Explaining Downward-rigid CEO Compensation: An Information Asymmetry Perspective

Yiqing Lü*

July 2017

Abstract: CEO compensation rarely gets cut, and almost every component increased in early 2000. I consider a two-period contracting problem in which a board is privately informed of its CEO's matching quality with the firm. The board faces a trade-off: Revealing good information makes the CEO work harder, but it is costly. To save the information revelation cost in the earlier period, the board commits to a back-loaded compensation plan that features only upward adjustments in fixed and performance-based pay. This paper also considers extensions in which CEOs have transferable skills and sheds light on bonus caps and compensation disclosure policies.

Key words: CEO compensation, informed principal, signaling, matching quality, bonus cap, compensation disclosure

JEL codes: G38, J24, J31, M12

*New York University Shanghai, yiqing.lu@nyu.edu. This paper is a chapter of my dissertation and was previously circulated under the title "Asymmetric Contractual Revision and Compensation Structure". I would like to thank Ulf Axelson, Mike Burkart, Maria Bustamante, Vicente Cuñat, Amil Dasgupta, Jason Donaldson, Mara Faccio, Daniel Ferreira, Igor Makarov, Daniel Paravisini, Ailsa Röell, Jesus Salas, David Webb, Moqi Xu, Kathy Yuan, Konstantinos Zachariadis, Jing Zeng and audiences at the European Finance Association 2014 Doctoral Tutorial, the Financial Management Association 2014 conference, IÉSEG, LSE and SFI for helpful comments. All remaining errors are my own.

During the 1990s and early 2000s, the total compensation of chief executive officers (CEOs) rose rapidly. Two phenomena have drawn researchers' attention. First, compensation for US CEOs was relatively flat in the decades leading up to the 1990s but increased dramatically during the 1990s and early 2000s (Frydman and Jenter, 2010). This sharp break occurred during the tech boom. Second, almost every component of executive compensation, including salary and incentive pay, went up (Shue and Townsend, 2017). Cash compensation was not adjusted downward to offset the dramatic rise in long-term incentive pay. While the extant literature (e.g., Harris and Holmstrom 1982) could justify downward rigidity in total compensation, it did not fit well with executive compensation data in which each component increased.

This paper blends a dynamic informed principal model with a standard moral hazard model to understand the changing patterns of CEO pay. The starting point of this paper is to observe that two characteristics plague the design of executive compensation, moral hazard and information asymmetry, especially during the tech boom. While moral hazard is extensively studied in the literature, the combination of these two problems is under-explored. Investigating the combined effect is crucial to understanding the behavior of CEO compensation during the tech boom, as this period was marked by substantial high-tech investment and disruptive technological progress that increase uncertainty in the market environment. Thus, apart from objective performance metrics, firms may have increasingly relied on private information to assess the (mis)match between the CEO's skills and the needs of the firm during that period.

The basic set-up of the model is as follows. The board (the principal, she) has private information about the matching quality between the firm and the CEO (the agent, he), which can be either high or low. The principal's private information arrives sequentially at the beginning of each period over two periods. The agent is risk neutral and needs to exert private effort in order to produce some output in each period. The matching quality is complementary with the agent's effort. In this paper, the principal can offer the agent a contract in order to motivate him to work. The general insight to be drawn from the model is that information asymmetry could be key to explaining recent patterns in CEO compensation. In addition to the traditional role of providing incentives, compensation plays another role: transferring information from the principal to the agent. In other words, the compensation structure serves as a means for the principal to convey her private information to the agent.

Several studies show that the board acquires and possesses private information about CEO performance. For example, Cornelli, Kominek, and Ljungqvist (2013) find that the board collects soft information to evaluate whether the CEO is a good match for the firm and upon which the board makes firing decisions. According to some recent surveys (Casal and Caspar, 2014; Larcker, Saslow, and Tayan, 2014), directors are well aware of the strengths and weaknesses of CEOs[1] and possess a great deal of expertise in analyzing the business environment due to their experiences serving on multiple boards in various industries[2]. Hermalin (1998) argues that leaders, due to their superior information, face a temptation to mislead their followers; thus, they must sacrifice or set an example (a costly action) in order to credibly signal their private information.

