to the agent. I show the condition under which the principal solely relies on the salary to signal her private information to the agent. Bonuses could also be information sensitive under certain conditions. In other words, the bonus may have dual roles, feedback provision and incentive provision. Firms either use more profit sharing or under-effort-provision to signal her private information.

I choose a specific production technology with zero log-supermodularity in skills and efforts. Such a technology allows me to assign the signalling role to the salary and the incentive role to the bonus in a one-period model. I first analyze a benchmark case in which the principal cannot commit to long-term contracts. Because the agent

[18] For instance, Section 953 requires additional disclosure about certain compensation matters, including pay-for-performance and the ratio between the CEO's total compensation and the median total compensation for all other company employees.

anticipates that the principal will make a new take-it-or-leave-it offer when new information comes, she will not agree to an arrangement which promises a high bonus in future. The equilibrium contract is thus stationary in the sense that the second period contract does not depend on the first period private information. The principal designs the contract as if she was in two separate one-period models.

If commitment, however, is possible, the principal could promise a higher bonus based on the second period performance measure as a way of providing signal for the initial high productivity, as the salary has zero incentive or signalling value for inducing the second period effort. If the productive efficiency continues to be high, the principal wants to provide an additional salary instead of high bonus, as salary has signalling value. Such a contract achieves greater efficiency by giving the agent more profit sharing opportunities and inducing greater efforts in the second period. The principal pays more rent in the second period in exchange for less paid to the agent in the first period.

This paper also sheds light on the attempts of recent regulations to curb managerial bonuses. My paper suggests that a limit on the ratio of bonus to salary may help improve efficiency by enforcing the principal to pay more for signalling. With a cap on the bonus to salary ratio, some firms which use too little bonus to provide feedback, will have to raise bonus and consequently adjust upward the salary even more. Such a change in compensation structure induces greater effort from the agent and leads to efficiency improvement. I also consider an extension in which the manager possesses general skills. It suggests that a shift in the relative importance of general skills versus firm-specific skills leads to higher pay-performance sensitivity, because the principal needs to provide stronger feedback to motivate and retain a high skilled agent.

Several important conclusions could be further drawn from this paper. First, bonus is not only sensitive to publicly observable information but also to private information. The mapping from objective measures to bonuses thus contains a principal's private information which is not observable to econometricians. Neglection on this important channel might lead to over-estimation of the incentive effect of performance based pay. Second, salary plays a crucial role in facilitating communication in organizations, especially in cases where bonus does not provide feedback. My paper suggests that salary can be performance sensitive as well.

2.8 References - Chapter 2 and 3

Daron Acemoglu. Directed technical change. *The Review of Economic Studies*, 69(4): 781–809, 2002.

Viral V Acharya, Kose John, and Rangarajan K Sundaram. On the optimality of resetting executive stock options. *Journal of Financial Economics*, 57(1):65–101, 2000.

Philippe Aghion, Philippe Bacchetta, and Abhijit Banerjee. Financial development and the instability of open economies. *Journal of Monetary Economics*, 51(6):1077–1106, 2004.

Philippe Aghion, Diego Comin, and Peter Howitt. When does domestic saving matter for economic growth? Working Paper 12275, National Bureau of Economic Research, 2006.

Philippe Aghion, Antoine Dechezleprêtre, David Hemous, Ralf Martin, and John Van Reenen. Carbon taxes, path dependency and directed technical change: evidence from the auto industry. CEP Discussion Papers No.1178, *forthcoming, Journal of Political Economy*, 2014.

George A Akerlof and Janet L Yellen. The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, pages 255–283, 1990.

Armen A Alchian and Harold Demsetz. Production, information costs, and economic organization. *The American economic review*, pages 777–795, 1972.

Linda Argote. *Organizational learning: Creating, retaining and transferring knowledge.* Springer Science and Business Media, 2012.

Kenneth Arrow. Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, NBER Chapters, pages 609–626. Princeton University Press, 1962.

George Baker, Robert Gibbons, and Kevin J. Murphy. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, 109: 1125–1156, 1994.

Jeffrey S Banks and Joel Sobel. Equilibrium selection in signaling games. *Econometrica*, 55(3):647–661, 1987.

Warren Bennis and Burt Nanus. *Leaders: The strategies for taking charge*. New York: Harper and Row, 1985.

Marianne Bertrand and Sendhil Mullainathan. Do ceos set their own pay? the ones without principals do. *Quarterly Journal of Economics*, 116, 2001.

Truman F. Bewley. Fairness, reciprocity, and wage rigidity. In Peter Diamond and Hannu Vartiainen, editors, *Behavioral Economics and Its Applications*, pages 157–188. Princeton and Oxford: Princeton University Press, 2007.

Patrick Bolton, Markus K Brunnermeier, and Laura Veldkamp. Leadership, coordination, and corporate culture. *The Review of Economic Studies*, 80(2):512–537, 2013.

Clive Bull. The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics*, 102:147–59, 1987.

Christian Casal and Christian Caspar. Building a forward-looking board: Mckinsey quarterly. http://www.mckinsey.com/insights/strategy/building_a_forward-looking_board, 2014.

In-Koo Cho and David M Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, pages 179–221, 1987.

In-Koo Cho and Joel Sobel. Strategic stability and uniqueness in signaling games. *Journal of Economic Theory*, 50(2):381–413, 1990.

Clayton M. Christensen. The rigid disk drive industry: A history of commercial and technological turbulence. *The Business History Review*, 67(4):531–588, 1993.

Simon C. Collinson and David C. Wilson. Inertia in japanese organizations: Knowledge management routines and failure to innovate. *The RAND Journal of Economics*, 27 (9):1359–1387, 2006.

Teresa Joyce Covin and Ralph H Kilmann. Participant perceptions of positive and negative influences on large-scale change. *Group and Organization Management*, 15 (2):233–248, 1990.

Wouter Dessein and Tano Santos. Adaptive organizations. *Journal of Political Economy*, 114(5):956–995, 2006.

James Dow and Enrico C. Perotti. Resistance to change. Working Paper 48.2012, FEEM, 2010.

Sunil Dutta. Managerial expertise, private information, and pay-performance sensitivity. *Management Science*, 54(3):429–442, 2008.

Florian Ederer. Feedback and motivation in dynamic tournaments. *Journal of Economics & Management Strategy*, 19(3):733–769, 2010.

Florian Ederer and Gustavo Manso. Is pay for performance detrimental to innovation? *Management Science*, 59(7):1496–1513, 2013.

Merle Ederhof, Madhav V. Rajan, and Stefan Reichelstein. Feedback and motivation in dynamic tournaments. *Foundations and Trends in Accounting*, 5(4):243–316, 2011.

Daniel Ferreira and Marcelo Rezende. Corporate strategy and information disclosure. *The RAND Journal of Economics*, 38(1):164–184, 2007.

Daniel Ferreira, Gustavo Manso, and Andre C Silva. Incentives to innovate and the decision to go public or private. *Review of Financial Studies*, 27(1):256–300, 2014.

Michael Fishman and Kathleen Hagerty. Mandatory versus voluntary disclosure in markets with informed and uninformed customers. *Journal of Law, Economics, and Organization*, 19(1):45–63, 2003.

Carola Frydman and Dirk Jenter. Ceo compensation. *Annual Review of Financial Economics*, 2(1):75–102, 2010.

William Fuchs. Contracting with repeated moral hazard and private evaluations. *American Economic Review*, 97:1432–1448, 2007.

William Fuchs. Subjective evaluations: Discretionary bonuses and feedback credibility. Technical report, IZA Discussion Paper, 2013.

William Fuchs. Subjective evaluations: Discretionary bonuses and feedback credibility. *American Economic Journal: Microeconomics*, 7(1):99–108, 2015.

Giovanni M. Gavetti, Rebecca Henderson, and Simona Giorgi. Kodak and the digital revolution (a). *Harvard Business School Case*, pages 705–448, 2004.

Mark Gertler and Kenneth Rogoff. North-south lending and endogenous domestic capital market inefficiencies. *Journal of Monetary Economics*, 26(2):245–266, 1990.

Robert Gibbons and Kevin J. Murphy. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy*, 100(3):pp. 468–505, 1992.

Armando Gomes. Going public without governance: managerial reputation effects. *The Journal of Finance*, 55(2):615–646, 2000.

Ilan Guedj and David Scharfstein. Organizational scope and investment: Evidence from the drug development strategies and performance of biopharmaceutical firms. Working Paper 10933, National Bureau of Economic Research, 2004.

Michael T Hannan and John Freeman. Structural inertia and organizational change. *American sociological review*, 49(2):149–164, 1984.

Milton Harris and Bengt Holmstrom. A theory of wage dynamics. *The Review of Economic Studies*, 49(3):315–333, 1982.

Rebecca Henderson. Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *The RAND Journal of Economics*, 24(2):248–270, 1993.

Benjamin E Hermalin. Toward an economic theory of leadership: Leading by example. *American Economic Review*, 88(5):1188–1206, 1998.

Bengt Holmström. Agency costs and innovation. *Journal of Economic Behavior and Organization*, 12(3):305–327, 1989.

Peggy Huang, Yiqing Lü, and Moqi Xu. Soft information, innovation, and stock returns. Working paper, London School of Economics, 2015.

Jens Josephson and Joel Shapiro. Credit ratings and structured finance. Working paper, 2014.

Sarah Kaplan and Rebecca Henderson. Inertia and incentives: Bridging organizational economics and organizational theory. *Organization Science*, 16(5):509–521, 2005.

Richard A Lambert. Long-term contracts and moral hazard. *The Bell Journal of Economics*, pages 441–452, 1983.

David F. Larcker, Scott Saslow, and Brian Tayan. How well do corporate directors know senior management? http://www.gsb.stanford.edu/faculty-research/publications/2014-how-well-do-corporate-directors-know-senior-management, 2014.

Josh Lerner. An empirical exploration of a technology race. *The Rand Journal of Economics*, 28(2):228–247, 1997.

Jonathan Levin. Relational incentive contracts. *The American Economic Review*, 93 (3):835–857, 2003.

Laurie K Lewis. Employee perspectives on implementation communication as predictors of perceptions of success and resistance. *Western Journal of Communication*, 70(1):23–46, 2006.

Bentley W. MacLeod. Optimal contracting with subjective evaluation. *The American Economic Review*, 93(1):216–240, 2003.

W Bentley MacLeod and James M Malcomson. Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, 57(2):447–480, 1989.

George J Mailath, Masahiro Okuno-Fujiwara, and Andrew Postlewaite. Belief-based refinements in signalling games. *Journal of Economic Theory*, 60(2):241–276, 1993.

George J Mailath, Volker Nocke, and Andrew Postlewaite. Business strategy, human capital, and managerial incentives. *Journal of Economics and Management Strategy*, 13(4):617–633, 2004.

Gustavo Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.

Henry Mintzberg. Crafting strategy. Technical report, McKinsey&Company, 1988.

Henry Mintzberg and James Waters. Of strategies, deliberate and emergent. *Strategic Management Journal*, 6(3):257–272, 1985.

Kevin J. Murphy and Ján Zábojník. Managerial capital and the market for ceos. Working paper, Queen's University, 2007.

Richard Nelson and Sidney Winter. *An evolutionary theory of economic change*. Harvard University Press, 1982.

Paul Oyer. Why do firms use incentives that have no incentive effects? *The Journal of Finance*, 59(4):1619–1650, 2004.

Joseph Quigley. *Vision: How leaders develop it, share it, and sustain it*. New York: McGraw-Hill, 1993.

Heikki Rantakari. Governing adaptation. *The Review of Economic Studies*, 75(4): 1257–1285, 2008.

Julio Rotemberg and Garth Saloner. Benefits of narrow business strategies. *American Economic Review*, 84(5):1330–49, 1994.

Yuliy Sannikov. A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984, 2008.

Georg Schreyögg and Jörg Sydow. Organizational path dependence: A process view. *Organization Studies*, 32(3):321–335, 2011.

Carl Shapiro and Joseph E Stiglitz. Equilibrium unemployment as a worker discipline device. *The American Economic Review*, 74(3):433–444, 1984.

Kelly Shue and Richard Townsend. Growth through rigidity: An explanation of the rise in CEO pay. Working paper, University of Chicago, 2014.

Joseph E. Stiglitz. Theories of wage rigidity. In James L. Butkiewicz, Kenneth J. Koford, and Jeffrey B. Miller, editors, *Keynes'Economic Legacy: Contemporary Economic Theories*, pages 153–221. Praeger, 1986.

Curtis Taylor. Time-on-the-market as a sign of quality. *The Review of Economic Studies*, 66(3):555–578, 1999.

Katrin Tinn. Technology adoption with exit in imperfectly informed equity markets. *The American Economic Review*, 100(3):925–957, 2010.

Eric Van den Steen. Organizational beliefs and managerial vision. *Journal of Law, Economics, and Organization*, 21(1):256–283, 2005.

Andrew Weiss. Job queues and layoffs in labor markets with flexible wages. *The journal of political economy*, 88(3):526–538, 1980.

Ján Zábojník. Subjective evaluations with performance feedback. *The RAND Journal of Economics*, 45(2):341–369, 2014.

2.9 Appendix 1

Proposition 8. Objective incentive compensation under symmetric information:

- If $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} = 0$, then $b_1^{s,h} = b_1^{s,l}$;
- If $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} > 0$, then $b_1^{s,h} > b_1^{s,l}$;
- If $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} < 0$, then $b_1^{s,h} < b_1^{s,l}$.

Proof The principal's problem:

$$\max_{e} P(\theta,e)(1-b_1(\theta))$$

$$s.t. \quad e^* = e(\theta, b_1(\theta)) \qquad ICa$$

The first order derivative is thus:

$$F.O.C. \quad \frac{\partial p}{\partial e}\frac{de}{db} - \frac{\partial p}{\partial e}b\frac{de}{db} - p(\theta, e(b)) = 0 \tag{2.9.1}$$

Take first order derivative of Equation 2.9.1 w.r.t. θ:

$$\{\frac{\partial^2 p}{\partial e\partial\theta}\frac{de}{db} + \frac{\partial^2 p}{\partial^2 e}(\frac{de}{db})^2\frac{db}{d\theta} + \frac{\partial p}{\partial e}\frac{d^2 e}{d^2 b}\frac{db}{d\theta}\}(1-b) - \frac{\partial p}{\partial e}\frac{de}{db}\frac{db}{d\theta} - \frac{\partial p}{\partial\theta} - \frac{\partial p}{\partial e}\frac{de}{db}\frac{db}{d\theta} = 0$$

Rearrange the equation, I obtain:

$$\frac{db}{d\theta} = -\frac{\frac{\partial^2 p}{\partial e\partial\theta}\frac{de}{db}(1-b) - \frac{\partial p}{\partial\theta}}{\{\frac{\partial^2 p}{\text{The salary is:}\partial^2 e}(\frac{de}{db})^2 + \frac{\partial p}{\partial e}\frac{d^2 e}{d^2 b}\}(1-b) - \frac{\partial p}{\partial e}\frac{de}{db} - \frac{\partial p}{\partial e}\frac{de}{db}} \tag{2.9.2}$$

Assume the second order condition of the principal's problem is satisfied, thus

$$\{\frac{\partial^2 p}{\partial^2 e}(\frac{de}{db})^2 + \frac{\partial p}{\partial e}\frac{d^2 e}{d^2 b}\}(1-b) - \frac{\partial p}{\partial e}\frac{de}{db} - \frac{\partial p}{\partial e}\frac{de}{db} < 0$$

If $\frac{db}{d\theta} = 0$, then

$$\frac{\partial^2 p}{\partial e\partial\theta}\frac{de}{db}(1-b) - \frac{\partial p}{\partial\theta} = 0$$

From Equation 2.9.1, I obtain:

$$p\frac{\partial^2 p}{\partial e\partial\theta} - \frac{\partial p}{\partial e}\frac{\partial p}{\partial\theta} = 0$$

This PDE is equivalent to $\frac{\partial^2 lnp}{\partial e\partial\theta} = 0$.
The solution to the above is $p(\theta, e) = h(\theta)f(e)$.

2.10 Appendix 2: Long-term contracts

The principal who hires an agent of type θ_h at date 0 commits to a menu of contracts $w\{..\} = \{w_{00}^{hh}, w_{10}^{hh}, w_{01}^{hh}, w_{11}^{hh}; w_{00}^{hl}, w_{10}^{hl}, w_{01}^{hl}, w_{11}^{hl}\}$. The principal pays w_{00}^{i} if $(y_1, y_2) = (0,0)$, w_{10}^{i} if $(y_1, y_2) = (1,0)$, w_{01}^{i} if $(y_1, y_2) = (0,1)$ and w_{11}^{i} if $(y_1, y_2) = (1,1)$, $i \in \{hl, hh\}$. Further define a different contract specification $\{b\{.\}, f\{.\}\} = \{b_1^{hh}, b_2^{hh}, b_3^{hh}, f^{hh}; b_1^{hl}, b_2^{hl}, b_3^{hl}, f^{hl}\}$. $b\{.\}$ are bonuses, $b_1^i > 0$, if $y_1 = 1$, $b_2^i > 0$, if $y_2 = 1$, $b_3^i > 0$, if $y_1 = y_2 = 1$, f^i is the fixed compensation.

Lemma Contract $\{w_{00}^{hh}, w_{10}^{hh}, w_{01}^{hh}, w_{11}^{hh}; w_{00}^{hl}, w_{10}^{hl}, w_{01}^{hl}, w_{11}^{hl}\}$ can be equivalently implemented by contract $\{b\{.\}, f\{.\}\} = \{b_1^{hh}, b_2^{hh}, b_3^{hh}, f^{hh}; b_1^{hl}, b_2^{hl}, b_3^{hl}, f^{hl}\}$.

Proof The agent of type θ_h at date 1 and type θ_h at date 2 maximizes effort e_1^h and e_2^{hh}. Given e_1^h and e_2^{hh}, the principal maximizes her profit w.r.t. the above eight parameters.

$$\max_{e_2^{hh}} \ (1-\theta_h e_1^h)(1-\theta_h e_2^{hh}) w_{00}^{hh} + \theta_h e_1^h (1-\theta_h e_2^{hh}) w_{10}^{hh} + \theta_h e_2^{hh}(1-\theta_h e_1^h) w_{01}^{hh} + \theta_h e_2^{hh}\theta_h e_1^h w_{11}^{hh} - \frac{1}{2} e_2^{hh^2}$$

$$\Longleftrightarrow$$

$$\max_{e_2^{hh}} \ w_{00}^{hh} + \theta_h e_1^h (w_{10}^{hh} - w_{00}^{hh}) + \theta_h e_2^{hh}(w_{01}^{hh} - w_{00}^{hh}) + \theta_h e_2^{hh}\theta_h e_1 (w_{11}^{hh} - w_{10}^{hh} - (w_{01}^{hh} - w_{00}^{hh})) - \frac{1}{2} e_2^{hh^2}$$

Set $w_{00}^{hh} = f^h$, $w_{10}^{hh} - w_{00}^{hh} = b_1^{hh}$, $w_{01}^{hh} - w_{00}^{hh} = b_2^{hh}$, and $w_{11}^{hh} - w_{00}^{hh} - (w_{01}^{hh} - w_{00}^{hh}) - (w_{10}^{hh} - w_{00}^{hh}) = b_3^{hh}$. It is easy to see that effort e_2^{hh} is only affected by b_2^{hh} and b_3^{hh}.

$$e_2^{hh} = \theta_h (b_2^{hh} + \theta_h e_1^h b_3^{hh})$$

Likewise, the agent of type θ_h at date 1 and type θ_l at date 2 maximizes effort e_1^h and e_2^{hl}. Set $w_{00}^{hl} = f^{hl}$, $w_{10}^{hl} - w_{00}^{hl} = b_1^{hl}$, $w_{01}^{hl} - w_{00}^{hl} = b_2^{hl}$, and $w_{11}^{hl} - w_{00}^{hl} - (w_{01}^{hl} - w_{00}^{hl}) - (w_{10}^{hl} - w_{00}^{hl}) = b_3^{hl}$. Similar to the above maximization program, the agent produces an effort level:

$$e_2^{hl} = \theta_l (b_2^{hl} + \theta_h e_1^h b_3^{hl})$$

When the agent makes effort e_1^h at date 0, neither the principal nor the agent knows the private information in period 2. As a result, the agent's effort e_1^h will not depend on second period private information and $b_1^{hl} = b_1^{hh}$. I then characterize the optimal effort level of e_1^h.

$$
\begin{aligned}
\max_{e_1^h} \quad & w_{00}^{hh} + \theta_h e_1^h (w_{10}^{hh} - w_{00}^{hh}) - \frac{1}{2} {e_1^h}^2 \\
& + q(\theta_h e_2^{hh} (w_{01}^{hh} - w_{00}^{hh}) + \theta_h e_2^{hh} \theta_h e_1^h (w_{11}^{hh} - w_{10}^{hh} - (w_{01}^{hh} - w_{00}^{hh})) - \frac{1}{2} {e_2^{hh}}^2) \\
& (1-q)(\theta_l e_2^{hl} (w_{01}^{hl} - w_{00}^{hl}) + \theta_l e_2^{hl} \theta_h e_1^h (w_{11}^{hl} - w_{10}^{hl} - (w_{01}^{hl} - w_{00}^{hl})) - \frac{1}{2} {e_2^{hl}}^2)
\end{aligned}
$$

Define $\beta^{hh} = (b_2^{hl} + \theta_h e_1^h b_3^{hh})$ and $\beta^{hl} = (b_2^{hl} + \theta_h e_1^h b_3^{hl})$.

$$
e_1^h = \theta_h b_1^h + q \theta_h^3 b_3^{hl} \beta^{hh} + (1-q) \theta_h \theta_l^2 b_3^{hl} \beta^{hl}
$$

Now let's turn to the principal's maximization program.

$$
\begin{aligned}
\max_{w\{..\}} \quad & q\{-(1-\theta_h e_1^h)(1-\theta_h e_2^{hh}) w_{00}^{hh} + \theta_h e_1^h (1-\theta_h e_2^{hh})(1-w_{10}^{hh}) \\
& + \theta_h e_2^{hh} (1-\theta_h e_1^h)(1-w_{01}^{hh}) + \theta_h e_2^{hh} \theta_h e_1^h (2 - w_{11}^{hh})\} \\
& + (1-q)\{-(1-\theta_h e_1^h)(1-\theta_l e_2^{hl}) w_{00}^{hl} + \theta_h e_1^h (1-\theta_l e_2^{hl})(1-w_{10}^{hl}) \\
& + \theta_l e_2^{hl} (1-\theta_h e_1^h)(1-w_{01}^{hl}) + \theta_l e_2^{hl} \theta_h e_1^h (2 - w_{11}^{hl})\}
\end{aligned}
$$

$$
\begin{aligned}
\Longleftrightarrow \max_{w\{..\}} \quad & q\{\theta_h e_1^h (1 - (w_{10}^{hh} - w_{00}^{hh})) + \theta_h e_2^{hh} (1 - (w_{01}^{hh} - w_{00}^{hh})) \\
& - \theta_h e_1^h \theta_h e_2^{hh} (w_{11}^{hh} - w_{00}^{hh} - (w_{01}^{hh} - w_{00}^{hh}) - (w_{10}^{hh} - w_{00}^{hh})) - w_{00}^{hh}\} \\
& + (1-q)\{\theta_h e_1^h (1 - (w_{10}^{hl} - w_{00}^{hl})) + \theta_l e_2^{hl} (1 - (w_{01}^{hl} - w_{00}^{hl})) \\
& - \theta_h e_1 \theta_l e_2^{hl} (w_{11}^{hl} - w_{00}^{hl} - (w_{01}^{hl} - w_{00}^{hl}) - (w_{10}^{hl} - w_{00}^{hl})) - w_{00}^{hl}\}
\end{aligned}
$$

It is obvious to see from the above system, variables $w\{..\}$ which can be transferred to $\{b\{.\}, f\{.\}\}$ in the agent's maximization problem can also be transferred to the same set of variables in the principal's maximization problem.

$$
\begin{aligned}
\max_{\{b\{.\},f\{.\}\}} \quad & \theta_h e_1^h (1 - b_1^h) + q\{\theta_h e_2^{hh} (1 - b_2^{hh}) - \theta_h e_1^h \theta_h e_2^{hh} b_3^{hh} - f^h\} \\
& + (1-q)\{\theta_l e_2^{hl} (1 - b_2^{hl})) - \theta_h e_1^h \theta_l e_2^{hl} b_3^{hl} - f^l\}
\end{aligned}
$$

Q.E.D.