In my model, the principal faces a trade-off. Revealing high matching quality to the agent makes him work harder, as he would realize that his productivity is higher than he initially perceived it to be. The principal, however, needs to incur a cost of information revelation in order to convince the agent. In a dynamic setting, to save the first-period information revelation cost, the principal with good information commits to a back-loaded compensation plan that consists of non-decreasing salary and non-

[1] According to a McKinsey report Casal and Caspar (2014), "Boards need to look further out than anyone else in the company,." The chairman of a leading energy company commented, "there are times when CEOs are the last ones to see changes coming". In a survey by Larcker, Saslow, and Tayan (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 the remainder (11.4%) understand them "slightly well" or "not at all well".

[2] According to McKinsey Quarterly from 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.

decreasing performance-based pay. In other words, the compensation structure is used to provide signals to the agent.

I first analyze four benchmark cases: a one-period model with and without information asymmetry, a two-period model without information asymmetry, and a two-period model with information asymmetry but no commitment. To facilitate the illustration in the introduction, I only explain the equilibrium compensation structure when the principal is prevented from offering equity pay. Nevertheless, the mechanism holds in both cases, as one can see in the section that allows for equity payment. In the two one-period benchmark cases, I find that if the production technology satisfies a certain condition that regulates the complementarity between the matching quality and the effort, the principal fully relies on fixed pay (or salary) to credibly communicate her private information to the agent. Performance-based (bonus) pay is set at a level as if the agent knew the matching quality. This is because, under this condition, signaling via bonus pay would involve sharing too much profit with the agent, outweighing the cost of signaling via salary. In other words, the principal does not use bonus pay to signal her private information, only to provide incentives. If the production technology does not satisfy the condition, the principal also uses bonus pay to provide signals, leading to either over- or under-provision of effort compared to the effort level under symmetric information.

The paper then considers a two-period model in which the matching quality may deteriorate in the second period. I choose a production technology that satisfies the aforementioned condition so that one can assign the signaling role to the salary and the incentive role to the bonus in a one-period model. Interestingly, the result that the bonus does not provide private information does not hold in a two-period setting. In other words, this technology allows me to attribute any increase in bonus pay in a two-period model over the amount paid under a one-period model to the signaling role, thus greatly simplifying the interpretation of my analysis.

Specifically, I consider the third benchmark case where information is symmetric in a two-period model. While the bonus is the same as in the one-period model, no salary is paid, as signaling is not needed. I then consider the fourth benchmark case in which information is asymmetric but commitment is impossible. In this case, the principal does pay salaries to provide signals. The equilibrium contract, however, is stationary in the sense that second-period compensation does not depend on first-period private information. It is a combination of two independent one-period contracts, which correspond to those offered under information asymmetry in a one-period model. Because 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 a package consisting of a higher bonus in the future but a lower salary today.

I then allow for commitment in the main analysis. In equilibrium, the principal pays either a higher salary or a higher bonus in the second period in exchange for a smaller payment in the first period in order to provide the first-period signal. In fact, the principal commits to a compensation schedule that consists of two basic compensation units. If the matching quality remains high in the second period, the principal will choose the unit that provides a larger salary (or at least not less) in the second period. If the matching quality declines, the principal will choose the unit that provides greater incentive pay. In other words, the principal uses the structure of compensation as a signal by allocating the signaling cost from the first period to the second period. Intuitively, such an arrangement is less costly for a principal with a higher matching quality to offer. For example, offering a higher bonus to the agent in the second period is costly, as the principal has to share more profit with the agent, but it induces greater effort from an agent with higher matching quality. The equilibrium contract therefore exhibits downward rigidity in terms of both salary and performance-based pay.

My model generates a rich set of empirical predictions. First, back-loaded long-term contracts are more likely to be observed in positions that require soft skills that are hard to quantify, for instance, R&D-oriented jobs and leadership positions. Second, signaling through salary is more likely if the matching quality improves, whereas signaling through incentive pay is more likely if the matching quality deteriorates. Third, the principal is more likely to offer a contract with high performance sensitivity to an agent with stronger bargaining power, as the principal has to resort to more costly signaling via bonus pay to retain the agent.

Internet technological innovations greatly expanded product markets and imposed new challenges on managers to be able to work in a more diverse environment (Murphy and Zábojník, 2007). I consider an extension in which the agent possesses transferable skills in the sense that the agent's outside option value depends on the matching quality. My model finds that the principal would offer a higher bonus instead of a higher salary to retain an agent who possesses highly transferable skills. However, if the agent's outside option value exceeds a certain level, the principal would not provide a signal, as attracting such an agent would imply too high a signaling cost.