1. **No cross-pledging**

Lemma In the equivalent contract specification $\{b\{.\}, f\{.\}\}$, if b_3^{hh}, b_3^{hl} and b_3^{ll} are set to 0, to induce effort, the following components are greater than zero: $b_1^h > 0$, $b_2^{hh} > 0$, $b_2^{hl} > 0$, $b_1^{ll} > 0$ and $b_2^{ll} > 0$. In order to satisfy the limited liability, $f^1 \geq 0$ and $f^h \geq 0$.

Proof Composite bonuses β^{hh} and β^{hl} defined in the above lemma are important auxiliary variables.

1. Assume the compensation paid to the agent of type θ_l at date 0 is $\{b_1^{ll}, b_2^{ll}, b_3^{ll}\}$. $b_1^{ll} > 0$, if $y_1 = 1$. $b_2^{ll} > 0$, if $y_2 = 1$. $b_3^{ll} > 0$, if $y_1 = y_2 = 1$. The principal does not need to pay salary to this agent. I first prove that in the separating equilibrium, if b_3^{ll} is set to zero for the agent of type θ_l at date 0, $b_1^{ll} > 0$, $b_2^{ll} > 0$.

   Define $\beta^{ll} = (b_2^{ll} + \theta_l e_1^l b_3^{ll})$, so $b_2^{ll} = \beta^{ll} - \theta_l e_1^l b_3^{ll}$. Following the above proof, it is easy to prove that $e_2^{ll} = \beta^{ll}\theta_l$ and $e_1^l = \theta_l b_1^{ll} + \theta_l^3 \beta^{ll} b_3^{ll}$. The principal's maximization program is thus:

$$\begin{aligned}
&\max_{\{b_1^{ll}, b_2^{ll}, b_3^{ll}\}} \quad \theta_l e_1^l (1 - b_1^{ll}) + \theta_l e_2^{ll}(1 - b_2^{ll}) - \theta_l e_1^l \theta_l e_2^{ll} b_3^{ll} \\
\Leftrightarrow &\max_{\{b_1^{ll}, \beta^{ll}, b_3^{ll}\}} \quad \theta_l(\theta_l b_1^{ll} + \theta_l^3 \beta^{ll} b_3^{ll})(1 - b_1^{ll}) + \theta_l^2 \beta^{ll}(1 - (\beta^{ll} - \theta_l e_1^l b_3^{ll})) - \theta_l^3 \beta^{ll} e_1^l b_3^{ll} \\
\Leftrightarrow &\max_{\{b_1^{ll}, \beta^{ll}, b_3^{ll}\}} \quad \theta_l^2 b_1^{ll}(1 - b_1^{ll}) + \theta_l^2 \beta^{ll}(1 - \beta^{ll}) + \theta_l^4 \beta^{ll}(1 - b_1^{ll}) b_3^{ll}
\end{aligned}$$

   b_3^{ll} only enters into the maximization program through term $\theta_l^4 \beta^{ll}(1 - b_1^{ll}) b_3^{ll}$. It can be easily verify that $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.

2. The second step is to prove that $b_1^h > 0$, $b_2^{hh} > 0$. A principal who hires an agent of type θ_h at date 0 maximizes the profit subject to two truth-telling constraints. The principal of type h solves the following maximization problem P^h:

$$\begin{aligned}
\max_{\{b\{.\}, f\{.\}\}} \quad &\theta_h e_1^h (1 - b_1^h) + q\{\theta_h e_2^{hh}(1 - b_2^{hh}) - \theta_h e_1^h \theta_h e_2^{hh} b_3^{hh} - f^h\} \\
&+ (1 - q)\{\theta_l e_2^{hl}(1 - b_2^{hl})) - \theta_h e_1^h \theta_l e_2^{hl} b_3^{hh} - f^l\}
\end{aligned}$$

$$s.t. \quad \theta_l e_1^h (1 - b_1^h) + \theta_l e_2^{hl}(1 - b_2^{hl}) - \theta_l e_2^{hl} \theta_l e_1^h b_3^{hl} - f^l \leq \frac{1}{4}\theta_l^2 + \frac{1}{4}\theta_l^2 \tag{2.10.1}$$

$$\theta_l e_2^{hh}(1 - b_2^{hh}) - \theta_l e_2^{hh} \theta_h e_1^h b_3^{hh} - f^h \leq \theta_l e_2^{hl}(1 - b_2^{hl}) - \theta_l e_2^{hl} \theta_h e_1^h b_3^{hl} - f^l \tag{2.10.2}$$

   In addition, the principal does not want to renegotiate the contract at the beginning of the second period when new private information arrives. I'll come back to this point later. Substituting b_2^{hh} and b_2^{hl} with β^{hh} and β^{hl}, the above two constraints are equivalent to:

$$f^l = \theta_l e_1^h(1-b_1^h) + \theta_l^2\beta^{hl}(1-\beta^{hl}) + \theta_l^2(\theta_h-\theta_l)e_1^h\beta^{hl}b_3^{hl} - \frac{1}{2}\theta_l^2$$
$$f^h = \theta_l\theta_h\beta^{hh}(1-\beta^{hh}) - \theta_l^2\beta^{hl}(1-\beta^{hl}) + f^l$$

The principal's maximization program is equivalent to the following program P'^h:

$$\max_{\{b\{.\},\beta\{.\}\}} \quad (\theta_h-\theta_l)\{e_1^h(1-b_1^h) - \theta_l^2 e_1^h\beta_l b_3^{hl}\} + q(\theta_h-\theta_l)\theta_h\beta^{hh}(1-\beta^{hh}) + \theta_l^2$$

From the previous lemma, we know that:

$$e_1^h = \theta_h b_1^h + q\theta_h^3 b_3^{hh}\beta^{hh} + (1-q)\theta_h\theta_l^2 b_3^{hl}\beta^{hl}$$

It's easy to see that b_3^{hh} only enters into the maximization program through term $(\theta_h-\theta_l)(1-b_1^h)q\theta_h^3\beta^{hh}b_3^{hh}$.

3. The last step is to prove that $b_2^{hl} > 0$. The principal who hires a deteriorating type of agent in the second period may want to renegotiate the contract. Assume if renegotiation happens, the renegotiated contract specification is given by $\{b_2'^{hl}, b_3'^{hl}, f'^l\}$. Define $\beta'^{hl} = (b_2'^{hl} + \theta_h e_1^h b_3'^{hl})$, thus effort $e_2'^{hl} = \theta_l\beta'^{hl}$. A renegotiation-proof contract must satisfy the following maximization program P^{hl}:

$$\begin{aligned}\{b_2^{hl}, b_3^{hl}, f^l\} \quad &\in arg \max_{\{b_2'^{hl}, b_3'^{hl}, f'^l\}} \quad \theta_l e_2'^{hl}(1-b_2'^{hl}) - \theta_h e_1^h\theta_l e_2'^{hl}b_3'^{hl} - f'^l \\ s.t.\quad &\theta_l e_2'^{hl}(b_2'^{hl} + \theta_h e_1^h b_3'^{hl}) + f'^l - \frac{1}{2}e_2'^{hl2} \geq \theta_l e_2^{hl}(b_2^{hl} + \theta_h e_1^h b_3^{hl}) + f^l - \frac{1}{2}e_2^{hl2}\end{aligned}$$

Assume $u = \theta_l e_{2l}(b_{2l} + \theta_h e_1 b_{3l}) + f_l - \frac{1}{2}e_{2l}^2$, the above program is equivalent to the following program P'^{hl}:

$$\begin{aligned}\{\beta^{hl}, b_3^{hl}, f^l\} \quad &\in arg \max_{\{\beta'^{hl}, b_3'^{hl}, f'^l\}} \quad \theta_l^2\beta'^{hl}(1-\beta'^{hl}) - f'^l \\ s.t.\quad &\frac{1}{2}\theta_l^2\beta'^{hl2} - f'^l \geq u\end{aligned}$$

Q.E.D.

Lemma Without the cross-pledging effect, pooling equilibrium does not survive the Intuitive Criterion.

Proof Proof by contradiction. Assume $b_2^{hl} = b_2^{hh} = b$ and $f^h = f^l = 0$. Further assume $\bar{\theta} = q\theta_h + (1-q)\theta_l$. The agent maximizes her own utility and chooses the optimal effort level:

$$\begin{aligned} e_2^* &\in arg \max_{e_2} q\theta_h e_2 b_2^{hh} + (1-q)\theta_l e_2 b_2^{hl} - \frac{1}{2}e_2^2 \\ e_2^* &= \bar{\theta} b \end{aligned}$$

The principal of type hh obtains profit $\pi = \theta_h \bar{\theta} b(1-b)$. If she deviates by paying an additional salary $f^h = \frac{1}{2}(\theta_h - \bar{\theta})\theta_h b(1-b)$, then she could obtain profit $\pi' = \theta_h^2 b(1-b)$. And $\pi' - f^h > \pi$.

Q.E.D.

Proposition Low separating profit ($\theta_h < 2\theta_l$). Define $\pi_{2l} = \frac{1}{4}\theta_l^2(2 - \frac{\theta_h}{\theta_l})$.

- The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_1 = \frac{1}{2}$; $b_2^{hh} = \frac{1}{2}$, $f^h = \frac{1}{4}\theta_l\theta_h - \pi_{2l}$; $b_2^{hl} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1})$, $f^l = 0$.
- The principal will offer two short-term contracts to type θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.

Proposition High separating profit ($\theta_h \geq 2\theta_l$).

- The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_1 = \frac{1}{2}$; $b_2^{hh} = \frac{1}{2}$, $f^h = \frac{1}{2}\theta_l(\theta_h - \theta_l)$; $b_2^{hl} = 1$, $f^l = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)$.
- The principal will offer two short-term contracts to type θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.

Proof In order to obtain the optimal contract, the principal of type h at date 0 needs to solve the program P'^h. In addition, the menu of contracts needs to be renegotiation-proof. A renegotiation-proof contract must satisfy the program P^{hl} and the following maximization program P^{hh} for a principal of type hh:

$$\{\beta^{hh}, b_3^{hh}, f^h\} \in arg \max_{\{b_2'^{hh}, b_3'^{hh}, f'^h\}} \theta_h e_2'^{hh}(1 - b_2'^{hh}) - \theta_h e_1^h \theta_h e_2'^{hh} b_3'^{hh} - f'^h$$

$$s.t. \quad \theta_l e_2'^{hh}(1 - b_2'^{hh}) - \theta_l e_2'^{hh}\theta_h e_1'^h b_3'^{hh} - f'^h \leq \theta_l e_2^{hl}(1 - b_2^{hl}) - \theta_l e_2^{hl}\theta_h e_1^h b_3^{hl} - f^l \tag{2.10.3}$$

$$\theta_h e_2'^{hh}(b_2'^{hh} + \theta_h e_1^h b_3'^{hh}) + f'^h - \frac{1}{2}(e_2'^{hh})^2 \geq \theta_h e_2^{hh}(b_2^{hh} + \theta_h e_1^h b_3^{hh}) + f^h - \frac{1}{2}(e_2^{hh})^2 \tag{2.10.4}$$

Assume $u = \theta_h e_2^{hh}(b_2^{hh} + \theta_h e_1^h b_3^{hh}) + f^h - \frac{1}{2}(e_2^{hh})^2$, and $\pi^{hl} = \theta_l e_1^h(1 - b_2^{hl}) - \theta_l e_2^{hl}\theta_h e_1^h b_3^{hl} - f^l$. The above program is equivalent to the following:

$$
\begin{aligned}
\{\beta^{hh}, b_3^{hh}, f^h\} \quad &\in arg \max_{\{\beta'^{hh}, b_3'^{hh}, f'^h\}} \theta_h^2 \beta'^{hh}(1-\beta'^{hh}) - f'^h \\
s.t. \quad & \frac{1}{2}\theta_h^2 \beta_h'^2 - f'^h \geq u
\end{aligned}
$$

Remember in this section, we examine a contract without cross-pledging.

1. If the Constraint 2.10.3 is satisfied, the Constraint 2.10.4 will be satisfied too. The argument is offered below. Assume that $\{b_2^{hh}, f^h\}$ is the contract that satisfies the following program. I'll prove that the principal will not want to renegotiate this contract as long as the truth-telling constraint is satisfied.

$$
\begin{aligned}
\{b_2^{hh}, f^h\} \quad &\in arg \max_{\{b_2'^{hh}, f'^h\}} \theta_h e_2'^{hh}(1-b_2'^{hh}) - f'^h \\
s.t. \quad & \theta_l e_1'^h(1-b_2'^{hh}) - f'^h \leq \theta_l e_1^h(1-b_2^{hl}) - f^l
\end{aligned}
$$

Assume $\{b_2'^{hh}, f'^h\}$ is the renegotiated contract, from which the principal obtains a higher profit than from the old contract $\{b_2^{hh}, f^h\}$. As a result, the following inequalities are met:

$$
\begin{aligned}
&\theta_h^2 b_2'^{hh}(1-b_2'^{hh}) - f'^h > \theta_h^2 b_2^{hh}(1-b_2^{hh}) - f^h \\
\Leftrightarrow & f^h - f'^h > \theta_h^2 b_2^{hh}(1-b_2^{hh}) - \theta_h^2 b_2'^{hh}(1-b_2'^{hh}) \\
\Leftrightarrow & f^h - f'^h > \theta_l\theta_h b_2^{hh}(1-b_2^{hh}) - \theta_l\theta_h b_2'^{hh}(1-b_2'^{hh}) \\
\Leftrightarrow & \theta_l\theta_h b_2'^{hh}(1-b_2'^{hh}) - f'^h > \theta_l\theta_h b_2^{hh}(1-b_2^{hh}) - f^h \\
\Leftrightarrow & \theta_l\theta_h b_2'^{hh}(1-b_2'^{hh}) - f'^h > \pi^{hl}
\end{aligned}
$$

The above inequality conflicts with the principal's truth-reporting constraint 2.10.3.

2. The second step is to show that the principal of type hl may or may not renegotiate the contract to zero salary paid to the agent, depending on the separating profit. Program P'^h can be simplified to the following without cross-pledging:

$$
\max_{\{b\{.\}} \ (\theta_h - \theta_l)\theta_h b_1^h(1-b_1^h) + q(\theta_h - \theta_l)\theta_h b_2^{hh}(1-b_2^{hh}) + \theta_l^2
$$

It is easy to see that b_1^h does not depend on the renegotiation as it is already sunk. So the principal will set $b_1^h = \frac{1}{2}$. It is easy to verify that without cross-pledging, $b_1^l = \frac{1}{2}$.

Program P'^{hl} can be simplified to the following program without cross-pledging:

$$\begin{aligned}\{b_2^{hl}, f^l\} \quad &\in arg \max_{b_2'^{hl}, f'^l} \theta_l^2 b_2'^{hl}(1 - b_2'^{hl}) - f'^l \\ s.t. \quad &\frac{1}{2}\theta_l^2 (b_2'^{hl})^2 - f'^l \geq u\end{aligned}$$

If $f^l > 0$, we could easily verify that $b_2^{hl} = 1$ by substituting the constraint into the objective function. This only happens if $\theta_h \geq \theta_l$. From Constraint 2.10.1, one could verify that if $b_2^{hl} = 1$, $f^l = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)$.

If $\theta_h < \theta_l$. The principal will not set $f^l > 0$, as it is not renegotiation-proof. The principal will always substitute it with more bonus. So the bonus will be set at the highest possible level with $f^l = 0$. According to Constraint 2.10.1, one could find that $b_{2l} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1})$.

Q.E.D.

2. **With cross-pledging**

$\theta_l^2 \geq 1/2$ **Lemma** Assume $\theta_l^2 \geq 1/2$ and superscript s denotes the bonus if information is symmetric:

- Any contracts $\{b\{.\}, f\{.\}\}$ for type hh, hl and ll under symmetric information can be replicated by contracts that only consist of b_3^{hh}, b_3^{hl} or b_3^{ll} respectively.
- $b_1^{hh} = b_2^{hh} = 0$, $b_3^{hh} = \frac{1}{\theta_h^2}$; $b_1^{hl} = b_2^{hl} = 0$, $b_3^{ll} = \frac{1}{\theta_l\theta_h}$; $b_1^{ll} = b_2^{ll} = 0$, $b_1^{ll} = \frac{1}{\theta_l^2}$.

Proof 1. For type ll, under contract $\{b\{.\}, f\{.\}\}$, the agent could obtain utility level based on Lemma 2.10:

$$\begin{aligned}&\theta_l b_1^{ll} e_1^l + \theta_l e_2^{ll} b_2^{ll} + \theta_l e_1^l \theta_l e_2^{ll} b_3^{ll} - \frac{1}{2}(e_1^l)^2 - \frac{1}{2}(e_2^{ll})^2 \\ \Leftrightarrow &e_1^l(\theta_l b_1^{ll} - \frac{1}{2}(e_1^l)) + \frac{1}{2}\theta_l^2 \beta^{ll} \\ \Leftrightarrow &\frac{1}{2}(\theta_l b_1^l + \theta_l^3 b_3^{ll} \beta^{ll})(\theta_l b_1^l - \theta_l^3 b_3^{ll} \beta^{ll}) + \frac{1}{2}\theta_l^2 \beta^{ll} \\ \Leftrightarrow &\frac{1}{2}(\theta_l^2 (b_1^l)^2 - \theta_l^6 (b_3^{ll})^2 (\beta^{ll})^2) + \frac{1}{2}\theta_l^2 \beta^{ll}\end{aligned}$$

One could find a contract which consists of only b_3^{ll} to incentivise the agent. $e_2^{ll} = \theta_l^2 e_1^l b_3^{ll}$. As a result, the agent's utility

$$\begin{aligned}&\frac{1}{2}\theta_l^4 (e_1^l)^2 (b_3^{ll})^2 - \frac{1}{2}(e_1^l)^2 \\ \Leftrightarrow &\frac{1}{2}(e_1^l)^2(\theta_l^4 (b_3^{ll})^2 - 1)\end{aligned}$$

b_3^{lls} is set at such a level that the following equation is satisfied:

$$\frac{1}{2}(e_1^{ls})^2(\theta_l^4(b_3^{lls})^2 - 1) = \frac{1}{2}(\theta_l^2(b_1^l)^2 - \theta_l^6(b_3^{ll})^2(\beta^{ll})^2) + \frac{1}{2}\theta_l^2\beta^{ll}$$

As a result, $b_3^{lls} = \frac{1}{\theta_l^2}$, $e_1^{ls} = e_2^{lls} = 1$. One could easily verify that the profit for the principal is $2\theta_l - 1$. If the principal does not use b_3 at all, the profit is $\frac{1}{2}\theta_l^2$, which is strictly smaller than $2\theta_l - 1$.

2. Following the same argument as in the previous step, for type θ_h at date 0, the principal only uses b_3^{hls} and b_3^{hhs}. The agent's expected utility is as follows:

$$\frac{1}{2}(e_1^h)^2\{q(\theta_h^4(b_3^{hh})^2 - 1) + (1-q)(\theta_h^2\theta_l^2(b_3^{hl})^2 - 1)\}$$

The minimum compensation paid to the agent in order to induce effort level 1 is by setting $q(\theta_h^4(b_3^{hh})^2 - 1) + (1-q)(\theta_h^2\theta_l^2(b_3^{hl})^2 - 1) = 0$. The principal's maximization problem is:

$$\max_{\{b_3^{hl}, b_3^{hh}\}} q(\theta_h e_1^h + \theta_h e_2^{hh} - \theta_h e_1^h \theta_h e_2^{hh} b_3^{hh}) + (1-q)(\theta_h e_1^h + \theta_l e_2^{hl} - \theta_h e_1^h \theta_l e_2^{hl} b_3^{hl})$$
$$s.t.\ \ q(\theta_h^4(b_3^{hh})^2 - 1) + (1-q)(\theta_h^2\theta_l^2(b_3^{hl})^2 - 1) = 0$$

Because $e_2^{hl} = \theta_l\theta_h e_1^h b_3^{hl}$ and $e_2^{hh} = \theta_h^2 e_1^h b_3^{hh}$, it is equivalent to the following program:

$$\max_{\{e_2^{hh}, e_2^{hl}\}} q\theta_h e_2^{hh} + (1-q)\theta_l e_2^{hl}$$
$$s.t.\ \ q(e_2^{hh})^2 + (1-q)(e_2^{hl})^2 = 1$$

Because $e_2^{hh}, e_2^{hl} \leq 1$, the principal will set$e_2^{hh}, e_2^{hl} = 1$ to maximize the profit. Thus $b_3^{hl} = \frac{1}{\theta_h\theta_l}$, and $b_3^{hh} = \frac{1}{\theta_h^2}$. $\beta^{hh} = \frac{1}{\theta_h}$, $\beta^{hl} = \frac{1}{\theta_l}$ and $\beta^{ll} = \frac{1}{\theta_l}$.

When the two parties receive new information in period two, the agent will not want to renegotiate. The agent's utility of type hh is $\theta_h e_2^{hh}\beta^{hh} - \frac{1}{2}(e_2^{hh})^2 = \frac{1}{2}\theta_h^2(\beta^{hh})^2$. Under the contract analyzed above, the agent's utility is $\frac{1}{2}$. If the principal wants to renegotiate, she will set for instance, $\beta^{hh} = \frac{1}{2}$. The agent's utility is thus $\frac{1}{8}\theta_h^2$. The agent thus will not want to renegotiate.

Q.E.D.

Lemma With the cross-pledging effect, pooling equilibrium does not survive the Intuitive Criterion.

Proof Proof by contradiction. Assume $b_2^{hl} = b_2^{hh} = b_2$, $b_3^{hl} = b_3^{hh} = b_3$ and $f^h = f^l = 0$. Further assume $\bar{\theta} = q\theta_h + (1-q)\theta_l$. The agent maximizes her own utility and chooses the optimal effort level:

$$\begin{aligned} e_2^* &\in arg\max_{e_2} q(\theta_h e_1 \theta_h e_2 b_3 + \theta_h e_2 b_2) + (1-q)(\theta_h e_1 \theta_l e_2 b_3 + \theta_l e_2 b_2) - \frac{1}{2} e_2^2 \\ e_2^* &= \theta_h \bar{\theta} e_1 b_3 + \bar{\theta} b_2 \end{aligned}$$

The principal of type hh obtains profit $\pi = \theta_h e_2^*(1-b_2-\theta_h e_1 b_3)$. If she deviates by paying an additional salary $f^h = \frac{1}{2}(\theta_h - \bar{\theta})(\theta_h e_1 b_3 + b_2)(1-b_2-\theta_h e_1 b_3)$, then she could obtain profit $\pi' = \theta_h e_2'(1-b_2-\theta_h e_1 b_3)$, in which $e_2' = \theta_h^2 e_1 b_3 + \theta_h b_2$. And $\pi' - f^h > \pi$.

Q.E.D.

Lemma A contract which contains b_1^h and b_2^{hh} offered by a high-type principal can be replicated by a contract which does not contain b_1^h and b_2^{hh}.

Proof

1. I first show that a contract which contains b_1^h offered by a high-type principal can be replicated by a contract which does not contain b_1^h.

   The principal's maximization program P'^h under asymmetric information can be rewritten as follows:

   $$\max_{\{b\{.\},\beta\{.\}\}} \quad (\theta_h - \theta_l)(\theta_h b_1^h + q\theta_h^3 b_3^{hh}\beta^h + (1-q)\theta_h\theta_l^2 b_3^{hl}\beta^l)(1 - b_1^h - \theta_l^2 b_3^{ll}\beta^l)$$

   Set $\theta_h^2 b_3'^{hh}\beta'^h = b_1^h + \theta_h^2 b_3^{hh}\beta^h$, and $\theta_l^2 b_3'^{hl}\beta'^l = b_1^h + \theta_l^2 b_3^{hl}\beta^l$. With $\{b_3'^{hh}, \beta'^h, b_3'^{hl}, \beta'^l\}$, the firm achieves the same profit. The agent will exert the same amount of effort e_1^h, but effort e_2 will increase due to an increase in β if b_3^{hh} and b_3^{hh} are kept constant. The principal could obtain the same profit using contract $\{b_3^{hh}, \beta'^h, b_3^{hl}, \beta'^l\}$ which induces a higher level of effort. This means that the principal could pay the agent less (less rent to the agent) in order to obtain a higher profit.

   The principal could use higher b_2 to keep the first period effort because of the cross-pledging effect while increasing the second period effort.

2. I then show a contract which contains b_2^{hh} offered by a high-type principal can be replicated by a contract which does not contain b_2^{hh}.

   b_3^{hh} enters into the maximization program through the term $(\theta_h - \theta_l)q\theta_h^3 b_3^{hh}\beta^h(1 - b_1^h - \theta_l^2 b_3^{ll}\beta^l)$. One could show that the principal will always want to use b_3^{hh} to substitute b_2^{hh}. Assume that $b_2'^{hh} = b_2^{hh} - \epsilon$, and $\theta_h e_1^h b_3'^{hh} = \theta_h e_1^h b_3'^{hh} + \epsilon$. The second equation implies $b_3'^{hh} > b_3^{hh}$, and $\beta'^h = \beta^h$ if e_1^h is not affected. However, $e_1^h = \theta_h b_1 + \theta_h^3\beta^h b_3^{hh} = \theta_h^3\beta^h b_3^{hh}$. When b_3^{hh} goes up to $b_3'^{hh}$, $e_1'^h > e_1'^h$ and $\beta'^h > \beta^h$, leading to a higher profit.

   The principal uses b_3^{hh} instead of b_2^{hh} as the former also induces higher first period effort because of the cross-pledging effect.

Q.E.D.

Proposition Assume $\theta_l^2 \geq 1/2$. Under information asymmetry,

- The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_3^{hh} = \sqrt{\frac{1-(1-q)\theta_l^2}{q\theta_h^4}}$, $b_2^{hl} = \frac{3q\theta_l^2 - 2q^2\theta_l^2 - q\theta_h + q^2\theta_h - \theta_l^2 + 1}{2q(1-q)\theta_l^2}$, and $b_3^{hl} = \frac{1}{2\theta_l^2} - \frac{\theta_h^2\beta^{hh}b_3^{hh}}{2(1-q)\theta_l^2}$. $f^h = f^l + \theta_l\theta_h^2 b_3^{hh}(1-\theta_h b_3^{hh})$, and $f^l = 1 - \theta_l + (\theta_h - \theta_l)\theta_l^2 b_3^{hl}$.
- The principal will offer a long-term contract to type θ_l at date 0, $b_3^{ll} = \frac{1}{\theta_l^2}$.

Proof The principal's maximization program P'^h under asymmetric information can be rewritten as follows:

$$\max_{\{b\{.\},\beta\{.\}\}} \quad (\theta_h - \theta_l)(\theta_h b_1^h + q\theta_h^3 b_3^{hh}\beta^{hh} + (1-q)\theta_h\theta_l^2 b_3^{hl}\beta^{hl})(1 - b_1^h - \theta_l^2 b_3^{ll}\beta^{hl})$$

Take first order derivative w.r.t. b_3^{hl}, one could find that:

$$b_3^{hl} = \frac{1}{2\theta_l^2}(1 - b_1^h) - \frac{b_1^h + \theta_h^2\beta^{hh}b_3^{hh}}{2(1-q)\theta_l^2}$$

In the second stage renegotiation for a principal of type hl, program P'^{hl} is as follows:

$$\begin{aligned} \{\beta^{hl}, b_3^{hl}, f^l\} \quad &\in arg \max_{\{\beta'^{hl}, b_3'^{hl}, f'^l\}} \theta_l^2\beta'^{hl}(1-\beta'^{hl}) - f'^l \\ s.t. \quad &\frac{1}{2}\theta_l^2\beta'^{hl2} - f'^l \geq u \end{aligned}$$

If $f^l > 0$, then $\beta^{hl} = 1$. The principal of type hh will not want to renegotiate the contract as proved in Proposition 2.10. As a result,

$$b_3^{hl} = \frac{1}{2\theta_l^2} - \frac{\theta_h^2\beta^{hh}b_3^{hh}}{2(1-q)\theta_l^2}$$

The agent hired by principal of type h at date 0 chooses effort level e_1^h:

$$\begin{aligned} e_1^h \quad &\in arg \max_{e_1^h} q(\theta_h e_1^h \theta_h e_2^{hh} b_3^{hh} - \frac{1}{2}(e_2^{hh})^2) + (1-q)(\theta_l e_2^{hl}\beta^{hl} - \frac{1}{2}(e_2^{hl})^2) - \frac{1}{2}(e_1^h)^2 \\ \Leftrightarrow &\in arg \max_{e_1^h} \frac{1}{2}q(\theta_h^4(b_3^{hh})^2 - 1)(e_1^h)^2 + \frac{1}{2}(1-q)(\theta_l^2(\beta^{hl})^2 - 1)(e_1^h)^2 \end{aligned}$$

To induce the agent to make an effort $e_1^h = 1$, the principal sets b_3^{hh} at:

$$\begin{aligned} b_3^{hh} &= \sqrt{\frac{1-(1-q)\theta_l^2(\beta^{hl})^2}{q\theta_h^4}} \\ &= \sqrt{\frac{1-(1-q)\theta_l^2}{q\theta_h^4}} \end{aligned}$$

It can be easily verified that $b_3^{hh} > \frac{1}{\theta_h^2}$. As a result, $b_3^{hl} < 0$. The principal sets b_2^{hl} at:

$$\begin{aligned} b_2^{hl} &= 1 - \theta_h b_3^{hl} \\ &= 1 - \theta_h (\frac{1}{2\theta_l^2} - \frac{\theta_h^2 \beta^{hh} (b_3^{hh})^2}{2(1-q)\theta_l^2}) \\ &= \frac{q\theta_h - q^2\theta_h + \theta_l^2 - q\theta_l^2 - 1}{2q(1-q)\theta_l^2} \\ &= \frac{3q\theta_l^2 - 2q^2\theta_l^2 - q\theta_h + q^2\theta_h - \theta_l^2 + 1}{2q(1-q)\theta_l^2} \end{aligned}$$

Because $b_3^{hl} < 0$, if the principal has no limited liability, then $b_2^{hl} > 1$.

$$\begin{aligned} f^l &= \theta_l + \theta_l^2(\theta_h - \theta_l) b_3^{hl} - (2\theta_l - 1) \\ &= 1 - \theta_l + (\theta_h - \theta_l)\theta_l^2 b_3^{hl} \\ f^h &= f^l + \theta_l - \theta_l\theta_h b_3^{hh} - \theta_l(1 - b_2^{hl}) + \theta_l\theta_h e_2^{hl} b_3^{hl} \\ &= f^l + \theta_l\theta_h^2 b_3^{hh}(1 - \theta_h b_3^{hh}) \end{aligned}$$

b_3^{hh} is decreasing in q. f^l is increasing in $\theta_h - \theta_l$.

Q.E.D.

Corollary Assume $\theta_l^2 \geq 1/2$. Under information asymmetry, The principal will commit to a menu of contracts to type θ_h at date 0 which specifies the following: $b_3^{hh} > \frac{1}{\theta_h^2}$, $b_2^{hl} > 1$, and $b_3^{hl} < 0$. $f^h > 0$, $f^l > 0$.

Proof Please refer to the proof above.

2.11 Appendix 3: Continuous Type

The Revelation Mechanism

Proof Consider a separating PBE. $x = X(\theta)$ is the contract the principal offers if she receives private information θ. $\hat{\alpha} = \mu(x)$ is the agent's belief, in a separating equilibrium, $\hat{\alpha} = \theta$. $e = e(\hat{\alpha}, x)$ is the effort the agent makes in the subgame. The principal's profit function is $V(\theta, \hat{\alpha}, e(\hat{\alpha}, x))$.

Assume $x = arg \max_x V(\theta, \mu(x), e(\mu(x), x))$ for type θ. And $\mu(x(\theta)) = g(\theta)$ For any $x' \neq x$,

$$\begin{aligned} V(\theta, \mu(x), e(\hat{\alpha}, x)) &\geq V(\theta, \mu(x'), e(\mu(x'), x')) \\ &= V(\theta, \mu(x'(\theta)), e(\mu(x'(\theta)), x'(\theta))) \\ &= V(\theta, \mu(x(\theta')), e(\mu(x(\theta')), x(\theta'))) \\ &= V(\theta, g(\theta'), e(g(\theta'), x(\theta'))) \end{aligned}$$

The above proof shows that for a separating PBE, one could find an equivalent direct mechanism in which the principal truthfully announce to a third party the private information, and the third party implements the contract for the principal.

At the beginning of the first period, the principal and the agent only know that the agent's ability is drawn from an interval $[\underline{\theta}, \overline{\theta}]$ according to a distribution function $H(\theta)$ with density $h(\theta)$. $H(\theta)$ is twice differentiable at each $\theta \in [\underline{\theta}, \overline{\theta}]$.[19] At date $t = 0$, the principal has a private signal η which belongs to the interval $[\underline{\theta}, \overline{\theta}]$.

Symmetric Information

In this benchmark case, the board and the CEO are both informed of the productivity θ at the end of the first period. Because the CEO knows her own type, the board has no incentive to use compensation as a means of providing feedback. Consequently, paying $f_1(m)$ is not necessary, and only the second period bonus $b_1(m, y)$ is needed.

Because the agent is protected by limited liability, $b_1(m, 0) = 0$ if $y = 0$. Define $b_1(m, 1) = b_1(m)$ if $y = 1$[20]. In the benchmark case, the message m is exactly the signal θ that both of the two parties receive, that is, $m = \theta$. As a result, $b_1(m, 0) = b_1(\theta, 0) = 0$, and $b_1(m) = b_1(\theta)$.

I first analyze the CEO's problem. Given $b_1(\theta)$, the CEO chooses effort e to maximize her utility:

$$\max_e P(\theta, e) b_1(\theta) - \psi(e)$$

[19] The result of separating equilibrium does not depend on the exact form of the distribution function $H(\theta)$.

[20] In the case of $y = 1$, the board's limited liability constraint will be satisfied as well, since the board could always improve the firm's profit from negative to at least zero by reducing the $b_1(m) > 1$ to $b_1(m) = 1$.

From the first order condition, the optimal level of effort as a function of θ and $b_1(\theta)$ is $e^* = e(\theta, b_1(\theta))$. The effort level depends on the CEO's productivity θ and the second period bonus b_1. Given the optimal effort level of the CEO, the board's problem P_0 is as follows[21]:

$$\begin{aligned} &\max_{b_1(\cdot)} P(\theta,e)(1-b_1(\theta)) && \\ s.t. \quad & e^* = e(\theta, b_1(\theta)) && IC_a \\ & P(\theta,e)b_1 - \psi(e) \geq 0 \qquad \text{for all } \theta && IR_a \end{aligned}$$

Constraint IC_a is the CEO's incentive constraint obtained from her own maximization program. Constraint IR_a is the participation constraint of the CEO. Using standard arguments of moral hazard (Laffont and Martimort, 2002), the limited liability constraint if the output is low ($y = 0$) will be binding. The objective function of Program P_0 already takes this into account.

The optimal level of second period compensation under symmetric information is denoted as $b_1^s(\theta)$. It is given by the first order condition of the principal's maximization program:

$$\frac{\partial P(\theta,e)}{\partial e}\frac{\partial e(\theta, b_1^s(\theta))}{\partial b_1^s(\theta)} = P(\theta,e) + \frac{\partial P(\theta,e)}{\partial e}\frac{\partial e(\theta, b_1^s(\theta))}{\partial b_1^s(\theta)} b_1^s(\theta) \tag{2.11.1}$$

The LHS of Equation 2.11.1 measures the marginal benefit that results from a unit increase in b_1 through an increase in the CEO's effort. The RHS of Equation 2.11.1 represents the marginal cost: the first term is the direct effect of an increase of b_1 on the marginal cost, and the second term is an indirect effect through the increased probability of higher output, thus higher payment. As argued at the beginning of this section, salary is not necessary under symmetric information. The firm, as a result, does not need to use the first period compensation to provide feedback to the CEO. This means that the bonus solely serves the purpose of incentivising the CEO. The following lemma characterizes the conditions under which the incentive effect becomes weaker or stronger as the type varies.

Proposition 16 *Objective incentive compensation under symmetric information:*

- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} = 0$*, then* $\frac{db_1^s}{d\theta} = 0$*;*
- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} > 0$*, then* $\frac{db_1^s}{d\theta} > 0$*;*
- *If* $\frac{\partial^2 lnP(\theta,e)}{\partial\theta\partial e} < 0$*, then* $\frac{db_1^s}{d\theta} < 0$.

This proposition shows that linear super-modularity between type and effort is not sufficient to give rise to increasing bonus with respect to the type θ. To suffice a positive

[21]It can be easily shown that the board will not set b_1 to less than zero. If b_1 is less than zero, the CEO will not exert effort. The board could increase b_1 to improve profit. As a result, b_1 less than zero cannot exist in equilibrium.

relationship, stronger super-modularity, more specifically, positive log super-modularity is required.

Under information symmetry, the variation of the bonus with respect to θ is derived from the increasing absolute value of Marginal Rate of Substitution (MRS) between effort and compensation as managerial skills θ improve. If the log super-modularity is zero, the MRS between effort and compensation does not vary with the type. If the log super-modularity is positive, the absolute value of MRS between effort and compensation increases with the CEO's ability. The board offers higher bonus to a CEO who has better managerial skills. Proposition 16 will have other important implications for later analysis.

Asymmetric Information with Informative Bonus This section characterizes the optimal contract under asymmetric information and the full spectrum of the contract is set free to provide feedback. Following Maskin and Tirole (1992), I invoke the revelation principal and focus on the direct revelation mechanism, which greatly simplifies the analysis by restricting the message space to be the type space and restricting attention to truth-telling constraint. As argued in Maskin and Tirole (1992), by appealing to the revelation principal, I'm not suggesting that they are realistic. What one typically sees in actual contracts is a schedule in which compensation is tied to output. This is equivalent to a direct revelation mechanism. For concreteness, I provide the proof in Appendix 1.

The model is formally a signalling game and as such can have multiple equilibria. For the purpose of this analysis, the most interesting among them is a separating Perfect Bayesian Equilibrium (PBE). The separating PBE in this section is defined as follows:

Definition A separating equilibrium is a Perfect Bayesian Equilibrium which satisfies:

1. The board offers a contract $[f_1(m), b_1(m, y)]$ that maximizes the firm's profit.
2. The CEO's belief of the board's evaluation is $\beta(\eta = m | f_1(m), b_1(m, y)) = 1$.
3. Given the contract $[f_1(m), b_1(m, y)]$ and the belief, the CEO chooses an effort level which maximizes her own utility.

I first analyze the CEO's problem. Given $b_1(m)$, the CEO chooses effort e to maximize her utility:

$$\max_e P(m, e) b_1(m) - \psi(e)$$

From the first order condition, the optimal level of effort as a function of m and $b_1(m)$ is $e^* = e(m, b_1(m))$. The effort level depends on the message m and the second period bonus $b_1(m)$. Unlike b_1 in the previous section, the bonus $b_1(m)$ under asymmetric information depends on the message the board wants to convey to the CEO.

Given the optimal effort level of the CEO, the board's problem ($P1$) is as follows:

$$\max_{f_1(\cdot),b_1(\cdot)} P(\theta,e)(1-b_1(m))-f_1(m)$$

$$s.t. \quad e^* = e(m,b_1(m)) \qquad IC_a$$

$$\theta \in arg\max_m P(\theta,e(m,b_1(m)))(1-b_1(m))-f_1(m) \qquad IC_p$$

$$P(\theta,e)b_1(m)-\psi(e)\geq 0 \qquad for\ all\ \theta \quad IR_a$$

Constraint IC_a is the CEO's incentive constraint obtained from her maximization program. Constraint IC_p is the board's truth-telling constraint.[22] If it is satisfied,the board will choose to truthfully report the signal it receives. Taking the first order derivative of IC_p w.r.t. message m, the following equation holds at $m=\theta$:

$$\frac{df_1(\theta)}{d\theta}=\frac{\partial P(\theta,e)}{\partial e}\left(\frac{\partial e(\theta,b_1)}{\partial\theta}+\frac{\partial e(\theta,b_1)}{\partial b_1}\cdot\frac{db_1(\theta)}{d\theta}\right)(1-b_1(\theta))-P(\theta,e)\frac{db_1(\theta)}{d\theta}$$

Lemma 11 *The principal's problem $P1$ is equivalent to the following maximization problem $P1'$ for each type of θ:*

$$\max_{f_1(\cdot),b_1(\cdot)} \frac{\partial P(\theta,e(\theta,b_1(\theta)))}{\partial\theta}(1-b_1(\theta))(1-H(\theta))$$

Substituting Constraint IC_a into the objective function and applying integration by parts, problem $P1$ with two constraints is simplified to problem $P1'$. The optimal level of second period compensation under asymmetric information is denoted as $b_1^a(\theta)$. It is given by the following equation:

$$\frac{\partial^2 P(\theta,e)}{\partial\theta\partial e}\frac{\partial e(\theta,b_1^a)}{\partial b_1^a}=\frac{\partial P(\theta,e)}{\partial\theta}+\frac{\partial^2 P(\theta,e)}{\partial\theta\partial e}\frac{\partial e(\theta,b_1^a)}{\partial b_1^a}b_1^a(\theta) \tag{2.11.2}$$

Similar to Equation 2.11.1, the LHS of Equation 2.11.2 measures the marginal benefit due to a unit increase in b_1 through an increase in the CEO's effort. The RHS of Equation 2.11.2 represents the marginal cost: the first term is the direct effect of an increase of b_1 on the marginal cost, and the second term is an indirect effect through increased probability of higher output, thus higher payment.

On the one hand, a higher evaluation outcome (high θ) directly results in higher output. On the other hand, a CEO who receives a higher evaluation will work harder, leading indirectly to higher output. In contrast to Equation 2.11.1, it is not the output $P(\theta,e)$ but the output sensitivity to private information $\frac{\partial P(\theta,e)}{\partial\theta}$ that matters for the characterization of the optimal level of bonus. When deciding the optimal bonus, the board maximizes the part of

[22] Since the production technology exhibits super-modularity, the concavity of the board's truth-telling constraint can be ensured.

profit that directly comes from the private information θ and only considers the effect of b_1 on this part of profit. Any profit coming indirectly from private information, namely, through an increase in effort, will be allocated to the CEO as a signalling cost in order to prevent the board of low evaluation from mimicking.