This extension also sheds light on disclosure polices related to executive pay. Mandatory disclosure of CEO compensation helps transmit boards' private information to the market, thus leading to increased competition for talented executives. Thus, the cost of providing signals to the agent increase with an increase in the CEO's outside option value. A disclosure policy may thus discourage the board from offering costly compensation plans to CEOs in order to provide signals. One implication of disclosure policies is that it could moderate executive pay. It also helps explain why instituting disclosure of the value of the CEO's option grants was followed by moderation in executive pay in the late 2000s (Shue and Townsend, 2017).

This paper also sheds light on the attempts of recent regulations to curb managerial compensation, for example, bonus caps.[3] I show that if the production function exhibits a certain form of complementarity between matching quality and effort, the

[3] EU regulators have decided to institute bonus caps in order to rein in executive compensation. These policies came into effect at the beginning of 2014. Under such a policy, certain bankers can only be paid bonuses equal to their annual salaries or twice as much, if their firm obtains approval from shareholders.

principal would increase both bonus and salary in response to the introduction of bonus caps as a way of providing signals while maintaining a low bonus-to-salary ratio. The policy, however, also increases the signaling costs of firms and may discourage principals from informing agents of their matching quality.

Related Literature. I mainly summarize two strands of the literature on managerial compensation. One strand of the literature (MacLeod, 2003; Levin, 2003; Fuchs, 2015; Zábojník, 2014) studies hidden information, and the other considers moral hazard models (Baker, Gibbons, and Murphy, 1994; Bull, 1987; MacLeod and Malcomson, 1989).

First, this paper is intellectually indebted to the literature on informed principal models (Myerson, 1983; Maskin and Tirole, 1992). The general intuition of their models is that when a party that designs a contract has private information, the structure of the contract may reveal some of that information to other parties. Compared to Maskin and Tirole (1992), Myerson (1983) considers a more general setup in which agents also have private information and make private decisions. My paper extends their static setup to a dynamic one and generates new implications regarding compensation structure.

Specifically, my paper is related to the managerial compensation literature that builds upon informed principal models. MacLeod (2003) generalizes the logic of repeated game models by demonstrating that subjective schemes can be feasible even without infinite interactions as long as workers can punish a deviation from the implicit contract by imposing some type of socially wasteful cost on the employer. This model was further developed by Fuchs (2007) and Zábojník (2014). Fuchs (2015) shows that discretionary salary can be used as a signaling device, but that paper only considers fixed compensation and thus leaves aside the moral hazard problem. Zábojník (2014) incorporates the moral hazard problem into Fuchs (2015)'s model. In contrast with my model, the principal in Zábojník (2014)'s model receives a private signal about the agent's expected contribution to the value of the firm in the interim period and can thus provide a contractible subjective evaluation to the agent. Neither paper considers a principal whose private information changes dynamically. In addition, I explore the roles of incentive pay in providing signals and incentives.

Second, this paper is related to the moral hazard literature. One strand of this literature studies optimal contracts when performance measures are observable to the principal and the agent but are not verifiable. It focuses primarily on how repeated interactions between the principal and the agent help overcome the reneging problem wherein the principal is tempted to underpay the agent in order to save on labor costs (Baker, Gibbons, and Murphy, 1994; Bull, 1987; MacLeod and Malcomson, 1989). Baker, Gibbons, and Murphy (1994) introduce a verifiable performance measure to study the interaction between an implicit bonus that is based on a non-verifiable performance measure and an explicit bonus that is contracted based on the verifiable performance measure. They show that depending on the value of the fall-back position after reneging on an implicit contract, the implicit bonus and the explicit bonus can be substitutes or complements. Combining moral hazard and learning, career concern models (Harris and Holmstrom, 1982; Gibbons and Murphy, 1992) find that increasing explicit incentives can be optimal as the implicit incentives decline over time. As a result, the equilibrium contract provides for wages that do not decline with age. Although these models predict increasing total compensation, they do not explain why every component of CEO compensation is downwardly rigid.