The following proposition shows the condition under which objective information does not provide feedback to the CEO:

Proposition 17 ***Information Invariant Condition***
If the following condition is satisfied, the bonus does not provide feedback, or is information insensitive:

$$\frac{\partial^2 lnP(\theta, e)}{\partial\theta\partial e} = 0$$

In other words, the production technology can be expressed as the product of two separating functions of θ and e respectively, that is, $P(\theta, e) = h(\theta)f(e)$. Under information asymmetry, the board only considers how the bonus affects the part of profit that directly comes from private information θ. If the information invariant condition is satisfied, the MRS between effort and bonus is the same as that under symmetric information. In other words, the information invariant condition mutes any effects of information asymmetry on the bonus. The board finds that maximizing the profit coming directly from θ is the same as maximizing the total profit.

As a result, this condition guarantees that bonus contains no informational value to the CEO, that is, it does not provide feedback. Its only role is to provide incentives. The board fully relies on the first period salary to provide feedback, while the bonus is paid at a level as if the CEO knew her own type (recall Proposition 16). The following corollary gives the first period salary.

Corollary 7 *If the information invariant condition is satisfied, the first period salary is:*

$$f_1(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial P(\widetilde{\theta}, e)}{\partial e} \frac{\partial e(\widetilde{\theta}, b_1(\widetilde{\theta}))}{\partial \widetilde{\theta}} (1 - b_1(\widetilde{\theta}))d\widetilde{\theta}.$$

And $\frac{df_1(\theta)}{d\theta} > 0$.

By offering $f_1(\theta)$, the board credibly communicates its evaluation to the CEO, which changes the CEO's belief and motivates her to make more efforts. This channel is different from the incentive effect provided by bonus b_1. It affects effort through convincing the CEO of her ability to achieving higher output, while the pay per unit of effort is held constant. The incentive channel affects the CEO's effort through raising the pay b_1, while the CEO's belief of her own type is held constant.

Salary is increasing in the signal that the board receives. This is to say, to prevent the board of lower evaluation from mimicking, the board with higher evaluation needs to pay more to provide credible feedback. Because the bonus is invariant to private information, how b_1 varies with θ does not concern the board. The salary only equals the profit coming from deceiving the agent to make more efforts when the board mimics.

Bonus Providing Feedback The above section analyzes the optimal contract under the information invariant condition. Once it is breached, it is unclear whether bonus will be the same as under information symmetry. The following proposition provides the answer.

Proposition 18 ***bonus Providing Feedback I***

$$If \quad \frac{\partial^2 lnP(\theta, e)}{\partial\theta\partial e} > 0, \qquad b_1^a(\theta) > b_1^s(\theta).$$

- *the first period salary is*

$$f_1(\theta) = \int_{\underline{\theta}}^{\theta} \{ \underbrace{\frac{\partial P}{\partial e}\frac{\partial e}{\partial \widetilde{\theta}}(1 - b_1^a(\widetilde{\theta}))}_{>0} + \underbrace{\frac{\partial P}{\partial e}\frac{\partial e}{\partial b_1^a}\frac{\partial b_1^a(\widetilde{\theta})}{\partial \widetilde{\theta}}(1 - b_1^a(\widetilde{\theta})) - P\frac{\partial b_1^a(\widetilde{\theta})}{\partial \widetilde{\theta}}}_{<0} \} d\widetilde{\theta},$$

If the information invariant condition is breached, Proposition 18 shows that bonus could provide feedback as well. A higher evaluation outcome (higher θ) improves the marginal productivity of effort in terms of log likelihood of high output. This is equivalent to saying that the MRS between effort and bonus, which determines the level of b_1, is higher than that under symmetric information. This leads to higher bonus under asymmetric information, that is, $b_1^a > b_1^s$ (recall Proposition16).

As a result, to prevent a board with a low evaluation outcome to mimic, a board with a good one will increase bonus in addition to its level under symmetric information. Thus, bonus, in addition to its incentive role, also plays a role in providing feedback to the CEO. I here impose a regulatory condition in order to focus on non-decreasing bonus.

Similar to the case under the information invariant condition, the salary could provide feedback to the CEO, as the first term of $f_1(\theta)$ is positive. However, the sum of the last two terms of $f_1(\theta)$ is negative, which implies that the importance of salary in providing feedback is undermined, because the bonus takes over the role of feedback provision.

If the sum of the last two terms is sufficiently negative, $f_1(\theta)$ will be decreasing in θ, which implies that bonus and salary are **substitutes**. In an extreme case in which the sum of the last two terms cancels out the first term, salary becomes zero and the board relies solely on the bonus to provide both incentives and feedback. In a different situation where the sum of the last two terms is not sufficiently negative, $f_1(\theta)$ will be increasing in θ, which implies that bonus and salary are **complements**.

Proposition 19 ***bonus Providing Feedback II***

$$If \quad \frac{\partial^2 lnP(\theta, e)}{\partial\theta\partial e} < 0, \qquad b_1^a(\theta) < b_1^s(\theta).$$

The first period salary is

$$f_1(\theta) = \int_{\underline{\theta}}^{\theta} \{ \underbrace{\frac{\partial P}{\partial e}\frac{\partial e}{\partial \widetilde{\theta}}(1 - b_1^a(\widetilde{\theta}))}_{>0} + \underbrace{\frac{\partial P}{\partial e}\frac{\partial e}{\partial b_1^a}\frac{\partial b_1^a(\widetilde{\theta})}{\partial \widetilde{\theta}}(1 - b_1^a(\widetilde{\theta})) - P\frac{\partial b_1^a(\widetilde{\theta})}{\partial \widetilde{\theta}}}_{>0} \} d\widetilde{\theta},$$

Proposition 19 shows that if log super-modularity is negative, bonus is even lower than the level under symmetric information. To prevent the board with a low evaluation from mimicking, the salary needs to be increased in addition to the level under the information invariant condition to counteract the lowered incentive compensation if the board mimics, as implied by the positive sum of the last two terms of $f_1(\theta)$. Compared to the case of symmetric information, this result characterizes the condition under which there is under-effort-provision.

Chapter 3

Soft Information, Innovation, and Stock Returns

3.1 Introduction

Investment in research and development (R&D) stimulates innovation and technological change. Yet it is difficult for outside investors to decipher how it will ultimately impact firm value (Eberhart, Maxwell, and Siddique, 2004; Cohen, Diether, and Malloy, 2012; Hirshleifer, Hsu, and Li, 2013). In this paper, we demonstrate that CEO compensation changes following internal subjective reviews contain soft information of CEO performance, and therefore predict future R&D successes and abnormal stock returns.

Subjective evaluations are usually based on soft performance measures, information that is either non-contractible or difficult to quantify. Firms conduct subjective evaluation to provide implicit incentives (Baker, Gibbons, and Murphy 1994; Hayes and Schaefer 2000; Prendergast 2002) as opposed to explicit incentives provided by compensation based on objective performance measures.[1] Because R&D activities usually have a long investment horizon and are explorative in nature, they naturally fall into the subject of internal reviews of a CEOs leadership in organizing firm activities. Thus compensation awards based on those reviews may contain soft information of R&D successes and can be used as early predictors of a firms long-run performance.

Yet it is usually difficult to infer ex ante incentive schemes from ex post compensation data, given that (a) subsequent negotiations and changes in the market environment are equally determinative of compensation and performance and (b) researchers can infer only those patterns that are consistent with certain types of guaranteed in-

[1] In addition to observable but non-verifiable information, firms may also have private information about the CEOs performance or ability (Levin 2003; Fuchs 2015).

centives. This critique is generally valid for empirical studies of compensation and especially so in the case of subjective evaluation, which by definition is neither verifiable nor even observable by outsiders. Hence any resolution of the dynamics in question benefits considerably from the study of ex ante compensation contracts.

We hand-collect 649 CEO contracts for S&P 500 firms along with reasons for compensation changes as given in their proxy statements. We find that CEO contracts are often both flexible (in terms of compensation adjustment) and explicitly subject to future reviews. Surprisingly, most of these clauses relate to changes in base salary, the compensation component that is typically regarded as fixed. In 55 percent of the sample contracts, the salary is subject to future reviews.[2] Such clauses are more prevalent for CEO-firm pairs with potentially more information asymmetry, such as firms that invest heavily in R&D or have more dispersed analyst forecasts.

Our identification of soft information regarding R&D successes is based on a simple framework of CEO compensation changes. We first classify salary increases if the CEOs real (i.e., inflation adjusted) salary growth is positive and then categorize them by the extent of contemporaneous changes in equity-based compensation. We then link stand-alone salary increases to ex-ante contracts. The average CEOs base salary (resp., equity-based compensation) increases in 69 percent (resp., 16 percent) of all years. Those CEOs who are subject to review are more likely to receive salary raises, which the firm is more likely to justify based on subjective reasons. Such raises are more prevalent in firms characterized by higher R&D investment. These positive correlations suggest that ex ante subjective review clauses may provide incentives that cannot be offered through strictly performance-based compensation in firms with high R&D investment.

If compensation changes due to subjective reviews indeed contain soft information of R&D successes, we should see that long-run returns improve following the changes. Indeed, a long-short portfolio strategy that invests into firms with stand-alone salary increases following scheduled subjective reviews earns abnormal returns of roughly 4 percent per year. Moreover, only those salary increases that are either based on subjective evaluations or offered by firms investing heavily in R&D predict favorable long-run performance. A long-short portfolio that invests into firms that have high R&D growth and offer stand-alone salary increases earns roughly 8 percent per year. This pattern of returns is practically unaffected when we adjust portfolio returns by

[2]For bonus and equity-based compensation, such flexible clauses are much rarer because these components are often linked to company-wide plans andin the case of equity-based compensationsubject to rules that protect against dilution of share value. In 5 percent (resp., 13 percent) of contracts, the awarding of bonuses (resp., equity) is at the firms sole discretion. Therefore, in this paper we focus on clauses related to the CEOs base salary.

size, value, momentum, and liquidity factors or by the Daniel, Grinblatt, Titman, and Wermers (1997; hereafter DGTW) characteristics-based benchmark.

In order to isolate further the marginal effect of compensation changes on future stock returns, we run return-forecasting regressions. One and two years after a stand-alone increase in CEO salary, the firms monthly stock returns are (respectively) 40-bps and 20-bps higher. This findingsalary increases might predict future returnsis robust to controlling for firm characteristics and to using different approaches when correcting standard errors.

We explore the channel through which salary increases are associated with long-run returns. Two years following CEO salary increases, the number of news articles about the firms new product developments increases by 17 percent, with positive abnormal returns after those announcements. Three years following CEO salary increases, the number of patent filings also increases. These findings strongly suggest that firms use subjective evaluationsand offer salary increasesfor early R&D investment success that is not yet reflected in explicit performance measures. In addition, we find that soft information deciphered from subjective evaluation is more predictive of future returns for firms with higher idiosyncratic risk and greater analyst forecast standard deviation. Finally, we demonstrate that the return predictability of CEO compensation changes is not affected by the inclusion of the persistent firm characteristic of innovation ability in our tests. In addition, we offer suggestive evidence that the board is more likely to offer a CEO stand-alone salary raise to reward her effort rather than ability. We also show that stand-alone bonus increase is not a good predictor of returns.

This paper provides the first explicit empirical analysis of subjective evaluation as a tool for incentivizing executives. Our results are in accord with theories on efficient contracting because they suggest that tying executive pay to subjective evaluations is no less important for incentives than are explicit performance measures. Previous literature (Bushman, Indjejikian, and Smith, 1996; Ittner, Larcker, and Rajan, 1997; Hayes and Schaefer, 2000; Murphy and Oyer, 2001) examines mainly subjective bonuses for CEOs. That research finds that subjective information is more useful in environments where objective performance measures (e.g., accounting information) are less indicative of true performance. Gibbs et al. (2004) and Ederhof (2010) analyze discretionary bonus payments that are paid in addition to a formula-based bonus component. Based on data collected from CEO contracts, we were surprised to discover that subjective evaluation is most often associated with changes in base salarythat is, the compensation component normally viewed as being insensitive to performance. Our paper shows that salary increases due to subjective evaluation is an important aspect of compensation and is predictive of future performance.

This paper also complements a growing literature highlighting the markets inability to properly value R&D investment. Our approach is picking up a new pattern in the cross-section of stock returns associated with the markets misevaluation of innovation. The recent evidence on firm-level R&D activities suggests that the market appears to underreact to the information contained in R&D investments. For example, Eberhart, Maxwell, and Siddique (2004) find that large increases in R&D investment predict positive future abnormal return; Cohen, Diether, and Malloy (2012) demonstrate that past information about firms success at R&D gives insight into their potential for future success; and Hirshleifer, Hsu, and Li (2013) show that firm-level innovative efficiency (measured as patents scaled by R&D) forecasts future returns. We show that our results are unaffected by the inclusion of these measures in our tests, suggesting that the market may undervalue innovation due to a lack of R&D related soft information that is only known to the firm.

Our paper also contributes to the developing empirical literature on executive contracts. Schwab and Thomas (2005) describe a sample of 375 contracts from a legal perspective. Gillan, Hartzell, and Parrino (2009) show that many CEOs operate without an explicit contract, and they study the choice between explicit and implicit contracts. We shall focus on the compensation section of CEO contractsexamining their clauses and linking them to ex post changes in compensation. Goldman and Huang (2014) document the ex-ante severance contracts and ex post separation pay of S&P 500 CEOs. They find evidence that, in forced departures, discretionary separation pay is used to facilitate a smooth transition from the discharged ex-CEO to a new CEO.

3.2 Data

We analyze the CEO compensation of all firms that were part of the S&P 500 in at least one of the years between 1994 and 2008. We obtain realized compensation for these CEOs from ExecuComp and then construct a sample of compensation contracts by screening proxy statements as well as forms 10K, 10Q, and 8K (along with their corresponding exhibits) for explicit employment agreements.[3] For cases in which those agreements are not available, we screen the same filings for indications of whether the CEO is subject to any agreement containing clauses related to compensation. Of all the S&P 500 CEOs in our sample, 649 employment agreements are publicly available. Our final data set consists of 8,190 firm-year observations, including 3,250 observations of firms that disclose the existence of a CEO employment agreement. Excluding the

[3] Securities and Exchange Commission (SEC) Regulation S-K (§ 229.601) requires the disclosure of any management contracts or any compensatory plan of named executive officers as defined by item 402(a)(3) (§ 229.402(a)(3)).

first and last years of each CEOs tenure leaves us with 5,242 observations. We exclude the first and the last year of a CEOs tenure because CEOs are often compensated for more months than their stated tenure and so compensation changes during those years could be incorrectly classified.

It is worth mentioning that even though we search many filings, we cannot be certain that the firms not disclosing information on employment agreements do not actually sign any. Hence a nondisclosing firm may be wrongly classified as one whose CEO operates without a contract. However, that would bias results concerning subjective valuation toward having no effect on changes in compensation, which means that our findings represent a lower bound on the strength of such effects. This is because some CEO compensation increases that do, in fact, result from subjective evaluations could be wrongly treated as occurring in the absence of such evaluation. The portion (40 percent) of our sample firms whose CEOs have an explicit contract is in line with the previous literature: Gillan, Hartzell, and Parrino (2009) report that, in 2000, about 46 percent of S&P 500 firms had a comprehensive written employment agreement with their CEO; Schwab and Thomas (2005) find that 42 percent of the firms they surveyed had a contract with their CEO. We hand-collect reasons for compensation changes from firms proxy statements. Companies are required to disclose not only the criteria underlying executive compensation decisions but also the relationship between their compensation practices and corporate performance.[4] This information is reported in the compensation table of the companys proxy statement. The other data used in our analysis come from standard sources. In particular, we obtain firms financial information from Compustat, stock returns from the Center for Research in Security Prices (CRSP), board and corporate governance information from Risk Metrics Corporation (RISKMETRICS), financial analyst estimates from the Institutional Brokers Estimate System (I/B/E/S), and product announcements from S&P Capital IQ.

Table 1 gives summary statistics of the explanatory variables that we usenamely, firm characteristics, CEO characteristics, and labor market characteristics. For each variable we report its mean, median, and standard deviation as well as minimum and maximum values. (See Table A.1 in Appendix 1 for the definitions of these variables.) Our sample firms have an average of 24$billioninassetsand$11 billion in sales (US dollars); their average leverage ratio is 33 percent and return on assets (ROA) is 7 percent. Return explained is the percentage of return that can be explained by the market factor; the average return explained of these firms is 29 percent. The sample firms have an average of 32 percent idiosyncratic risk (as defined by Wurgler and Zhuravskaya

[4] The SEC website (http://www.sec.gov/answers/execomp.htm) gives a detailed account of regulations on executive compensation.

2002). Analyst forecast standard deviation is 13 percent, and the segment number is 2.7 on average. The mean of CEO tenure is seven years and of CEO age is 55. About 67 percent of a typical board is occupied by independent directors, and 29 percent of all boards are busy boards (as compared with 21 percent in Fich and Shivdasani 2006). Some 13 percent of CEOs are either hired from outside the firm or have worked in the firm for less than a year. Industry CEO turnover averages about 12 percent but varies, across industries, from a minimum of no turnover to a maximum of 75 percent turnover.

3.3 Subjective Reviews and Compensation Changes

In this section, we explore contract clauses related to subjective reviews and provide a brief discussion of the role of subjective reviews to guide our empirical tests. We then introduce our compensation change variable and reasons for those changes.

3.3.1 Review Clauses and Subjective Reviews

Review clauses mostly appear in the salary section of executive contracts. Panel A of Table 2 shows that the salary section is most indicative of the flexibility of executive contracts: more than 75 percent of contracts explicitly allow for salary adjustment, as compared with 5 percent and 13 percent allowing for adjustments in (respectively) bonus and equity. More than half of the sample contracts contain no explicit rules governing adjustments of bonus and equity. One possible reason for this is that bonus and equity compensation is often subject to company-wide plans, and equity compensation is subject to rules protecting shareholders form dilution. Since 2003, both the New York Stock Exchange (NYSE) and NASDAQ require shareholder approval of all equity-based compensation plans; furthermore, there is no longer a de minimis dilution exception for nonofficer and nondirector plans (Lund 2006).[5] One implication of this new standard is that the board of directors must convince shareholders before adjusting the CEOs equity-based compensation.

We can group such review clauses into three categories: flexibility clauses, which govern the direction of compensation changes; review clauses, which indicate whether or not the compensation must undergo subjective reviews; and factor clauses, which indicate the basis on which a salary level is set.

Flexibility clauses. Panel B of Table 2 documents the frequency of clauses that

[5]Prior to 2003, the NYSE required shareholder approval of equity-based compensation plans covering officers and directors but allowed exceptions for such plans that were broadly basedthat is, those that offered equity to a large segment of the firms employees.

characterize the rigidity of contractual salary and govern the direction of future adjustments (see Table A.2 for examples). Examining these clauses reveals that many contracts already have built-in flexibility; that is, the contract itself allows for future adjustment. In our sample, 76 percent of contracts have discretionary clauses concerning the flexibility of adjustments to the contractual base salary. We find only two fully rigid contracts (i.e., that explicity preclude both upward and downward adjustments); another 30 contracts are identified as fully flexible in the sense that base salary can be adjusted both upward and downward. About 71 percent of contracts are partly flexible. Panel B also classifies these contracts further in terms of how their clauses are worded, and Panel D shows that a significant number (though far from the majority) of contracts explicitly indicate how compensation should change in the future.

Review clauses. From a legal standpoint, many of the flexibility clauses just described are nonbinding. In particular, phrases such as the Company may increase and an annual base salary of not less than leave insufficient grounds for litigation should the firm choose not to raise the CEOs salary. One explanation for these legally nonbinding flexibility clauses is the existence of concurrent review requirementsthat is, language explicitly indicating that compensation levels are subject to future review. Some contracts specify the review frequency (annually, in most cases), and some require the affected executives consent before a pay reduction. We give examples of such phrases in Panel E. More than half (55 percent) of the contracts require future reviews. For CEOs hired under such a contract, review of the base salary is mandatory. Thus contracts may build in not only the possibility but also the frequency of reviews, which are usually at the sole discretion of the board but need not be one-sided; we found five contracts that include the CEO as one of the review parties.

Review factors. Only some 9 percent of contracts delineate the factors considered by the board when adjusting the CEOs base salary; see Panel F of the table for an overview. Examples of such factors include the firms financial condition and the CEOs performance. The specifying of these factors provides useful guidance for our multivariate study, in which we control for factors that could affect CEO compensation. Salaryunlike bonus and equity-based compensationis not directly linked to explicit performance measures; hence the fact that most contracts do not contain review factors also indicate that some firms subjectively evaluate the CEOs contribution and adjust his (or her) base pay accordingly.

In short, nearly all contracts feature some flexibility and few are fully rigid (in either direction). Contract clauses that govern compensation nearly always account for possible changeseither by specifying them in advance or making the contract subject to review. We next provide a brief discussion of the role played by subjective reviews in influencing compensation changes to guide our empirical tests. In order to incentivize risk-averse CEOs, their compensation should be closely linked to performance (Holmstrom 1979). Yet real-world incentives are frequently informal and based on nonverifiable performance. Hence the evaluation of individual performance requires both quantitative and qualitative analysis. In line with this argument, contracts explicitly plan for subjective reviews so as to evaluate the CEOs contribution as evidenced by the contractual results above.

When firms use both verifiable and nonverifiable performance measures, the latter may be given more weight if the former are noisier. Bushman, Indjejikian, and Smith (1996), Ittner, Larcker, and Rajan (1997), Hayes and Schaefer (2000), and Murphy and Oyer (2001) provide empirical evidence for this argument by examining how CEO bonuses are affected by subjective factors. Baker, Gibbons, and Murphy (1994) study subjective performance measures in a repeated game framework and show that incentive provision schemes relying on subjective performance measures is less costly if financial performance measures are noisy. Therefore firms are more inclined to undertake subjective reviews.

Provided that contracts offer incentives based on subjective evaluations, we should be able to use the relevant clauses to predict compensation changes. That is, CEOs with contracts that call for periodic review should be more likely than those whose contracts do not to have their compensation adjusted based on subjective evaluations.

The logic behind subjective review is that the firm rewards its CEO for good performance before that performance is impounded into verifiable objective measures. Such compensation raises are justified only if future performance does, in fact, improve (or is highly probable to improve). The lack of an actual link between subjective evaluationbased rewards and future performance would be indicative of governance problems, since in that case boards would be doling out rewards that are unfounded and thus arbitrary. We therefore expect that (well-governed) firms in which CEO compensation is based on subjective evaluations are likely to achieve better long-run performance than are firms that compensate their CEOs in terms of other criteria. We explore the leading channel through which long-run performance can be improved without being immediately evident in objective measures: investments in R&D and product development.

3.3.2 Ex-post Salary Changes and Subjective Evaluation

This section introduces our compensation change variable. Salary changes are categorized as either raises or cuts. In defining changes, we take a conservative approach: we classify a change in salary as a raise only if the CEOs real (i.e., inflation adjusted) salary growth is positive; in contrast, our salary cut classification is simply based on nominal salary growth. That is, an upward adjustment that does not exceed the inflation rate is not classified as a raise.

Table 3 gives summary statistics of changes in salary, which can be decreasing, stable, or increasing (as shown, respectively, in columns 1, 2, and 3). Salary cuts are rare; they occur in only 5 percent of all years and average 13 percent. Salary raises are frequent; they occur in 69 percent of all years and average 9 percent. In only 25 percent of years do CEOs receive the same salary or increase in salary less than inflation. Table 3 also provides the average compensation for CEOs that received cuts or raises. Those CEOs who received salary cuts have a lower average salary ($0.6 million) than those who did not.

Because equity changes are subject to rules protecting shareholders form dilution, we further categorize salary increases by the extent of contemporaneous changes in equity-based compensation in Panel C.[6]

Equity-based compensation is typically granted in multiyear cycles (Hall 1999), and recipients are not entirely vested until a prespecified period of time has elapsed (Cadman, Compbell, and Klasa 2011). Because our objective is to study compensation decisions rather than realized changes in wealth, we focus on changes in grant values. Therefore, we assume that if a CEO receives no equity in years between two grants, it is an instance of no change in equity-based compensation. We then compare the current grant value to the last previous grants value. Our analysis ignores trivial changesthat is, changes in equity-based compensation that fail to exceed (in absolute value terms) that years change in salary.

The pattern observed is most often (46 percent of years) an increase in CEO salary but with no change in equity-based pay. In 13 percent of our sample years we observe CEOs receiving more salary and equity. Finally, in 10 percent of all years, CEOs receive a salary increase but a contemporaneous cut in equity-based pay.

Any changes in compensation that are built into an incentive scheme should also be written into the CEOs employment contract. We therefore expect that compensation changes ex post will be related to contract characteristics ex ante. Table 4 tests this idea by linking compensation changes to review clauses.

Since not all CEOs sign contracts and since not all firms that sign contracts disclose

[6] In most of the analysis we focus on stand-alone salary increases. First, most review and flexibility clauses relate only to changes in that base salary. Second, adjustments to equity-based compensation are subject to mandated rules that aim to protect shareholders from undue dilution of share value (see Section II). Third, a firm for which there are no significant salary changes is not likely to change the equity-based compensation of its CEO, as shown in Table 2. Salary cuts are also very rare. Fourth, increases in every component are more indicative of a CEOs better outside option than good soft performance.

their particulars, we control for the possibility of selection into our contract sample. For this purpose we use a Heckman (1979) approach and report the inverse Mills ratio for all second-stage regressions. Appendix 4 describes the first-stage regression in detail. In the second stage, we regress indicator variables for subsequent compensation change on our explanatory variable: contractual clauses requiring periodic review.

We start by regressing salary changes on such contract clauses. We then (a) categorize salary increases by the extent of contemporaneous changes in equity-based compensation and (b) report results of the second-stage regressions on contract clauses. The dependent variables in columns 15 of Table 4 are indicators for a stand-alone salary increase; those used in columns 610 are indicators for an increase in overall compensation.[7]

Our main finding is that the review requirement clause predicts stand-alone salary raises. The inclusion of that clause in a CEOs contract increases by 7.5 percent the likelihood of a stand-alone salary increase when the only control is for year fixed effects. This result is robust to regression specifications that also control for CEO tenure and age, the inverse Mills ratio, and industry fixed effects. We also include the review factor dummy to control for salary increases that are based on factors explicitly written into the contract; the results are robust to controlling for this indicator variable.

However, in none of our regression specifications are review requirement clauses are significantly associated with salary raises that concur with raises or cuts in equity-based pay. This means that stand-alone salary raises are more likely to be part of an incentive scheme that is based on nonverifiable performance; otherwise, such raises would also be positively linked to overall compensation increases.

3.3.3 Ex-post Salary Changes and Subjective Reasons

Public firms in the US need to provide a narrative for compensation changes in the proxy statement. To distinguish compensation changes following subjective evaluation from others we study these narratives in this section. For firms in our sample, the boards of directors provide reasons for 67 percent of compensation changes. We can categorize these reasons to three types: good subjective performance, good objective financial performance, and benchmarking to peers. Table 5 gives summary statistics for these reasons and lists the keywords that we consider to signify different types.

We first discuss the reasons for compensation changes that are not based on performance. In the sample, 29 percent of changes result from the boards benchmarking

[7] In unreported tests, we regress two other factors on contract clauses: (i) a salary increase combined with a decrease in equity-based compensation; and (ii) change in total compensation. We find no association between review requirement clauses and either of these compensation changes.

of CEO compensation to other executives who work in the same industry. The board may also increase compensation upon contract renewal or to adjust for inflation.

Second, some increases are the direct result of good financial performanceas reflected by net income, ROA and so on. However, such changes account for only 7 percent of increased compensation instances. At the same time, about 40 percent of changes in compensation are rewards for general financial performance, which is not tied to a specific, financial-based metric.

Third, the board of directors may reward the CEO for good subjective performance as measured by so-called soft criteria, consistent with the contractual clauses for subjective performance reviews. Nearly 17 percent of salary increases are claimed to be given as a reward for subjectively evaluated performanceleadership, strategic planning, accomplishing an expansion or restructuring and so on. These narratives do not link the subjective to tangible financial performance as direct outcomes those activities.

No reason is given (in proxy statements) for a full third of all salary increases. There are two possible causes of our not observing an explicitly stated reason. First, the board can arbitrarily increase CEO pay. In that case, there should be no systematic differences between firms that increase CEO salary with versus without giving reasons for doing so, for those increases are entirely arbitrary. Second, if the ex ante contract already requires periodic subjective review of that compensation, then the board is not obliged to offer a specific reason. It might be hard to describe the subjective reason for the increase. That reticence can also be beneficial if the firmsay, for competitive reasonswould prefer not to disclose its motivation for increasing CEO compensation until a more advantageous time. In either case, however, there should be systematic differences between firms that do and do not give reasons for increasing CEO pay. A comparison between compensation increases that are justified in terms of (good) subjective performance with those for which no explicit justification is given will allow us to determine whether the respective firms involved exhibit any detectable systematic differences.

Are CEOs with subjective review clauses indeed more likely to receive salary increases based on their performance that is yet to be impounded into objective measures? We link compensation changes to subjective evaluation and test whether stand-alone salary increase is truly related to subjective reviews.

In Panel A of Table 6, we show that they are: CEOs working under contracts that contain review requirement clauses receive salary increases in 45.3 percent of all years when no reason is given and in 61.1 percent of years when either no reason is given or their performance is evaluated subjectively. The corresponding numbers for CEOs whose contracts do not incorporate review requirement are lower: 31.5 percent and

51.9 percent, respectively.

On the contrary, CEOs working under contracts that contain review requirement clauses receive salary increases in 5 percent of all years when an objective reason is given and in 31 percent of years when CEO has achieved good financial performance overall. The corresponding numbers for CEOs whose contracts do not incorporate subjective reviews are higher: 8 percent and 41 percent, respectively. Results for the CEOs with required annual reviews are in Panel B and similar to those reported in Panel A.

Thus, salary increases are unlikely to reflect an arbitrary board decisioneven when no specific reasons are given for the raise. Otherwise, we would observe similar frequencies of stand-alone salary increases for CEOs with and without review clauses. Alternative explanations for a salary increase are that it is simply part of an overall company compensation plan or is due to a contract renewal. Yet as shown in Panel F of Table 5, we find that fewer than 1 percent of compensation changes are attributable to these reasons.

Following regression specifications in Table 4, we conduct similar tests in Table 7. We include the dependent variables of stand-alone salary increases based on no reason and subjective reasons. This is to test whether CEOs working under contracts that include review clauses are more likely to have their compensation adjusted based on subjective reasons.

In column 1 of the table we see that CEOs with subjective review clauses are 5 percent more likely to receive stand-alone salary increases unaccompanied by any reasons. In column 2 we add stand-alone salary increases following good performance, as evaluated subjectively, and find that CEOs with subjective review clauses are 8 percent more likely to receive stand-alone salary increases. The values reported in column 3 indicate that contracts with subjective review clauses are not more highly predictive of stand-alone salary increases based on good performance as evaluated objectively. These results support our hypothesis that CEO contracts account for future subjective evaluation and thus predict stand-alone salary increases as rewards following such evaluation

3.4 Linking Compensation Changes to Firm Performance

The previous results suggest that stand-alone salary increases reflect positive subjective evaluations. If these increases are indeed justified by the CEOs good performance, not yet impounded into objective performance measures, then the result should be a long-run improvement in returns.

3.4.1 Portfolio Returns

To test this hypothesis, we examine average returns on portfolios formed using information about compensation changes. Specifically, we conduct a calendar-time portfolio analysis in which stocks are sorted by the previous years changes in compensation. At the end of each year, we sort stocks into two portfolios, one consisting of firms that offer stand-alone salary increases and the other consisting of firms that do not offer such increases. The portfolios so constructed are held for three years and are rebalanced yearly.

We compute three- and four-factor alphas (as in Fama and French (1996), and Carhart (1997)) by running time-series regressions of excess portfolio returns on the market (MKT), size (SMB), value (HML), and momentum (UMD) factor returns. In addition, we characteristically-adjust the portfolio returns using 125 size/book to market/momentum benchmark portfolios as in Daniel, Grinblatt, Titman, and Wermers (1997). In short, those benchmarks are constructed from the returns of 125 passive portfolios that are matched with stocks held in the evaluated portfolio on the basis of market capitalization, book-to-market ratio, and prior-year stock return characteristics.

Table 8 reports the average monthly returns to these portfolios, and illustrates our main return result: firms that offer stand-alone salary increases in the past outperform those that do not in the future. This result holds for three- and four-factor alphas and for characteristically-adjusted returns. As can be seen from Panel A, a long-short portfolio spread (Spread) between stocks in the portfolio that offers stand-alone salary increases and the portfolio that does not is significant and large under all risk adjustment specifications. For example, when three-factor adjustment is used, abnormal returns are most pronounced in year 1 after the salary increase. The magnitude of abnormal returns to the long-short portfolio is about 27-bps (t=2.24), which translates to 3.2 percent annually. Abnormal returns are still significant but less so in year 2 with a smaller magnitude of 19-bps (t=2.18), which translates to 2.3 percent annually. Significance of the long-short portfolio disappears in year 3.

To test whether subjective evaluation is related to abnormal returns, in Panel B of Table 8 we further sort firms which offer stand-alone salary increases into two subportfolios based on the reasons given for compensation changes: one consisting of firms that offer stand-alone salary increases evaluated subjectively and the other consisting of firms that offer those increases evaluated objectively. A long-short portfolio spread (Spread-subjective performance) between stocks in the portfolio that evaluates stand-alone salary increases subjectively and stocks in the portfolio that does not offer stand-alone salary increases is significant and large under all risk adjustment specifications. When four-factor adjustment is used, the magnitude of abnormal returns to the long-short portfolio is about 35-bps (t=2.58), which translates to 4 percent annually. Abnormal returns are still significant but less so in year 2 with a smaller magnitude of 24-bps (t=2.49), which translates to 2.9 percent annually. Significance of the long-short portfolio also disappears in year 3.

In contrast, a long-short portfolio spread (Spread-objective performance) between stocks in the portfolio that evaluates stand-alone salary increases objectively and stocks in the portfolio that does not offer stand-alone salary increases is not significant. This result suggests that compensation changes based on subjective performance do contain soft information which is not captured by objective performance measures.

As R&D activities usually have a long horizon and come to fruition late, firms with a recent and substantial increase in R&D expenditures are most likely to rely on subjective evaluations and offer standalone-salary increases. Panel C of the table test this idea. We further sort firms that offer stand-alone salary increases based on yearly percentage increases in R&D expenditures. We rank those firms by R&D growth above and below industry median in that year and report returns of the two portfolios. When four-factor adjustment is used, a long-short portfolio spread (Spread-R&D growth high) between stocks in the portfolio that offers stand-alone salary increases and have high R&D growth and stocks in the portfolio that does not offer stand-alone salary increases predicts positive abnormal returns of 82-bps (t=3.68) in year one and 56-bps (t=3.31) in year 2, which respectively translate to 9.8 percent and 6.7 percent annually. In contrast, a long-short portfolio spread (Spread-R&D growth low) between stocks in the portfolio that offers stand-alone salary increases and have low R&D growth and stocks in the portfolio that does not offer stand-alone salary increases is not significant. The results gleaned from Panels B and C of Table 8 lend support to our hypothesis that firms increase CEO compensation following subjective reviews for good performance that is not yet manifest in standard financial measures.

While R&D growth is indicative of investment in potentially new research projects, R&D/sales reflects the amount of R&D projects in place. As a robustness check, in Panel D of the table we instead sort firms that offer stand-alone salary increases based on the ratio of R&D/sales. We rank them by R&D/sales above and below industry median and report returns of the two portfolios. When four-factor adjustment is used, a long-short portfolio spread (Spread-R&D/sales high) between stocks in the portfolio that offers stand-alone salary increases and have high R&D/sales and stocks in the portfolio that does not offer stand-alone salary increases predicts positive abnormal returns of 32-bps (t=2.39) in year one and 21-bps (t=2.23) in year 2. In contrast, a long-short portfolio spread (Spread-R&D/sales low) between stocks in the portfolio that offers stand-alone salary increases and have low R&D/sales and stocks in the portfolio that does not offer stand-alone salary increases is not significant. Sorting based on quintiles produces very similar (even stronger) results.

As a robustness check, we exclude from the full sample years from 2001 to 2003 around which the stock market crashed. The crash may lend equity grants less attractive (Frydman and Jenter, 2010) and lead to a substitution of equity grants with cash-based pay. As shown in Panels E, after removing those years, the returns are more statistically significant and economically substantial than those in the full sample. For instance, when four-factor adjustment is used, the magnitude of abnormal returns to the long-short portfolio is about 57-bps (t=2.68) in year one, which is 30-bps higher than the full sample, and 25-bps (t=2.26) in year 2, which is 6 bps higher.

In sum, the results from Table 8 demonstrates that our classification scheme, which is designed to capture soft information of subjective evaluation, produces a large and significant spread in future abnormal returns. This finding also highlights the fact that it is critical to understand subjective evaluation as an implicit means of providing CEO incentives.

3.4.2 Cross-Sectional Regressions

To isolate further the marginal effect of compensation changes on future stock returns, we perform return forecasting regressions; results are reported in Table 9. Because residuals may be correlated across firms or across time, we estimate standard errors clustered by firm and by year-month (Petersen 2009). We also conduct Fama-MacBeth return forecasting regression (Fama and Macbeth, 1973). The dependent variable is monthly stock return in the subsequent period, and the independent variable of interest is the indicator variable of stand-alone salary increase in year t. Additional control variables include firm size (Banz, 1981), book-to-market ratio (Rosenberg, Reid, and Lanstein, 1985; Fama and French, 1992), and past returns (to capture the liquidity and microstructure effects documented by Jegadeesh (1990)). The monthly cross-sectional regression estimates in Table 9 confirm our earlier portfolio results: firms that offer stand-alone salary increases in the past outperform in the future. Specifically, stand-alone salary increase significantly predicts stock returns after one year in both one- and two-way clustering and Fama-Macbeth regressions. It still predicts after two years, but only in return forecasting regressions with one- and two-way clustering and also less significantly.

The coefficients reported in columns 2 and 4 of the table imply that, a stand-alone salary increase of any nontrivial amount results in a 40-bps (t=3.82) increase in stock returns in one year under one-way clustering, a 40-bps (t=2.95) under two-way clustering and a 20-bps (t=1.95) in Fama-Macbeth regression. These increases are significant when we control for firm characteristics. The magnitude and significance decline two years after the stand-alone salary increase to only a 20-bps (t=1.9) increase in stock returns under one-way clustering, a 20-bps (t=2.04) under two-way clustering and 0-bps (t=-0.1) in Fama-Macbeth regression.

The return regressions offer further confirmation of our hypothesis that nonperformance-based compensation is used to reward CEOs for good performance that is yet to be evident in the firms stock returns.

3.5 Mechanism

We show in the previous section that compensation increases based on subjective evaluations predict long-run stock performance. Here we explore a channel through which subjective evaluations of CEO performance affect long-term but not immediately verifiable returns.

3.5.1 Innovation as a Channel

Because R&D activities usually have a long investment horizon and are explorative in nature, they naturally fall into the subject of internal reviews of a CEOs ability in planning and organizing firm activities. Thus compensation changes based on those reviews may contain soft information of R&D successes and can be used as early predictors of a firms long-run performance. If financial measures have not yet absorbed the effect of novel research and/or new product development, we should observe that such firm activities come to fruition following rewards based on subjective evaluation of those activities. Table 10 tests this idea.

Following the regression specifications in Table 4, in Table 10 we differentiate stand-alone salary increases by yearly percentage increases in R&D expenditures above or below the industry median one year before the compensation change, as used in Panel C of Table 8. Column 1 to 5 show that CEOs with subjective review clauses are 10 percent more likely to receive stand-alone salary increases in firms with a high increase in R&D investment, but not (as shown in column 6 to 10) in firms with a low increase in R&D investment. This result is significant and robust after we control for inverse Mills ratio, review factors and different fixed effects. Because firms with high levels of R&D spending are usually difficult to evaluate, our findings is therefore consistent with the argument that firms rely more on subjective assessments if the performance is not yet impounded into objective measures.

If compensation increases due to subjective evaluation are justifiable, not only should returns improve in the long-run but also firm activities. Table 11 summarizes the outcome of R&D activities, specifically, the extent to which compensation changes predict the number of future product announcements, abnormal returns to those announcements and the number of future patent filings. In particular, we control for other forms of salary increases, namely, with contemporaneous increases and decreases in equity compensation. This is to validate our presumption that the board should prefer salary to equity in subjective evaluations as the latter is subject to shareholder approval.

In Panel A, we regresson compensation changesthe number of product announcements at one, two, and three years after those changes in compensation. Farrell and Saloner (1986)s theory argues that a firms product development can greatly influenced by competing firms, as early adopters bear a disproportionate share of transient incompatibility costs. According to Hendricks and Singhal (2008), the effect of product introduction delays on performance is significantly related to industry size and profitability. To avoid any inflation in the number product announcement due to various industry effects, we divide the number of each firms product announcements by the average amount of product announcements made in the same year by all firms that operate in the same industry. Since firm activity may be affected by variations in time and in firm characteristics, we control for both year and firm fixed effects. We find that, two years after an increase in stand-alone salary, the number of product announcement increases. Although stand-alone raises are thus positively associated with the number of product announcements, other changes in compensation exhibit no such pattern.

In the event of a positive subjective evaluation, we expect that compensation changes predict an improvement in returns to new product announcements. In Panel B of Table 11, we calculate the average abnormal return changes before and after each product announcement date (using a 5-day window) and then take the mean for all product announcement events over each fiscal year. Doing so enables us to show how compensation changes predict future return changes. Stand-alone salary increases predict returns that increase significantly (by about 0.6 percent) over the 5-day windows that we observe.[8]

In Panel C, we regresson compensation changesdummies which indicate an increase in patent filings one, two, and three years after those changes in compensation. We

[8] As a robustness check, we also look at both 25-day and 45-day windows. Some predictive power remains, it is both statistically and economically weaker though.

do not use percentage increases in patent filings as dependent variables, as some firms have zero patent filings in some years. To avoid co-linearity between the dependent variable and firm fixed effects, we only control for year and industry fixed effects in Panel C. We find that, three years after an increase in stand-alone salary, the firm is significantly more likely to have an increase in patent filings by 12 percent. But other changes in compensation exhibit no such pattern.

These results indicate that stand-alone raises are a good predictor of the future success of a firms research activities. Moreover, firm activities improve after two years, which is largely consistent with the results in portfolio analysis and return forecasting regressions.

3.5.2 Information Availability

Because soft information regarding early R&D successes is, by its very nature, hard to quantify, outside investors usually rely on public information sources. If compensation changes due to subjective evaluation truly contain soft information, then they should be more predictive of future returns for firms with less information availability. Table 12 tests this idea.

We first examine idiosyncratic risk in Panel A of the table. Idiosyncratic risk reflects the availability and volatility of firm-specific information (Campbell et al. 2001). We introduce an interaction term between idiosyncratic risk and stand-alone salary raise. If firm-specific information is impacting whether investors are able to decipher the future successes of firm activities, then stand-alone salary increases offered by firms with lower idiosyncratic risk should have less return predictability.

Panel A of the table shows that compensation changes due to subjective evaluation are indeed more predictive of future returns for firms with higher idiosyncratic risk. A 1 percent increase in idiosyncratic risk for firms which offer stand-alone salary raise leads to 1.4-bps increase in monthly stock returns after year 1 and 2.1-bps after year 2. The results remain robust under one-way and two-way clustering. We then examine analyst forecast STD in Panel B of the table. Variations in analyst forecasts reflect divergence of the market opinions in a firms future success. The greater the analyst forecast STD is, the more divergent the opinions are. Similar to Panel A, we introduce an interaction term between analyst forecast STD and stand-alone salary raise. If analyst forecast STD is impacting whether investors are able to form a consensus view in the future successes of firm activities, then stand-alone salary increases offered by firms with higher analyst forecast STD should have more return predictability.

Panel B shows that compensation changes due to subjective evaluation are indeed more predictive of future returns for firms with higher analyst forecast STD. A 1 percent increase in analyst forecast STD for firms which offer stand-alone salary raise leads to 0.3-bps increase in monthly stock returns after year 1. The results remain robust under one-way and two-way clustering.

3.6 Robustness

In this section, we provide a series of additional tests aimed at isolating the mechanism that drives our main results.