The other strand of the moral hazard literature studies dynamic contracting problems between a risk-neutral principal and a risk-averse agent. According to Lambert (1983), increasing explicit incentives provides insurance to the agent in order to reduce incentive cost. Rogerson (1985), among others, has found that when the principal can dictate the agent's consumption/saving decisions, the optimal consumption pattern tends to be front loaded. Sannikov (2008) and He (2012) extend previous discrete-time principal-agent models to continuous time and identify the conditions under which the optimal compensation process becomes back loaded. The literature, however, only derives qualitative predictions regarding total compensation.

The structure of this paper is as follows. Section 2 describes the model. Section 3 analyzes the one-period model. Section 4 considers the case in which the principal's private information dynamically changes over two periods. Section 5 discusses recent bonus cap regulations. Section 6 considers transferable skills and the implications of disclosure policies. The last section concludes.

1 The Model

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

1.1 Dynamic Environments

In each period, the market condition m_t can be in one of two possible states, $m_t \in \{h, l\}$. h (l) represents a good (bad) market condition.

At the beginning of the first period, the prior probabilities of m_1 being h and l are r and $1 - r$, respectively, and $0 < r < 1$. The market condition might change in the second period. With probability q, a good market condition remains good, $Pr(m_2 = h|m_1 = h) = q$, and $0 < q < 1$. With probability $1 - q$, a good market condition deteriorates, $Pr(m_2 = l|m_1 = h) = 1 - q$. Parameters r and q are known to both parties. The deterioration can be caused by increased competition in the product market. A bad market condition in the first period remains bad in the second period, $Pr(m_2 = l|m_1 = l) = 1$. This assumption greatly simplifies the analysis by reducing the number of states of the economy. It also provides a robust setting in which to study signaling: If a firm in the bad market does not mimic the firm in the good market, the agent will learn that the firm in the bad market will remain in the bad market in the second period. The firm in the bad market therefore has the strongest incentive to mimic, which gives the firm in the good market the strongest incentive to separate itself from the other firm.

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

1.2 Production Technology

The principal supervises the agent over two periods. The agent's output in each period y_t is verifiable and can take on two possible values, $y_t \in \{0, 1\}$. The probability of achieving output 1 is $p_t = P(\theta_t, e_t)$, a function of matching quality θ_t and the agent's private effort e_t. Depending on the market conditions, θ_t can take two values: θ_l if $m_t = l$, and θ_h if $m_t = h$. Assume that $0 < \theta_l < \theta_h \leq 1$ and $e_t \in [0, 1]$. Output in each period becomes observable only at the end of period 2 (Zábojník, 2014). This assumption simplifies the analysis, as the agent infers matching quality not from interim performance but from the contract offered by the principal.

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) $\frac{\partial^2 P}{\partial \theta \partial e_t} > 0$; and (4) $P(\theta_h, 1) \leq 1$. Feature (1) means that zero effort leads to zero output. Feature (2) means that the probability of achieving a high output increases with matching quality and effort. Feature (3) means that supermodularity exists between effort and matching quality. In other words, the marginal productivity of the agent's effort increases with the matching quality. The last assumption ensures that the maximum probability of achieving a high output does not exceed 1.

The principal may want to hire the agent even when the matching quality is low. First, the search cost of a high-quality match can be high, and the firm needs a stop-gap agent to work for the firm. Second, in my model, the principal could still make a positive profit by hiring an agent with low matching quality.

1.3 An Informed Principal

As noted in the introduction, the principal is better informed than the agent about matching quality for various reasons. By virtue of monitoring many inputs, a supervisor gains superior information about the worker's talents (Alchian and Demsetz, 1972).

At the beginning of each period, the principal privately receives a perfect signal η_t about the productivity in that period, $\eta_t \in \{l, h\}$. The agent, however, does not observe signals in either period.[4] The principal will decide whether to convey her private information to the agent at the beginning of the first period. Due to the non-observability of the signal to the agent, it is impossible to write a contract contingent on the signal.

1.4 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 expected profit after deducting the compensation paid to the agent.

The agent's effort cost function is $\psi(e_t)$ for either period. It is twice differentiable in e_t. Assume that $\psi(0) = 0$, and $\psi'(e) > 0$. The agent maximizes the expected compensation after deducting effort disutility. I further assume that $\frac{\partial^2 P}{\partial^2 e} - \psi''(e) < 0$.[5] The agent has zero initial wealth and is protected by limited liability. The agent's reservation utility is assumed to be zero for all θ.[6]

[4]Unlike the agent in MacLeod (2003) who receives a private signal, the agent in this model does not receive private signals to abstract from the problem of opinion disagreement between the principal and the agent.