3.6.1 Determinants of Review Clauses

In section III, we use review clauses to show that CEOs whose contracts contain those clauses are more likely to receive stand-alone salary increases. One might argue that those clauses may also be written into contracts for other reasons. For instance, competitive labor market conditions require the board to frequently review the CEOs performance and adjust her compensation accordingly. Additionally, a powerful CEO could demand more favourable clauses. Therefore, they do not necessarily represent the need for subjective evaluations. To alleviate this concern, we directly investigate the determinants of review clauses here to show that firm-CEO pairs that need subjective evaluations are more likely to sign contracts with review clauses.

According to Baker, Gibbons, and Murphy (1994), the firms that should use subjective evaluation are those with noisy objective performance measures. For example, firms with high levels of R&D investment tend to reward CEOs based on subjective evaluation, and typically considerable time elapses before R&D investments come to fruition. A firm that anticipates using compensation raises to reward subjectively evaluated performance would be well advised to sign a CEO employment contract that mandates periodic reviews. In addition, firms characterized by more information asymmetry, and/or more volatile returns can be expected to sign flexible contracts with their CEOs.

Of course, subjective reviews are not the only reasons for contractual flexibility and reviews. A large literature studies the various causes, and most such causes reflect the existence of outside options, the extent of managerial power, and financial constraints on the firm.

Compensation changes may result from ex post renegotiation, which might occur in response to changes in a CEOs outside options. As shown in Table 5, about 29 percent of salary increases are explained by benchmarking to peer groups. Firms must offer compensation high enough that CEOs are willing to forgo their outside options. Along these lines, matching theories (e.g., Gabaix and Landier 2008) argue that larger firms need more able CEOs and so must offer higher compensation to attract them. We follow Gabaix and Landier in using total assets to proxy for firm size. We use industry CEO turnover and homogeneity to control for labor market depth (as in Gillan, Hartzell, and Parrino 2009).

Monitoring subjective performance reviews is difficult for investors and perhaps even for outside board members. Such reviews can thus be manipulated more easily (than objective criteria) by CEOs. Indeed, Bebchuk and Fried (2004) argue that managers wield substantial influence over their own pay arrangements. We therefore follow Fich and Shivdasani (2006) and use both the proportion of independent board directors and a busy board indicator variable as proxies for managerial power.

Firms facing financial constraints have less cash to offer as salary and so may prefer to offer more equity-based pay than do less constrained firms. Babenko, Lemmon, and Tserlukevich (2011) posit that financially constrained firms may finance investments using cash inflows from employees exercising their stock options. Consistently with this argument, Core and Guay (2001) document a greater use of options for compensation by firms that face financing constraints. We use a dummy variable for distress (based on Altman 1968) to control for financial constraints.

Table 13 reports the results of our Probit regressions. The dependent variable is an indicator for the review requirement clause. If a contract specifies that periodic review is required, then clearly the board demands that executive compensation be evaluated (and perhaps adjusted) on a regular basis. Whether a contract contains review clauses thus reflects the boards ex ante willingness to adjust CEO compensation. A principal component analysis of contract clauses (see Appendix 3) confirms that the review requirement is a viable indicator of contract flexibility, since that factor has the largest loading. The explanatory variables used in our regressions include proxies for information asymmetry, firm characteristics, corporate governance, and CEO characteristicsbut only for years in which the CEOs contract is effective. Column 3 and 4 in the table include industry characteristics, while columns 2 and 4 include industry fixed effects.

Columns 1 and 4 in Table 13 show that a firm investing heavily in R&D is more likely to have review clauses in its CEOs contract. This finding is consistent with our hypothesis because such firms are the most likely to realize their performance gains (or losses) after some delay. As a result, the board factors this consideration into the contract and so allows subjective evaluation to predominate in reviews of the CEOs performance. We also find that outside CEOs are more likely (12 percent higher) subject to review requirements; this result is significant across all specifications. In other words, the board also relies on subjective evaluations when assessing a CEO about whom they have scant prior information. We believe that this finding is of greater relevance than the managerial power argument because the coefficient for percentage of independent directors is not significant.

Our hypothesis is further buttressed by the results for idiosyncratic risk. Columns 1 and 3 in Table 13 reveal that firms characterized by higher levels of idiosyncratic risk are also more likely to offer CEO contracts that include the review requirement. For instance, a 1 percent increase in idiosyncratic risk increases the likelihood of a review clause by 0.24 percent in the third specification. In contrast, it is less likely that review clauses will be required by distressed firms. This finding could be explained by the asymmetry of adjustments that result from compensation review (i.e., since upward adjustments are far more common than downward ones).

Industry characteristics are also significantly related to contract characteristics. For example, we find that firms operating in a more homogeneous industry are less likely to write review requirement clauses into the contract. This result is likely explained by the greater ease of assessing managerial skills within industries that are relatively less heterogeneous. Collectively, these findings suggest that firms featuring strongly asymmetric information are more inclined to offer flexible CEO contracts which require subjective reviews, which reinforces our identification strategy using standalone salary increases.

3.6.2 Firm Innovation Ability

Persistent firm characteristics, for instance, high innovation ability (Cohen, Diether, and Malloy, 2012), may explain early R&D success. For subjective evaluation to have an incentive effect, the board should offer stand-alone salary raise to CEOs based on their performance. Otherwise, our compensation change variable is not founded on subjective evaluation and only captures soft information regarding persistent firm characteristics.

We thus conduct the same regressions as in Table 9 with the inclusion of the innovation ability variable introduced by Cohen, Diether, and Malloy (2012). This control variable is constructed based on a firms past sales over R&D investment and measures the firms ability to turn R&D investment into sales. Ability estimate is constructed as follows: we run separate regressions for 5 different lags of R&D from year t-1 to t-5; we then take the average of five R&D regression coefficients as ability. Ability high equals one for a stock if its ability estimate is in the top quartile in a given month. R&D high equals one for a stock if its R&D scaled by sales is above 70th percentile.

Table 14 shows that the return predictability of CEO compensation changes is not affected by the inclusion of the persistent firm characteristic of innovation ability.

With the inclusion of the innovation ability, standalone salary increase still leads to 30-bps increase in monthly returns after year 1 and 20-bps after year 2. This result suggests that our compensation change variable captures soft information that cannot be explained by persistent firm characteristics.

3.7 Discussion

3.7.1 Nature of Soft Information

In previous sections, we argue that compensation changes contain soft information that is indicative of early R&D successes and cannot be explained by firm innovation ability. But what is the exact nature of the soft information? What does stand-alone salary increase reward for? Although it is not the focus of the paper, we propose two candidates, namely, CEO ability and effort, and discuss which one is more likely to be captured by our compensation change measure.

In theory, both CEO ability and effort could contribute to early R&D successes, both of which therefore justify compensation increases. If the board offer the CEO salary increase to reward her ability, then stand-alone salary increases should be more predictive of future returns for outside CEOs. In addition, assuming that CEO ability is a relatively persistent characteristic, the board does not need to reward the CEO once it has learnt her ability. Therefore, the likelihood of receiving such increase should decline over a CEOs tenure. In contrast, CEO effort could change over time depending on both explicit and implicit incentive schemes. Therefore, if the nature of soft information concerns CEO effort, then stand-alone salary raise should not be more predictive of future returns for outside CEOs and the likelihood of receiving such increase should be stable or, at least, not exhibit a clear declining pattern over a CEOs tenure.

In unreported tests, we conduct otherwise the same regressions as in Table 12 but on outside CEO and CEO tenure. We construct an interaction term between the outside CEO dummy and stand-alone salary raise, but we do not find that stand-alone salary raise is significantly more predictive of future returns for outside CEOs. We also construct an interaction term between CEO tenure and stand-alone salary raise. Neither do we find that stand-alone salary raise is significantly more predictive of future returns for CEOs in their early years of tenure. As robust checks, we construct several dummies by categorizing CEO tenure into two groups and using different cut-off years, none of which produce any significant results.

In Figure 1, we plot the frequency of stand-alone salary increase over CEO tenure.

Specifically, Figure 1.1 shows the percentage of CEOs who receive stand-alone salary raise over their tenures among all CEOs, and Figure 1.2 shows the percentage of CEOs who receive stand-alone salary based on subjective reasons among all CEOs. In both figures, we do not observe a clear declining pattern. In fact, Figure 1.2 show that the frequency of stand-alone salary raise goes down from year 2 to year 4 and then goes up again almost to the same level in year 6 as in year 4, and the variation between year 4 and year 6 is only about 5 percent.

Our results suggest that stand-alone salary raise is less likely to contain soft information regarding CEO ability but rather CEO effort. One might also argue that the nature of soft information is project-specific and has nothing to do with either CEO ability or effort, for example, positive productivity shock (or luck). Although this argument remains theoretically sound, we simply cannot think of such shocks in reality that happen so frequently over CEO tenure and could lead to such a fairly stable pattern as shown in Figure 1.

3.7.2 Bonus Changes

In previous sections, we focus on stand-alone salary increase as a measure of positive subjective evaluation outcome. In Table 10, we also show that overall increase in salary and equity does not explain the improvement in firm activity. Gibbs et al. (2004) and Ederhof (2010) analyze discretionary bonus payments paid in addition to any bonus warranted by a prespecified formula. The board of directors may use such discretionary bonuses to reward the CEOs good performance (as subjectively evaluated). In this section, we discuss whether salary and bonus are not perfectly substitutable forms of compensation and whether stand-alone bonus increase is a better predictor of returns. In short, we find that the results reported here for stand-alone salary increases do not, in general, apply to bonus increase.

The bonus is often calculated as a multiple of base salary, where the multiple is determined by a formula that incorporates performance factors (De Angelis and Grinstein 2014). We therefore identify the actual change in bonusrather than the mechanical change arising simply from any base salary changeby viewing each bonus strictly as a multiple of salary. Panel A of Table 3 shows that the change in this bonus multiple is highest (42 percent) for CEOs who received salary cuts and lowest (6 percent) for CEOs with unchanged base salary. This indicates that changes in salary are not always in parallel to changes in bonus and thus likely the outcome of separate review processes, consistent with the contracts. In unreported results, we also find that a CEOs salary and bonus both increase in 36 percent of all years but that they both decline in only 4 percent of all years. In 30 percent of all years, salary increases but bonus declines.

Table 15 reports results from an analysis that mimics the ones described in Tables 11but for changes in bonus compensation, not stand-alone salary. Panel A of Table 15 shows that indeed an increase in stand-alone bonus compensation is not correlated with the firms number of product announcements. In other words, remuneration that is based on stand-alone salary increases due to subjective evaluation is more indicative of future firm activity. Panel B shows that, unlike stand-alone salary increases, stand-alone bonus increases are not significantly related to abnormal returns during the 5-day window around new product announcements. One possible reason for this finding is that there is a nondiscretionary component to bonuses that, like equity-based compensation, depends explicitly on objective performance measures. This feature of bonuses weakens their power to predict firm activity. The CEOs base salary, in contrast, is not tied to any explicit performance metric and so should be considered more discretionary than a bonus. Panel C of the table, which address the patent filings, reinforces our hypothesis that a salary increase based on subjective evaluation is a better indicator of the firms future activity and performance.

One possible reason for the finding that salary and the bonus multiple are neither substitutes nor complements is that, as pointed above, bonus is sometimes tied to explicit performance metric, which weakens the flexibility of adjusting it. Another possible reason is that many firms have a bonus pool which puts a cap on the total bonus that is allowed to be given to all the managers who participate in the bonus program. This again limits a firms flexibility of adjusting the CEOs bonus, as doing so will affect other managers.

3.8 Conclusions

This paper introduces a novel early predictor of R&D successes and abnormal stock returns: stand-alone salary increases following internal CEO performance reviews. We demonstrate that one motive for increasing salary is to reward CEOs following subjective evaluation. We document that executive contracts explicitly schedule subjective reviews of performance. A long-short portfolio strategy that invests into firms with salary increases following scheduled subjective reviews earns abnormal returns of roughly 4-8 percent per year. Importantly, these positive review outcomes also predict future product announcements, returns to such announcements, and increases in patent filings.

First, we establish that CEO contracts are usually both flexible and subject to future review. If some compensation changes represent rewards following subjective reviews, then they should be part of the firms incentive scheme. Our results confirm this hypothesis, as we document that CEOs with more flexible contracts are more likely to receive increases in compensation. Second, if compensation changes are indeed rewards for good subjective performance then the firms stock price should eventually increase. In line with this hypothesis, we find that monthly abnormal portfolio returns of firms that give compensation increases are significantly positive both one year and two years after stand-alone salary increases. Return forecasting regressions further confirm that rewards for good subjective performance are positively correlated with the firms long-run returns.

Third, we find that firm activities improve following stand-alone salary increases. Specifically, product announcements and patent filing increase in firms that give stand-alone salary raises following subjective evaluation, abnormal returns around subsequent announcements of product developments are positively associated with compensation changes.

We also conduct several robustness checks to show that stand-alone salary increases are indeed related to subjective evaluation. We show that firms with more R&D investment are more likely to sign contracts with explicit review clauses, and firms characterized by greater dispersion among analyst forecasts and greater return volatility are more likely to incorporate review clauses based on prespecified factors. We also show that the positive relation between stand-alone salary increases and returns is not driven by firm innovation ability.

Lastly, we offer suggestive evidence that the board awards a CEO stand-alone salary increases that are indicative of early R&D successes mainly due to her effort rather than ability. Furthermore, we show that stand-alone bonus increase is not a good predictor of returns.

Our paper contributes to the literature on subjective evaluation of executives by providing evidence gathered from CEO contracts and an in-depth analysis of changes in firm activity and returns following rewards based on subjective review. Instead of studying compensation based on explicit performance measures, we focus on how compensation contracts whose terms do not rely on such measures play a key role in incentivizing CEOs. It also complements a growing literature highlighting the markets inability to properly value R&D investment. Our approach is picking up a new pattern in the cross-section of stock returns associated with the markets misevaluation of innovation.

There is still much scope for future work on the channels through which contract clauses affect CEO compensation. It would be worthwhile also to study how explicit and implicit performance measures interact, since that would help us better understand the performance sensitivity of executive compensation.

3.9 References

Altman, Edward. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance 23(4): 589-609.

De Angelis, David, and Yaniv Grinstein. 2014. "Performance Terms in CEO Compensation Contracts." Review of Finance: rfu014.

Babenko, Lemmon and Tserlukevich. 2011. Employee Stock Options and Investment. Journal of Finance 66(3): 981-1009.

Baker, George, Gibbons, Robert and Kevin J Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts," Quarterly Journal of Economics 109(4): 1125-56.

Banz, Rolf W.. 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9: 318.

Bebchuk, Lucian and Jesse Fried. 2004. Pay Without Performance: The Unfulfilled Promise of Executive Compensation. Harvard Business Press.

Bharath, Sreedhar T., Paolo Pasquariello, and Guojun Wu. 2009. Does Asymmetric Information Drive Capital Structure Decisions? Review of

Bushman, R., R. Indjejikian, and A. Smith. 1996. CEO compensation: the role of individual performance evaluation. Journal of Accounting and Economics 21(2): 161-193.

Cadman, Brian D., John L. Compbell, and Sandy Klasa. 2011. Are Ex-ante CEO Severance Pay Contracts Consistent with Efficient Contracting? Mimeo, University of Utah.

Carhart, Mark M.. 1997. On persistence in mutual fund performance. Journal of Finance 52(1), 5782.

Cohen, Lauren, Karl Diether, and Christopher Malloy. 2013. Misvaluing innovation. Review of Financial Studies 26(3): 635-666.

Core, John E. and Wayne R. Guay. 2001. Stock option plans for non-executive employees. Journal of Financial Economics 61(2): 253-287.

Daniel, Kent, Grinblatt, Mark, Titman, Sheridan and Russ Wermers. 1997. "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks." Journal of Finance 52(3): 1035-1058.

Eberhart, Allan, Maxwell, William, and Siddique, Akhtar. 2004, An Examination of the Long-Term Abnormal Stock Returns and Operating Performance Following R&D Increases, Journal of Finance 59, 623-650.

Ederhof, Merle. 2010. Discretion in Bonus Plans. The Accounting Review 85(6): 1921-1949.

Fama, E. and MacBeth, J.. 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81(3): 607-636.

Fama, Eugene F., and French, Kenneth R.. 1992. The cross-section of expected stock returns. Journal of Finance 46(2): 427466.

Fama, Eugene F., and French, Kenneth R.. 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51(1): 55-84.

Farrell, J., and Garth Saloner, 1986. Installed base and compatibility: innovation, product preannouncements and predation. The American Economic Review 76 (5): 940-955.

Fich, Eliezer M., and Shivdasani, Anil. 2006. "Are busy boards effective monitors?" The Journal of Finance 61(2): 689-724.

Frydman, Carola, and Raven E. Saks. 2010. Executive Compensation: A New View from a Long-term Perspective, 19362005. Review of Financial Studies 23(5), 2099-2138.

Fuchs, William. 2015. "Subjective Evaluations: Discretionary Bonuses and Feedback Credibility." American Economic Journal: Microeconomics 7(1), 99-108.

Gabaix, X. and A. Landier. 2008. Why has CEO pay increased so much? Quarterly Journal of Economics 123(1): 49-100.

Gibbs, Michael, Kenneth A. Merchant, Wim A. Van der Stede, and Mark E. Vargus. 2004. "Determinants and effects of subjectivity in incentives." The Accounting Review 79 (2): 409-436.

Gillan, Stuart L., Jay C. Hartzell, and Robert Parrino. 2009. Explicit vs. Implicit Contracts: Evidence from CEO Employment Agreements. Journal of Finance 64(4): 1629-1655.

Goldman, Eitan and Peggy Huang. 2014. Contractual Versus Actual Separation Pay following CEO Turnover. Management Science, forthcoming.

Gompers, Paul, Joy Ishii, and Andrew Metrick. 2003. Corporate Governance and Equity Prices. Quarterly Journal of Economics 118(1): 107-156.

Hall, Brian. 1999. The Design of Multi-year Stock Option Plans. Journal of Applied Corporate Finance 12(2): 97-106.

Hayes, Rachel M. and Scott Schaefer. 2000. "Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance." The Rand Journal of Economics 31(2): 273-293.

Heckman, James J. 1979. Sample selection bias as a specification error. Econometrica 47(1): 153-161.

Hendricks, Kevin B., and Vinod R. Singhal. The Effect of Product Introduction Delays on Operating Performance. Management Science 54 (5): 878-892.

Holmstrom, Bengt. 1979. "Moral Hazard and Observability." Bell Journal of Economics 10 (1): 74-91.

Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li. 2013. "Innovative efficiency and stock returns." Journal of Financial Economics 107(3): 632-654.

Ittner, Christopher D., David F. Larcker, and Madhav V. Rajan. 1997. The Choice of Performance Measures in Annual Bonus Contracts. The Accounting Review 72(2): 231-255.

Jegadeesh, N.. 1990. Evidence of predictable behavior of security returns. Journal of Finance 45(3): 881-898.

Levin, Jonathan, 2003. "Relational Incentive Contracts." American Economic Review 93(3): 835-857.

Lund, Andrew, 2006. What Was the Question? The NYSE and Nasdaq's Curious Listing Standards Requiring Shareholder Approval of Equity-Compensation Plans. Connecticut Law Review, 39(1): 119-158.

Martin, Kenneth J. and Randall S Thomas. 2005. When is enough, enough? Market reaction to highly dilutive stock option plans and the subsequent impact on CEO compensation. Journal of Corporate Finance 11(1-2): 61-83.

Murphy K. and Oyer P. 2001. Discretion in executive incentive contracts: Theory and evidence. Available at SSRN 294829.

Parrino, Robert. 1997. CEO Turnover and Outside Succession: A Cross-sectional Analysis. Journal of Financial Economics 46, 165-197.

Petersen, Mitchell A. 2009. "Estimating standard errors in finance panel data sets: Comparing approaches." Review of financial studies 22.1, 435-480.

Prendergast, Canice. 2002. "Uncertainty and incentives." Journal of Labor Economics 20(2): 115-137.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein. 1985. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11: 917.

Schwab, Stewart J., and Randall S. Thomas. 2005. An Empirical Analysis of CEO Employment Contracts: What Do CEOs Bargain For? Washington and Lee Law Review 63: 231-70.

Taylor, Lucian A. 2013. "CEO wage dynamics: Estimates from a learning model." Journal of Financial Economics 108(1): 79-98.

Wurgler, Jeffrey, and Ekaterina Zhuravskaya. 2002. Does Arbitrage Flatten Demand Curves for Stocks? Journal of Business 75 (4): 583-608.

3.10 Tables and Figures

Table 1 Summary statistics: firm, CEO and labor market characteristics

	Variable	Mean	Median	STD	Min	Max
		(1)	(2)	(3)	(4)	(5)
Firm characteristics	Total assets	24,787	7,387	40,404	7	153,413
	Total sales	10,789	5,284	12,934	0	46,090
	ROA	0.07	0.07	0.06	-0.62	0.17
	ROE	0.17	0.16	0.40	-2.27	1.82
	EPS	4.14	3.43	7.96	-5.95	564.90
	Product announcement	3.85	0	15.80	0	295
	Return explained	0.29	0.26	0.16	0	1
	Idiosyncratic risk	0.32	0.27	0.17	0.06	2.16
	Analyst forecast STD	0.13	0.05	0.29	0.00	3.32
	Segment number	2.70	2	1.28	1	6
	R&D/sales	0.03	0	0.21	0	16.44
	Leverage (net)	0.33	0.36	0.25	-0.88	4.27
	Distress	0.32	0	0.47	0	1
CEO characteristics	Outside CEO	0.13	0	0.34	0	1
	Tenure CEO	7.15	5	6.55	1	46
	Age CEO	54.98	56	7.86	36	74
	Chairman CEO	0.69	1	0.46	0	1
	Independent directors fraction	0.67	0.66	0.15	0	1
	Busyboard (dummy)	0.29	0.00	0.45	0.00	1.00
	Gindex	9.50	9.44	1.48	3	15
Labor market characteristics	Industry homogeneity	0.06	0.05	0.02	0.04	0.14
	Industry CEO turnover	0.12	0.11	0.07	0	0.75
	Industry outside CEO	0.58	0.58	0.07	0.17	0.86

Note: This table presents firm/CEO characteristics for the whole sample. Column 1, 2, 3, 4 and 5 shows the mean, median, standard deviation, minimum and maximum value respectively for each variable.