[5]This assumption ensures that the second-order condition of the agent's utility is satisfied.

[6]In an extension, I analyze the case in which the agent's reservation utility is type dependent.

1.5 Contract

Here, I characterize the contracting space. Based on the private signal η_1, the principal offers the agent a contract $\mathcal{C}_{t=1}$ at the beginning of the first period. While output is contractible, neither the principal's private information nor the agent's effort is. The contract $\mathcal{C}_{t=1}$ is a subset of $\mathbb{R}^4_+$, $\mathcal{C}_{t=1} \subseteq \mathbb{R}^4_+$. It consists of a set of compensation plans. I call $c^i_{t=2}$, an element of $\mathcal{C}_{t=1}$, a compensation plan. Note that $\mathcal{C}_{t=1}$ may contain more than one compensation plan, i.e., $\mathcal{C}_{t=1} = \{c^i_{t=2}, i = 1, 2, ...n\}$, where $c_{t=2} = \{w(0,0), w(1,0), w(0,1), w(1,1)\}$. The principal pays the agent $w(0,0)$ if $(y_1, y_2) = (0,0)$, $w(1,0)$ if $(y_1, y_2) = (1,0)$, $w(0,1)$ if $(y_1, y_2) = (0,1)$, and $w(1,1)$ if $(y_1, y_2) = (1,1)$. Limited liability constraints imply that all payments are non-negative. The principal commits to the contract $\mathcal{C}_{t=1}$. After receiving the private signal η_2 at the beginning of the second period, the principal, at her sole discretion, chooses a single compensation plan $c^i_{t=2}$ from $\mathcal{C}_{t=1}$.

Because this setting involves a signaling problem, the payment scheme will fully reveal the principal's private information under a separating Perfect Bayesian Equilibrium (PBE). Also, because this PBE setting might have multiple equilibria, I apply Cho and Kreps (1987)'s Intuitive Criterion to refine separating PBEs. For the purpose of this analysis, I focus on separating PBEs, as they are the most interesting ones. I also prove that a pooling equilibrium does not survive the Intuitive Criterion.

1.6 Timing

Figure 1 presents the timeline. At the beginning of the first period, the principal is privately informed of the market condition m_1 and offers a contract $\mathcal{C}_{t=1}$ to the agent. The agent could leave or stay. If he leaves, he obtains a reservation utility of zero. If he accepts the contract, he exerts effort. At the beginning of the second period, after observing market condition m_2, the principal chooses a single compensation plan $c^i_{t=2}$ from contract $\mathcal{C}_{t=1}$ and offers it to the agent. Again, the agent could leave or stay. If

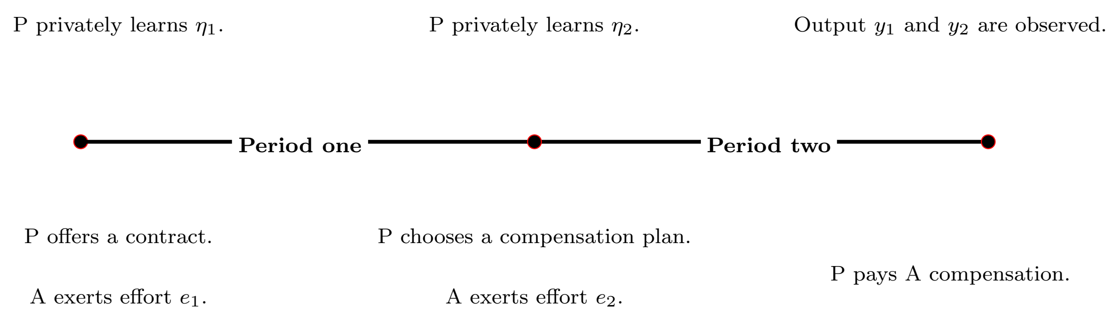

Figure 1: The Timeline

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

he leaves, he obtains a reservation utility of zero. If he accepts the compensation plan, he again exerts effort. At the end of the second period, the two parties observe the realized values of y_1 and y_2. Finally, compensation is paid.

2 A One-Period Model

Before I characterize the optimal contract in the two-period model, I analyze the one-period model. Figure 2 provides the timeline of the one-period model.

Denote fixed salary as f_1 and performance-based pay as (y_1). While f_1 is paid regardless of performance, $b_1(1)$ is paid when performance $y_1 = 1$ and zero otherwise. The subscript denotes one period. It is easy to show that the contract {w(.)} can be characterized by $\{f_1, b_1(y_1)\}$.[7] Specifically, $f_1 = w(0)$, $b_1(1) = w(1) - w(0)$, and $b_1(0) = 0$.

2.1 Symmetric Information

In this section, I analyze the first benchmark case in which the agent receives the the same signal as the principal does in a one-period model. In this benchmark case, both the principal and the agent are informed of the matching quality θ. I use the superscript s to denote compensation $\{f_1^s, b_1^s(y_1)\}$ under symmetric information.

[7]For the proof, please refer to Lemma 2.

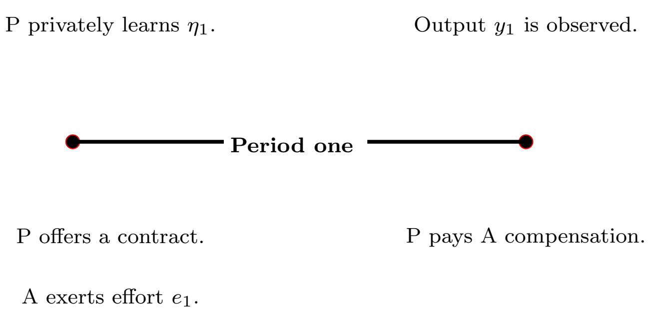

Figure 2: The Timeline

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

Because the agent is protected by limited liability, the principal cannot punish agent for poor performance. The principal thus chooses to pay the agent the minimum under a low output, that is, $b_1^s(0) = 0$ if $y = 0$. Due to the zero outside option value, the individual participation constraint will be automatically satisfied. Because both parties receive the signal η, there is no need for the principal to provide signal to the agent. As a result, paying f_1^s is not necessary, as it has neither incentive value nor signaling value.

I first analyze the agent's problem. Given the contract, the agent chooses the optimal effort level to maximize his utility:

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

Given the optimal level of effort $e^* = e(\theta, b_1^s)$ as a function of θ and b_1, the principal solves the following maximization program $P0$:

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

Constraint IC_a is the agent's incentive constraint obtained from his own maximization program. Constraint IR_a is the participation constraint of the agent. The limited liability constraint is satisfied if output is low ($y = 0$), as the objective function takes this into account.

The following equation is obtained based on the first-order condition of the principal's maximization program:

$$\frac{\partial P(\theta, e)}{\partial e} \frac{\partial e(\theta, b_1^s)}{\partial b_1^s}(1 - b_1^s) = P(\theta, e^*) \tag{2.1}$$

The left-hand side of Equation 2.1 measures the marginal benefit from an increase in the bonus: It indirectly leads to an increase in the probability of achieving high output through an increase in the agent's effort. The right-hand side of Equation 2.1 represents the marginal cost of an increase in the bonus: It directly increases the expected incentive cost. As argued at the beginning of this section, firms do not pay a salary under symmetric information. In this case, the bonus serves solely the role of incentivizing the agent. The following lemma characterizes the conditions under which the incentive effect becomes weaker or stronger as the production technology varies.

Proposition 1 *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* $b_1^{s,h} < b_1^{s,l}$.

Proposition 1 shows that linear supermodularity between matching quality and effort is insufficient to give rise to a bonus that increases in the matching quality. To induce a positive relationship, stronger supermodularity, specifically, positive log-supermodularity is required. I will revisit Proposition 1, as it has important implications for the subsequent analysis. Intuitively, if log-supermodularity is positive, the increase in the marginal benefit outweighs the increase in the marginal cost as the bonus increases.

2.2 Asymmetric Information with an Informative Bonus

This section studies the optimal contract for the second benchmark case: a one-period model with asymmetric information. A separating PBE is defined as follows:

Definition A separating Perfect Bayesian Equilibrium satisfies the following:

1. The principal that hires an agent with matching quality m offers a contract $\{f_1, b_1(y)\}$ that maximizes the firm's profit.

2. The agent's belief regarding the actual matching quality conditional on the contract offered is $\hat{q}(\eta = m | f_1, b_1(y)) = 1$.

3. Given the contract $\{f_1, b_1(y)\}$ and the belief, the agent chooses an effort level that maximizes his own utility.