Table 2 Summary statistics: contracts

(1)	(2)	(3)
Panel A: An overview		
Contract clause	Number	% of Total
Salary		
Explicit discretion	490	75.5%
Bonus		
Explicit discretion	32	5%
Equity grants		
Explicit discretion	87	13%
Panel B: Decomposition of contracts based on salary rigidity:		
Partly flexible	460	70.88%
Upward flexible	378	58.24%
Downward rigid	186	28.66%
Lower bound	185	28.51%
Fully rigid	2	0.31%
Fully flexible	30	4.62%
No discretionary clauses	157	24.19%
Panel C: Downward rigid—conditional		
Salary cut for other executives	7	1.08%
Salary cut for everyone	27	4.16%
CEO Consent	33	5.08%
Panel D: References		
# of Contracts with References	59	9.09%
Amount	2	0.31%
Reference to rate (CPI etc.)	15	2.31%
Reference to top 5 executives	4	0.62%
Reference to the precedent CEO	1	0.15%
Panel E: Review clauses		
Review requirement	355	54.70%
Review frequency Mentioned:	327	50.39%
Regular (Annually,15 Months and 18 Months)	256	39.45%
Irregular	64	9.86%
As often as other officers	7	1.08%
Not specified	28	4.31%
Panel F: Review factors considered in adjustment explicitly expressed in contracts		
Performance of the company and the CEO	56	8.63%
Comparable executives in the firm and industry	23	3.54%
Market conditions	3	0.46%
Financial condition of the firm	3	0.46%
Cost of living	7	1.08%

Note: this table presents the summary statistics of contract clauses. Specific contract clauses are listed in Column 1, the number of contracts that contain such clauses are shown in Column 2, and Column 3 presents the percentage of such clauses. Panel A provides an overview of contractual discretion that the board has over each compensation component. Panel B to C detail the flexibility of the salary component. Panel D shows how salary is adjusted as specified in the contract if any. Panel E and F detail the clauses regarding review requirement, frequency and factors on which the review is based.

Table 3 Summary statistics: compensation changes

Panel A: Change in salary			
Change in salary	**-**	**0**	+
	(1)	(2)	(3)
% of all years	5%	25%	69%
Salary (thousands)	646.04	721.14	712.96
Bonus (thousands)	553.06	793.24	648.94
Equity-based compensation (thousands)	4,082.80	4,677.45	3,850.01
Change in salary	-13.3%	-2.6%	9.5%
Change in bonus multiple	41.7%	5.8%	11.2%
Change in equity-based compensation	0.0%	0.1%	0.5%
Entry salary to industry level	104.5%	97.8%	83.9%

Panel B: Change in salary and equity			
Change in salary		+	
Change in equity-based pay	**-**	**0**	+
% of all years	10%	46%	13%
Salary (thousands)	713.68	699.61	766.99
Bonus (thousands)	677.39	665.56	655.67
Equity-based compensation (thousands)	3,170.69	4,009.93	5,405.40
Change in salary	6.7%	10.8%	6.5%
Change in bonus multiple	19.7%	54.1%	2.8%
Change in equity-based compensation	-25.8%	0.0%	23.7%
Entry salary to industry level	86.6%	81.9%	88.7%

Note: This table presents compensation statistics for salary cut, no change and raise in Panel A. Panel B presents the frequency equity increase/no change/cut when salary goes up or down or stays constant. We take TDC1 in COMPUSTAT as the total compensation. We classify a change in salary as a raise only if the CEO's "real" (i.e., inflation adjusted) salary growth is positive; in contrast, our salary cut classification is simply based on nominal salary growth. That is, an upward adjustment that does not exceed the inflation rate is not classified as a raise. We assume that if a CEO receives no equity in years between two grants, it is an instance of "no change" in equity-based compensation. We then compare the current grant value to the last previous grant's value. Our analysis ignores trivial changes—that is, changes in equity-based compensation that fail to exceed (in absolute value terms) that year's change in salary. Bonus multiple is defined as bonus divided by salary. Industry classification is based on the first two digits of SIC.

Table 4 Compensation changes and contract clause

Dependent variable	Standalone salary increase					Overall compensation increase				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Review requirement	0.075**	0.061*	0.061*	0.077**	0.067*	0	-0.017	-0.014	0.004	0.011
	(0.038)	(0.033)	(0.033)	(0.038)	(0.039)	(0.022)	(0.019)	(0.019)	(0.020)	(0.020)
Mills			0.034	0.024	0.062			0.028	0.037*	-0.049**
			(0.037)	(0.045)	(0.057)			(0.020)	(0.022)	(0.020)
Review factor					0.021					0.042*
					(0.045)					(0.022)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure group	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Age group	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Industry fixed effects	No	No	No	Yes	Yes	No	No	No	Yes	Yes
N	954	954	954	954	954	862	862	862	862	862

Note: This table presents marginal effects for the contract group from Probit regressions and standard errors (in parenthesis) that are heteroskedasticity robust. Dependent variables are dummy variables -- standalone salary raises from column 1 to 5, and overall raises from column 6 to 10. Review requirement dummy is the explanatory variable. Others are control variables, including Mills ratio and review factor dummy. Age group consists of five dummies for CEO age under 45, between 45 and 50, between 50 and 55, between 55 and 60, and above 65. Tenure group consists of three dummies for a CEO who has worked in the same firm for at most 2 year, 3-6 years and more than 6 years. Industry fixed effects are based on the first two digits of SIC.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 5 Summary statistics: reasons for compensation changes

	Keywords	N	Percentage
	(1)	(2)	(3)
Subejctive performance	Leadership	421	9.68%
	Strategy	298	6.85%
	Organizational development	40	0.92%
	Expansion	37	0.85%
	Restructure	3	0.07%
	Subjective	130	2.99%
	Total	731	16.81%
Objective performance	Revenue	291	6.69%
	Net income	41	0.94%
	EPS	95	2.18%
	ROE	54	1.24%
	ROA	11	0.25%
	Total	322	7.41%
General performance	Performance	1,701	39.12%
	Merit increase	67	1.54%
	Total	1,735	39.90%
Indexed to peer	Peer	419	9.64%
	Median	348	8.00%
	Survey	331	7.61%
	Competitive rate	504	11.59%
	Attract	61	1.40%
	Benchmark	36	0.83%
	Total	1,251	28.77%
Others	More responsibility	448	10.30%
	Become CEO	227	5.22%
	Part of the plan	19	0.44%
	Unchanged since	4	0.09%
	Unchanged since	4	0.09%
	Contract renewal	4	0.09%
No reasons given		1,446	33.26%

Note: this table presents the summary statistics of reasons for compensation changes. Keywords that summarize the reason for compensation changes are presented in Column 1. The number of observations that contain those keywords are shown in Column 2, and the percentage of such changes out of total changes provided in Column 3.

Table 6 Reasons for compensation changes and contract clauses

Panel A: Review requirement clause			
	Without review requirement	With review requirement	
Variable	Mean	Mean	t-stats
	(1)	(2)	(3)
No reasons given	0.315	0.453	-6.375 ***
Soft measures of performance	0.519	0.616	-4.227 ***
Objective performance	0.077	0.050	2.293 **
General performance	0.412	0.305	4.785 ***
Benchmarking	0.292	0.259	1.589
Panel B: Review annual clause			
	Without annual review	With annual review	
No reasons given	0.318	0.464	-6.081 ***
Soft measures of performance	0.520	0.628	-4.229 ***
Objective performance	0.077	0.047	2.263 **
General performance	0.409	0.309	4.001 ***
Benchmarking	0.291	0.258	1.447

Note: this table presents the summary statistics of salary increases based on listed reasons as percentage of all salary increases by contract clauses, namely review requirement clause in Panel A and review annual clause in Panel B. We then compare the difference between the two percentages and present the t-statistics in Column 3.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table 7 Standalone salary increases and reasons

Dependent variable	Standalone salary increase*No reason	Standalone salary increase*Subjective Reason	Standalone salary increase*Objective reason
	(1)	(2)	(3)
Review requirement	0.051**	0.081***	0.001
	(0.021)	(0.027)	(0.022)
Review factor	0.021	0.009	-0.037
	(0.041)	(0.053)	(0.040)
Mills	0.039	0.037	0.033
	(0.081)	(0.055)	(0.202)
Tenure group	Yes	Yes	Yes
Age group	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
N	937	937	937

Note: This table presents marginal effects from Probit regressions and standard errors (in parenthesis) that are heteroskedasticity robust. The dependent variable in Column 1 is an indicator variable for standalone salary increases with no reasons provided, with either no reasons provided or with subjective reasons in Column 2, and with objective reasons in Column 3. Age group consists of five dummies for CEO age under 45, between 45 and 50, between 50 and 55, between 55 and 60, and above 65. Tenure group consists of three dummies for a CEO who has worked in the same firm for at most 2 year, 3-6 years and more than 6 years. Industry fixed effects are based on the first two digits of SIC.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 8 Calendar-time portfolio returns

Compensation changes	3-factor alpha	4-factor alpha	DGTW adjusted	3-factor alpha	4-factor alpha	DGTW adjusted	3-factor alpha	4-factor alpha	DGTW adjusted
	Year 1 after portfolio formation			Year 2 after portfolio formation			Year 3 after portfolio formation		
Panel A: Standalone salary increases									
Standalone salary increase	0.46%	0.57%	0.71%	0.48%	0.51%	0.57%	0.57%	0.62%	0.50%
No change in salary	0.26%	0.30%	0.39%	0.28%	0.32%	0.43%	0.78%	0.83%	0.32%
Spread	**0.20%**	**0.27%**	**0.32%**	**0.20%**	**0.19%**	**0.15%**	-0.22%	-0.21%	0.18%
T-stat	2.71	2.24	2.93	2.41	2.18	1.73	-0.14	-0.32	0.29
Panel B: Standalone salary increases --reasons									
Subjective performance	0.51%	0.65%	0.61%	0.53%	0.57%	0.50%	0.58%	0.64%	0.50%
Objective performance	0.31%	0.40%	1.05%	0.28%	0.29%	0.95%	0.54%	0.54%	0.69%
Spread_subjective performance	**0.26%**	**0.35%**	**0.22%**	**0.25%**	**0.24%**	**0.07%**	-0.21%	-0.19%	0.18%
T-stat	2.96	2.58	3.05	2.71	2.49	2.32	-0.08	-0.21	0.02
Spread_objective performance	0.05%	0.10%	**0.66%**	0.00%	-0.03%	0.53%	-0.24%	-0.29%	0.38%
T-stat	1.21	0.88	2.25	0.76	-0.5	1.15	-0.21	-0.54	0.75
Panel C: Standalone salary increases -- R&D growth									
R&D growth high	0.98%	1.12%	0.64%	0.78%	0.88%	0.46%	0.69%	0.79%	0.49%
R&D growth low	0.26%	0.21%	0.62%	0.35%	0.42%	0.41%	0.51%	0.56%	0.46%
Spread_R&D growth high	**0.73%**	**0.82%**	**0.25%**	**0.50%**	**0.56%**	0.04%	-0.09%	-0.04%	0.17%
T-stat	4.00	3.68	1.72	3.23	3.31	0.56	-0.48	-0.47	0.88
Spread_R&D growth low	0.00%	-0.09%	0.23%	0.07%	0.10%	-0.01%	-0.27%	-0.27%	0.14%
T-stat	0.41	-0.23	1.27	1.06	1.05	-0.17	-0.36	-0.5	0.55

Panel D: Standalone salary increases -- R&D/sales									
R&D/sales high	0.50%	0.62%	0.78%	0.50%	0.53%	0.59%	0.55%	0.60%	0.51%
R&D/sales low	0.32%	0.41%	0.60%	0.43%	0.46%	0.53%	0.58%	0.60%	0.64%
Spread_R&D/sales high	**0.24%**	**0.32%**	**0.39%**	**0.22%**	**0.21%**	0.16%	-0.23%	-0.23%	0.19%
T-stat	2.82	2.39	3.18	2.45	2.23	1.48	-0.22	-0.41	0.08
Spread_R&D/sales low	**0.07%**	0.11%	0.21%	**0.15%**	0.14%	0.11%	-0.20%	-0.23%	0.33%
T-stat	1.93	1.43	1.49	1.83	1.62	0.95	-0.02	-0.14	0.79
Panel E: Standalone salary increases -- excluding 2001-2003									
Standalone salary increase	0.46%	0.61%	0.67%	0.43%	0.46%	0.29%	0.65%	0.71%	0.37%
No change in salary	-0.01%	0.04%	0.22%	0.17%	0.22%	-0.02%	0.86%	0.89%	0.28%
Spread	**0.47%**	**0.57%**	**0.45%**	**0.25%**	**0.25%**	**0.31%**	-0.21%	-0.19%	0.09%
T-stat	3.12	2.68	3.12	2.45	2.26	1.65	-0.25	-0.12	0.4

Note: This table shows calendar-time equal-weighted monthly returns and t-statistics to portfolios sorted by changes in compensation in the previous year. In Panel A, we sort stocks into two portfolios at the end of each year, one consisting of firms that offer stand-alone salary increases and the other one consisting of firms that do not offer such increases. The portfolios so constructed are held for three years and are rebalanced yearly. We further sort firms that offer standalone salary increases based on reasons for salary changes in panel B, namely subjective reasons and objective reasons. In Panel C, we instead sort firms that offer stand-alone salary increases based on yearly percentage increases in R&D expenditures. We rank those firms by R&D growth above and below industry median in that year. In Panel D, we sort firms that offer stand-alone salary increases based on R&D/sales. We rank those firms by R&D/sales above and below industry median in that year. Panel E report subsample analysis excluding years 2001 to 2003. We compute three- and four-factor alphas (as in Fama and French (1996), and Carhart (1997)) by running time-series regressions of excess portfolio returns on the market (MKT), size (SMB), value (HML), and momentum (UMD) factor returns. In addition, we characteristically-adjust the portfolio returns using 125 size/book to market/momentum benchmark portfolios as in Daniel, Grinblatt, Titman, and Wermers (1997). Spreads of long-short portfolios are in **bold** if they are positive and 10% significant.

Table 9 Stock return regressions

Dependent variable	Monthly stock return in year 1			Monthly stock return in year 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Standalone salary increase	0.004***	0.004***	0.002**	0.002*	0.002**	-0.000
	(0.001)	(0.001)	(0.012)	(0.001)	(0.001)	(0.015)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm cluster	Yes	No	No	Yes	No	No
Two way cluster	No	Yes	No	No	Yes	No
Fama-Macbeth	No	No	Yes	No	No	Yes
N	96,695	96,695	96,695	96,683	96,683	96,695

Note: This table reports coefficients and standard errors (in parenthesis) of forecasting regressions of stock returns on changes in compensation changes and other control variables that are known to predict stock returns. The dependent variable in columns 1, 2, and 3 is the monthly stock return in the year following the fiscal year end; in columns 4, 5 and 6, it is the monthly stock return in the second year following the fiscal year end. The independent variable of interest is the dummy—standalone salary increases in the previous fiscal year. Control variables include cumulative stock returns at various horizons, firm size and the market-to-book ratio. Standard errors are clustered by firm in Columns 1 and 4 and by firm and year-month in Columns from 2 to 5. I conduct Fama and Macbeth (1973) regressions in Columns 3 and 6.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 10 Compensation changes and contract clause – R&D

Dependent variable	Standalone salary increase									
	High R&D grwoth					Low R&D growth				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Review requirement	0.101**	0.093**	0.100**	0.106**	0.107**	0.013	0.015	0.006	0.06	0.058
	(0.047)	(0.048)	(0.048)	(0.054)	(0.055)	(0.047)	(0.048)	(0.049)	(0.062)	(0.063)
Mills			0.045	0.048	0.05			0.04	0.024	0.024
			(0.051)	(0.085)	(0.083)			(0.095)	(0.089)	(0.088)
Review factor					-0.116					0.061
					(0.097)					(0.165)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure group	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Age group	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Industry fixed effects	No	No	No	Yes	Yes	No	No	No	Yes	Yes
N	502	502	502	502	502	464	464	464	464	464

Note: This table presents marginal effects for the contract group from Probit regressions and standard errors (in parenthesis) that are heteroskedasticity robust. Dependent variables are dummy variables -- standalone salary raises. In Columns from 1 to 5, we take firms with R&D increase one year prior to standalone salary increase higher than the industry median based on the first two digits of SIC. In Columns from 6 to 10, we take firms with R&D increase one year prior to standalone salary increase lower than the industry median based on the first two digits of SIC. Review requirement dummy is the explanatory variable. Others are control variables, including Mills ratio and review factor dummy. Age group consists of five dummies for CEO age under 45, between 45 and 50, between 50 and 55, between 55 and 60, and above 65. Tenure group consists of three dummies for a CEO who has worked in the same firm for at most 2 year, 3-6 years and more than 6 years. Industry fixed effects are based on the first two digits of SIC.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 11 Firm activity

Year	t+1	t+2	t+3
	(1)	(2)	(3)
Panel A: Number of product announcements			
Standalone salary increase	-0.046	0.169**	0.016
	(0.106)	(0.085)	(0.089)
Overall compensation increase	-0.298*	0.107	-0.059
	(0.166)	(0.16)	(0.164)
Salary increase & equity decrease	-0.158	-0.012	-0.008
	(0.147)	(0.109)	(0.135)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
R-squared	0.248	0.321	0.276
N	2,576	2,569	2,588
Panel B: Abnormal returns to product announcements ±5-day window			
Standalone salary increase	-0.001	0.006***	0.001
	(0.003)	(0.002)	(0.004)
Overall compensation increase	-0.002	0.003	-0.005
	(0.003)	(0.003)	(0.005)
Salary increase & equity decrease	0.004	-0.001	-0.001
	(0.003)	(0.004)	(0.005)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
R-squared	0.261	0.373	0.134
N	1,022	1,003	984
Panel C: Patent filings increase			
Standalone salary increase	0.048	0.045	0.121**
	(0.057)	(0.056)	(0.057)
Overall compensation increase	0.017	0.044	0.012
	(0.078)	(0.078)	(0.078)
Salary increase & equity decrease	0.037	0.075	0.151*
	(0.080)	(0.082)	(0.082)
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
N	5,022	5,022	5,022

Note: This table reports coefficients of OLS regressions in panel A and B and Probit regressions in panel C. Stand errors (in parenthesis) are heteroskedasticity robust. Dependent variables in Panel A are the numbers of product announcements normalized by industry average in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t. Specifically, we divide each firm's number of product announcements by the average amount of product announcements made in the same year by all firms that operate in the same industry. Industry classifications are based on the first two digits of SIC. Dependent variables in panel B are average abnormal return changes ±5 days 5 days before and after product announcements in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t. Abnormal returns are calculated by taking the residuals of the regression of the daily stock return on Fama-French three factors. Dependent variables in panel C are dummies indicating whether the number of patent filings has increased in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 12 Information availability

	(1)	(2)	(3)	(4)
Panel A: Idiosyncatic risk				
Dependent variable	Monthly stock return in year 1	Monthly stock return in year 2	Monthly stock return in year 1	Monthly stock return in year 2
	(1)	(2)	(3)	(4)
Standalone salary increase*Idiosyncratic risk	0.014*	0.021***	0.014*	0.021***
	(0.008)	(0.008)	(0.007)	(0.007)
Standalone salary increase	0.006***	0.006**	0.006***	0.006**
	(0.002)	(0.003)	(0.002)	(0.002)
Idiosyncratic risk	-0.005	-0.024***	-0.005	-0.024***
	(0.006)	(0.006)	(0.006)	(0.005)
Other controls	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	No	No
Two way cluster	No	No	Yes	Yes
R-squared	0.004	0.003	0.004	0.003
N	46,025	46,025	46,025	46,025
Panel B: Analyst forecast STD				
Standalone salary increase*analyst forecast STD	0.003**	0.000	0.003**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Standalone salary increase	-0.005	-0.001	-0.005	-0.001
	(0.005)	(0.006)	(0.005)	(0.005)
analyst forecast STD	0.001	0.004	0.001	0.004
	(0.003)	(0.005)	(0.003)	(0.004)
Other controls	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	No	No
Two way cluster	No	No	Yes	Yes
R-squared	0.004	0.002	0.004	0.002
N	42636	42636	42636	42636

Note: This table reports coefficients and standard errors (in parenthesis) of forecasting regressions of stock returns on changes in compensation changes. The dependent variable in columns 1 and 2 is the monthly stock return in the year following the fiscal year end; in columns 3 and 4, it is the monthly stock return in the second year following the fiscal year end. In Panel A, the independent variable of interest is the dummy—standalone salary increases in the previous fiscal year, idiosyncratic risk as introduced by Wurgler and Zhuravskaya (2002), and an interaction term between the dummy and idiosyncratic risk. In Panel B, the independent variable of interest is the dummy—standalone salary increases in the previous fiscal year, analyst forecast STD, and an interaction term between the dummy and the analyst forecast STD. Control variables include cumulative stock returns at various horizons, firm size and the market-to-book ratio. Standard errors are clustered by firm in Columns 1 and 2 and both by firm and year-month in Columns 3 and 4.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 13 Determinants of contract clauses

	Dependent variable	Review requirement			
		(1)	(2)	(3)	(4)
Information asymmetry	R&D/sales	1.095***	0.456**	1.068***	0.439**
		(0.253)	(0.201)	(0.253)	(0.198)
	Outside CEO	0.120***	0.117***	0.120***	0.114***
		(0.029)	(0.033)	(0.029)	(0.033)
	Idiosyncratic risk	0.172*	0.12	0.239**	0.094
		(0.093)	(0.113)	(0.095)	(0.113)
	Depr. &amort.%	-0.172	0.951*	0.03	0.950*
		(0.368)	(0.497)	(0.372)	(0.496)
	Distress	-0.098***	-0.103**	-0.119***	-0.108***
		(0.033)	(0.040)	(0.034)	(0.041)
Industry	Industry homogeneity			-1.034**	2.727
				(0.486)	(2.11)
	Industry outside CEO			0.890***	0.16
				(0.187)	(0.449)
Controls	Independent directors%	0.137	0.157	0.087	0.155
		(0.084)	(0.098)	(0.084)	(0.0982)
	Net leverage	0.194*	0.06	0.181*	0.071
		(0.109)	(0.097)	(0.105)	(0.100)
	Log assets	0.009	0.002	0.006	0.002
		(0.011)	(0.014)	(0.011)	(0.014)
	Tenure group	Yes	Yes	Yes	Yes
	Age group	Yes	Yes	Yes	Yes
	Year fixed effects	Yes	Yes	Yes	Yes
	Industry fixed effects	No	Yes	No	Yes
	N	1,876	1,693	1,875	1,693

Note: This table presents marginal effects from Probit regressions and standard errors (in parenthesis) that are heteroskedasticity robust. The dependent variable is review requirement, a dummy equal to 1 if the contract contains review requirement clause and zero otherwise. Column 3 and 4 include industry characteristics. Age group consists of five dummies for CEO age below 45, between 45 and 50, between 50 and 55, between 55 and 60, and above 65. Tenure group consists of three dummies for a CEO who has worked in the same firm for at most 2 years, 3-6 years and more than 6 years. Industry fixed effects are based on the first two digits of SIC.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 14 Innovation ability

Dependent variable	Monthly stock return in year 1	Monthly stock return in year 2	Monthly stock return in year 1	Monthly stock return in year 2
	(1)	(2)	(3)	(4)
Standalone salary increase	0.003***	0.002*	0.003***	0.002*
	(0.00096)	(0.00101)	(0.00095)	(0.00094)
R&D high * ability high	0.000	0.001	-0.000	0.001
	(0.00183)	(0.0019)	(0.00206)	(0.00209)
Ability high	-0.003*	-0.003*	-0.003	-0.003
	(0.00156)	(0.00161)	(0.00175)	(0.00179)
R&D high	0.002	0.001	0.002	0.001
	(0.0013)	(0.0013)	(0.00145)	(0.00144)
Other controls	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	No	No
Two way cluster	No	No	Yes	Yes
R-squared	0.003	0.002	0.003	0.002
N	96,683	96,671	96,683	96,671

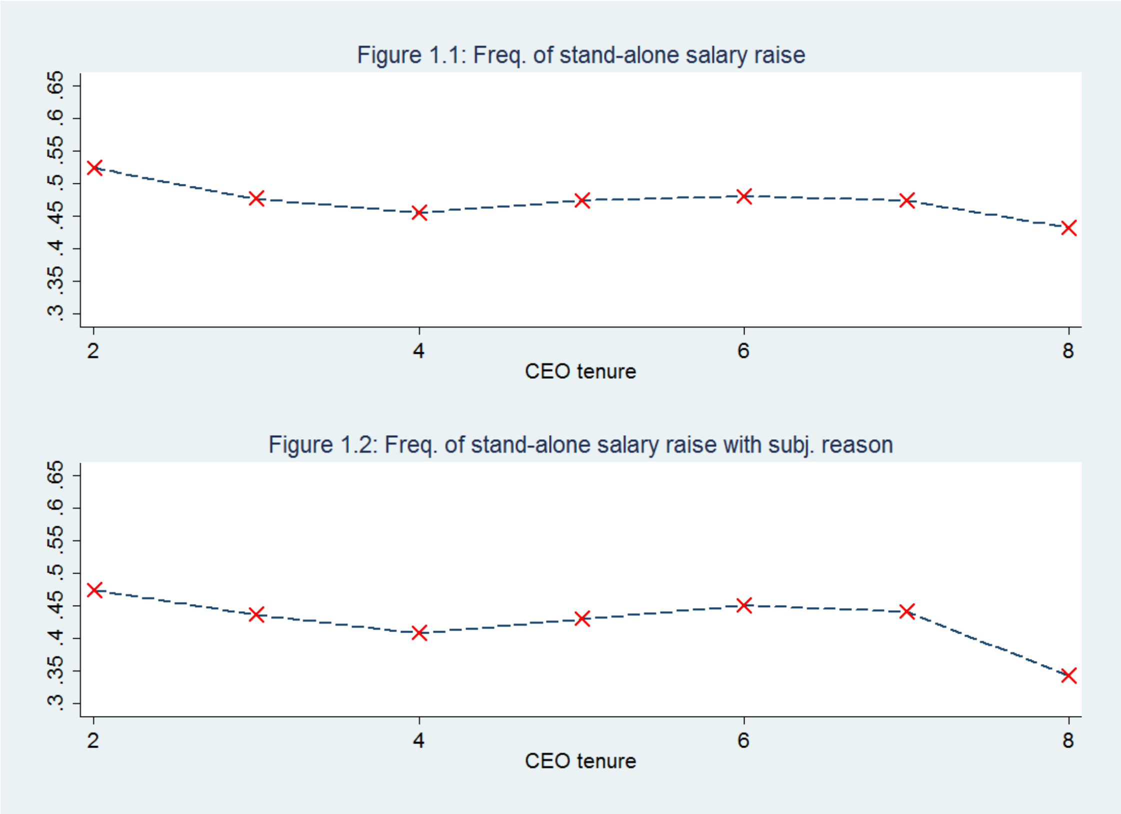
Note: This table reports coefficients and standard errors (in parenthesis) of forecasting regressions of stock returns on stand-alone salary raise with the inclusion of innovation ability as introduced by Cohen, Diether, and Malloy (2012). It is computed by running rolling firm-by-firm regressions of firm-level sales growth on lagged R&D over sales. We run separate regressions for 5 different lags of R&D from year t-1 to t-5; we then take the average of five R&D regression coefficients as ability. Ability high equals one for a stock if its ability estimate is in the top quartile in a given month. R&D high equals one for a stock if its R&D scaled by sales is above 70^{th} percentile. The dependent variable in columns 1 and 2 is the monthly stock return in the year following the fiscal year end; in columns 3 and 4, it is the monthly stock return in the second year following the fiscal year end. Additional control variables are changes in assets, cumulative stock returns at various horizons, firm size and the market-to-book ratio. Standard errors are clustered by firm in Columns 1 and 2 and both by firm and year-month in Columns 3 and 4.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Figure 1 Stand-alone salary raise over CEO tenure



Note: This figure shows the frequency of stand-alone salary increases over CEO tenure. In Figure 1.1, we plot the percentage of CEOs who receive stand-alone salary over their tenures among all CEOs. In Figure 1.2, we plot the percentage of CEOs who receive stand-alone salary based on subjective reasons over their tenures among all CEOs. CEO tenure ranges from second year of their tenure to eighth year.

Table 15 Firm activity (bonus)

Year	t+1	t+2	t+3
	(1)	(2)	(3)
Panel A: Number of product announcements			
Standalone bonus increase	0.087	0.227	0.09
	(0.292)	(0.297)	(0.335)
Overall compensation increase	0.048	0.27	-0.039
	(0.306)	(0.277)	(0.318)
Bonus increase & equity decrease	0.204	-0.057	-0.06
	(0.292)	(0.271)	(0.323)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
R-squared	0.368	0.429	0.424
N	2,576	2,569	2,588
Panel B: Abnormal returns to product announcements ±5-day window			
Standalone salary increase	-0.004*	-0.001	0.001
	(0.002)	(0.002)	(0.003)
Overall compensation increase	0.001	-0.001	-0.003
	(0.004)	(0.004)	(0.003)
Salary increase & equity decrease	0.001	-0.004*	-0.002
	(0.003)	(0.003)	(0.003)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
R-squared	0.264	0.365	0.135
N	1,022	1,003	984
Panel C: Patent filings increase			
Standalone bonus increase	0.053	0.052	0.056
	(0.051)	(0.051)	(0.051)
Overall compensation increase	0.004	0.013	-0.039
	(0.085)	(0.086)	(0.086)
Bonus increase & equity decrease	-0.116	-0.056	-0.044
	(0.089)	(0.092)	(0.091)
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
N	5,022	5,022	5,022

Note: This table reports coefficients of OLS regressions in panel A and B and Probit regressions in panel C. Stand errors are heteroskedasticity robust. The dependent variables in Panel A are the numbers of product announcements normalized by industry average in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t. Specifically, we divide each firm's number of product announcements by the average amount of product announcements made in the same year by all firms that operate in the same industry. Industry classifications are based on the first two digits of SIC. Dependent variables in panel B are average abnormal return changes ±5 days before and after product announcements in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t. Abnormal returns are calculated by taking the residuals of the regression of the daily stock return on Fama-French three factors. Dependent variables in panel C are dummies indicating whether the number of patent filings has increased in year t+1, t+2, and t+3 in Column 1, 2 and 3 after compensation changes in year t.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

3.11 Appendix 1

Table A.1—Variable Definitions

Variables	Definitions
Age dummy	We create five dummies for CEO age under 45, between 45 and 50, between 50 and 55, between 55 and 60, and above 65
Analyst forecast STD	Information Asymmetry measures based on Bharath, Pasquariello, and Wu (2009) or similar. We use the standard deviation of EPS estimates scaled by the actual value.
Atwill exceptions	1 if the contract is governed by the law of a state with a good faith and fair dealing at-will exception
Busyboard	1 if the fraction of busy directors who are in more than 2 outside public boards over the number of independent directors is greater than 0.5
Cashflow/assets or sales	Cash flow over total assets or sales
CEO Age	Executive's age in years
Chairman CEO	1 if the CEO is also the chairman
CEO ownership	Percentage of firms' common stock owned by the CEO
CEO tenure	Number of years the CEO has been in office
Change in Employee number	Growth of number of workers
Change in staff expense	Growth of labor cost
Depr. and Amort. %	Depreciation and amortization as percentage of assets
Distress	Distress indicator based on Altman (1968)
Garmaise	Index of Garmaise (2006)
Gindex	The index is based on Gompers, Ishii, and Metrick (2003)
Idiosyncratic risk	Idiosyncratic risk based on Wurgler and Zhuravskaya (2002). We regress daily firm excess return on four factors and get the volatility of residuals.
Independent directors fraction	Percentage of independent directors on the board
Industry adjusted return	Log annualized return adjusted by industry average or median return (compounded)
Industry CEO turnover	Industry turnover ratio of CEOs based on the first two SIC
Industry homogeneity	Homogeneity of industry (Parrino 1997). We calculate the correlation between common monthly stock returns within two-digit SIC industries
Industry outside CEO	Industry ratio of outside CEOs based on the first two SIC (see definition of outside CEO below)
Leverage net	Debt minus cash over assets
Log assets	Log book assets (in $ millions)
Outside blockholder ownership	Percentage of shares held by the outside shareholders who held more than 5% of total number of shares outstanding
Outside CEO	1 if the CEO is hired from the outside or works in the firm for less than a year
Product Announcement	The number of product announcement in each year of each firm
R&D expenditure/sales	R&D expenditure as percentage of sales

Renewal	Indicator variable for CEOs who were in office at the time of the contract start
Return explained	The percentage of return that could be explained by market factor
ROA	Return on assets
ROE	Return on equity
Segment number	Number of business segments within a firm
Tenure group	Three dummies for a CEO who has worked in the same firm for at most 2 year, 3-6 years and more than 6 years
Total risk	Daily log stock return volatility per year

3.12 Appendix 2

Table A. 2 Examples of discretionary clauses

Panel A: Partly Flexible	
Upward flexible	shall be increased Be entitled to such increase May pay additional compensation to The board of directors may give consideration to increasing Such larger amount Subject to increase may be increased/may increase Be reviewed for possible increase Shall consider increasing In the form of an increase in salary
Lower bound	A minimum base salary of At least Not less than If so increased, the Regular Salary shall not be decreased to less than In no event ... be reduced to ...less than Not lower than
Downward rigid	Shall not be reduced
Panel B: Examples of conditional downward rigid	
Salary cut for other executives	All executive down by same percentage Reduction does not exceed that of the other Executives A salary reduction generally and ratably applicable to substantially all senior executives of the Company.
Salary cut for everyone	Cross-the-board reduction Ageneral salary reduction program for non-union employees and applicable to all officers
Consent	Written consent of the CEO for downward adjustment
Panel C: Examples of fully rigid	
Shall not be increased and shall not be reduced	
Panel D: Examples of fully flexible	
Subject to adjustment up or down May increase or decrease Will be adjusted Subject to adjustment	

3.13 Appendix 3 Principal component analysis

Panel A: Eigenvalue

Component	Eigenvalue	Difference	Proportion	Cumulative
	(1)	(2)	(3)	(4)
Comp1	3.526	0.908	0.122	0.122
Comp2	2.618	0.563	0.090	0.212
Comp3	2.055	0.267	0.071	0.283
Comp4	1.787	0.208	0.062	0.344
Comp5	1.580	0.080	0.055	0.399
Comp6	1.499	0.149	0.052	0.451
Comp7	1.351	0.089	0.047	0.497
Comp8	1.261	0.101	0.044	0.541
Comp9	1.161	0.042	0.040	0.581
Comp10	1.119	0.088	0.039	0.619
Comp11	1.031	0.014	0.036	0.655
Comp12	1.017	0.078	0.035	0.690
Comp13	0.939	0.017	0.032	0.722
Comp14	0.922	0.073	0.032	0.754
Comp15	0.849	0.007	0.029	0.783
Comp16	0.842	0.066	0.029	0.812
Comp17	0.776	0.116	0.027	0.839
Comp18	0.659	0.085	0.023	0.862
Comp19	0.575	0.041	0.020	0.882
Comp20	0.534	0.018	0.018	0.900
Comp21	0.516	0.011	0.018	0.918
Comp22	0.505	0.059	0.017	0.935
Comp23	0.446	0.042	0.015	0.951
Comp24	0.404	0.070	0.014	0.965
Comp25	0.334	0.052	0.012	0.976
Comp26	0.282	0.054	0.010	0.986
Comp27	0.228	0.042	0.008	0.994
Comp28	0.186	0.186	0.006	1.000

Panel B: Eigen vectors

	Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Comp8
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Entry compensation	Contract length	-0.041	-0.045	0.258	0.275	0.294	0.148	0.199	-0.266
	Entry salary to industry	-0.062	-0.025	0.072	0.209	-0.027	-0.085	-0.093	0.507
	Entry equity to industry	-0.023	-0.083	0.287	0.018	0.198	-0.383	0.111	0.229
	Entry bonus multiple to industry	-0.039	0.034	0.030	-0.138	0.274	0.297	0.234	0.150
	Entry PPS	0.048	-0.168	-0.038	-0.010	0.083	-0.137	0.332	0.296
Bonus clause	Participation in a firm-level bonus plan	0.124	0.000	-0.050	-0.315	0.264	0.104	-0.137	-0.236
	Explicit discretion	0.110	-0.045	0.014	0.217	-0.347	-0.146	0.210	-0.119
	Multiples of salary	0.191	0.086	-0.011	-0.134	0.124	-0.047	0.376	-0.219
	Given as a value	-0.008	-0.038	0.247	0.031	-0.389	-0.068	0.162	-0.102
	Functions of performance measures	-0.022	-0.067	0.312	0.240	0.214	-0.222	-0.150	-0.153
Equity clause	Future equity grant specified	0.158	-0.072	0.427	-0.303	-0.109	0.246	0.052	0.021
	Discretionary future equity grant	-0.027	-0.059	0.196	0.396	0.226	0.102	0.184	-0.270
	Equity grant as a function of salary	0.089	-0.050	0.383	-0.303	0.007	-0.084	-0.131	-0.060
	Equity grant as a function of performance	0.030	0.006	0.068	-0.124	-0.183	0.093	-0.053	-0.320
	Have vest information	0.060	-0.059	0.267	-0.193	-0.100	0.363	0.204	0.303
Flexibility clause	No flexible clause	0.367	0.016	0.040	0.190	0.089	0.135	-0.183	0.158
	Upcan clause	0.370	0.022	0.066	0.158	0.111	0.007	-0.103	0.135
	Lower bound clause	0.035	-0.063	0.108	0.196	0.073	0.407	-0.390	0.098
	No cut clause	0.255	-0.070	0.110	0.277	-0.165	0.093	0.100	-0.014
Review clause	Review requirement	0.443	0.050	-0.102	-0.103	0.016	-0.113	-0.062	-0.040
	Review annual clause	0.394	-0.050	-0.040	-0.001	-0.052	-0.131	-0.008	-0.052
	Review party - Compensation committee	0.301	-0.031	-0.015	0.059	-0.258	0.018	-0.054	-0.017
	Review party -Board	0.292	0.004	-0.274	0.035	0.173	-0.037	0.091	0.004
	Review party - Human resource committee	0.049	-0.048	0.269	-0.204	0.156	-0.429	-0.210	0.018
Review factor	Factor CEO performance	0.124	0.191	-0.102	-0.073	0.315	0.026	0.157	0.055
	Factor financial condition	0.002	0.577	0.117	0.053	-0.034	-0.017	-0.059	0.041
	Factor market condition	-0.016	0.456	0.111	0.029	-0.059	-0.032	0.122	-0.002
	Factor firm performance	0.002	0.577	0.117	0.053	-0.034	-0.017	-0.059	0.041

Note: This table presents the results of principal component analysis of contract clauses. Eigenvalues for each principal component are shown in Column 1 of Panel A. Difference, proportion of variance explained and cumulative proportion of variance explained are shown in Column 2, 3 and 4 respectively. Panel A lists the eigenvectors and the loading on each contract clauses.

3.14 Appendix 4: Selection into a Contract

To control for the selection bias arising from this non-random exclusion, we follow the approach of Heckman (1979) and use the choice regression described below to compute the Mills ratio.

We choose a state law characteristic for the identifying restriction: the at-will exception rule of good faith and fair dealing (here forth "exception rule"). This state-wide rule prohibits terminations made in bad faith or motivated by malice.[1] This rule protects rank-and-file employees with shorter contracts or without contracts, which makes such forms of employment more attractive. The ensuing popularity of shorter contracts makes it difficult for executives to negotiate longer contracts for themselves.

The direct judicial consequences of the rule to CEOs are likely to be limited, however, since they are protected by individual contracts. The listing of these so-called at-will exceptions is reported in Table A.2 as in Walsh and Schwarz (1996) and Muhl (2001). In most states, the rules were adopted between 1960 and 1980, following debates that were driven by political sentiments of that time as well as the particularities of isolated precedent cases.

To ensure that geographical effects are due to the at-will exceptions and not to other legal differences across states, we control for other geographical indexes such as the anti-takeover index of Bertrand and Mullainathan (1999) and the anti-competition enforceability index of Garmaise (2011). All regressions contain industry and year fixed effects to control for exogenous shocks to the labor market.

We run Probit regressions of contract disclosure and results are reported in Table A.3.

[1] There are two other exceptions that are less relevant for us. Under the public policy exception, dismissal is not allowed if it violates the state's public policy or a statute. Under the implied contract exception, an employee can dispute his/her dismissal if he/she can prove the existence of an implicit (i.e., not written) contract.

Table A.3 At-will exceptions

		At-will exceptions					
Code	State	Public policy	Implied contract	Good faith and fair dealing	Garmaise	Anti-takeover	Patents
AL	Alabama	0	1	1	5	0	9,017
AK	Alaska	1	1	1	3	0	1,075
AZ	Arizona	1	1	1	3	1	27,065
AR	Arkansas	1	1	0	5	0	3,867
CA	California	1	1	1	0	0	303,592
CO	Colorado	1	1	0	2	0	31,339
CT	Connecticut	1	1	0	3	1	45,008
DC	District of Columbia	1	1	0	6	0	1,576
DE	Delaware	1	0	1	7	1	10,827
FL	Florida	0	0	0	9	0	55,303
GA	Georgia	0	0	0	5	1	23,774
HI	Hawaii	1	1	0	3	0	1,946
ID	Idaho	1	1	1	6	1	14,903
IL	Illinois	1	1	0	5	1	92,974
IN	Indiana	1	0	0	5	1	33,766
IA	Iowa	1	1	0	6	0	13,330
KS	Kansas	1	1	0	6	1	9,086
KY	Kentucky	0	1	0	6	1	9,738
LA	Louisiana	0	0	0	4	0	11,803
ME	Maine	0	1	0	4	1	3,099
MD	Maryland	1	1	0	5	1	29,470
MA	Massachusetts	1	0	1	6	1	69,616
MI	Michigan	1	1	0	5	1	82,589
MN	Minnesota	1	1	0	5	1	48,550
MS	Mississippi	1	1	0	4	0	3,597
MO	Missouri	1	0	0	7	1	20,864
MT	Montana	1	0	1	2	0	2,623
NE	Nebraska	0	1	0	4	1	4,697
NV	Nevada	1	1	1	5	0	5,591
NH	New Hampshire	1	1	0	2	0	10,766
NJ	New Jersey	1	1	0	4	1	95,136
NM	New Mexico	1	1	0	2	0	6,345
NY	New York	0	1	0	3	1	139,544
NC	North Carolina	1	0	0	4	0	31,587
ND	North Dakota	1	1	0	0	0	1,603
OH	Ohio	1	1	0	5	1	83,265
OK	Oklahoma	1	1	0	1	0	16,955
OR	Oregon	1	1	0	6	0	23,386
PA	Pennsylvania	1	0	0	6	1	84,618
RI	Rhode Island	0	0	0	3	1	6,413
SC	South Carolina	1	1	0	5	1	12,229
SD	South Dakota	1	1	0	5	1	1,385
TN	Tennessee	1	1	0	7	1	17,301
TX	Texas	0	0	0	3	0	106,463
UT	Utah	1	1	1	6	0	12,413
VT	Vermont	1	1	0	5	0	5,613
VA	Virginia	1	0	0	3	1	23,797
WA	Washington	1	1	0	5	1	32,901
WV	West Virginia	1	1	0	2	0	4,321
WI	Wisconsin	1	1	0	3	1	36,818
WY	Wyoming	1	1	1	4	1	1,282

Note: This table presents the at-will exceptions, anti-takeover regulations, the Garmaise (2011) index, and the number of patents issued between 1977 and 2004 by state.

Table A.4 First stage

	Dependent variable	Contract
Geography	At-will exceptions	0.035
		0.0545
	Garmaise	-0.018*
		0.0102
Disclosure quality	Restatements	0.056
		0.0937
	Assets	-0.008
		0.0155
Governance	Renewal	-1.430***
		0.0467
	Gindex	0.033***
		0.0121
Risk	Analyst forecast STD	0.03
		0.058
	Industry homogeneity	-0.73
		1.99
Control variables	Tenure dummy	Yes
	Age dummy	Yes
	Year fixed effects	Yes
	Industry fixed effects	Yes
	N	7804

Note: This table presents marginal effects from Probit regressions and standard errors that are heteroskedasticity robust. The dependent variable is has contract, a dummy equal to 1 if the CEO has a disclosed contract and zero otherwise.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.