I first analyze the agent's problem. Let $\hat{m}$ be the message that the principal sends to the agent via contract $\{f_1, b_1(y)\}$. The agent chooses an optimal level of effort e to maximize her utility given the contract:

$$\max_{e} P(\hat{m}, e)b_1 + f_1 - \psi(e)$$

From the first-order condition, we obtain the optimal level of effort, which is a function of $\hat{m}$ and b_1, $e^* = e(\hat{m}, b_1)$. That is, the effort level depends on the bonus b_1 and the agent's perceived matching quality or the message $\hat{m}$ sent by the principal.

Given the optimal effort level of the agent, a principal with high matching quality has the following maximization problem $P1$. Here, we solve for the separating equilibrium in which the contract offered by the principal sends a truthful message to the agent: $\hat{q}(m_h|f_1^h, b_1^h) = 1$, and $\hat{q}(m_l|f_1^l, b_1^l) = 1$.

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

$$s.t. \quad e^* = e(\theta_h, b_1^h) \qquad IC_a$$

$$P(\theta_h, e^*)b_1^h - \psi(e^*) + f_1^h \geq 0 \qquad IR_a$$

$$P(\theta_h, e(\theta_h, b_1^h))(1 - b_1^h) - f_1^h \geq P(\theta_h, e(\theta_l, b_1^l))(1 - b_1^l) - f_1^l \qquad for \quad IC_h$$

$$P(\theta_l, e(\theta_l, b_1^l))(1 - b_1^l) - f_1^l \geq P(\theta_l, e(\theta_h, b_1^h))(1 - b_1^h) - f_1^h \qquad for \quad IC_l$$

Constraint IC_a is the agent's incentive constraint obtained from her maximization program. Constraints IC_h and IC_l are truth-telling constraints for principals with high and low matching quality, respectively. These two constraints guarantee that the principal with high matching quality has no incentive to offer the contract offered by the principal with low matching quality, and vice versa.[8] If these constraints are satisfied, the contracts offered truthfully reveal the matching quality.

To solve for the equilibrium contracts, we first show that if Constraint IC_l is satisfied, then Constraint IC_h is automatically satisfied. Applying Cho and Kreps (1987)'s Intuitive Criterion, the least costly separating equilibrium is the one under which $f_1^l = 0$ and IC_l binds.

Lemma 1 *The principal's problem $P1$ is equivalent to the following maximization problem $P1'$:*

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

The optimal bonus under asymmetric information is thus given by the following equation:

$$\left(\frac{\partial P(\theta_h, e(\theta_h, b_1^h))}{\partial e} - \frac{\partial P(\theta_l, e(\theta_h, b_1^h))}{\partial e}\right)\frac{\partial e(\theta_h, b_1^h)}{\partial b_1^h}(1 - b_1^h) = P(\theta_h, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_h, b_1^h)) \qquad (2.2)$$

[8]Since the production technology exhibits supermodularity, the concavity of the principal's truth-telling constraint is guaranteed.

Similar to Equation 2.1, the left-hand side of Equation 2.2 measures the marginal benefit due to a unit increase in b_1 through an increase in the agent's effort after deducting compensation. The right-hand side of Equation 2.2 represents the associated marginal cost. In contrast with Equation 2.1, it is the sensitivity of output to private information that matters for the characterization of the optimal bonus level, which can be seen from the difference in the two partial derivatives of high and low perceived matching quality in Equation 2.2.

Intuitively, an agent, after receiving a better signal, would work harder, which leads to higher output. When deciding the optimal bonus, a principal with high matching quality would appropriate the profit derived from the agent's improved belief to the agent in order to prevent a principal of lower matching quality from mimicking the higher quality agent. The principal then maximizes profit after deducting the cost associated with signaling.

The following proposition characterizes the condition under which the bonus does not provide signal to the agent:

Proposition 2 ***Information Invariant Condition (IIC)***
If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} = 0$*, then the bonus is information insensitive,* $b_1^h = b_1^l = b_1^{s,h} = b_1^{s,l}$*. Only the salary provides a signal,* $f_1^h = (P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l)))(1-b_1^l)$*,* $f_1^l = 0$*.*

The IIC condition mutes any effects of information asymmetry on the bonus. The principal fully relies on the salary to provide the signal, while the bonus is paid as if the agent knew her own type (recall Proposition 1). The principal would not want to offer a higher bonus to substitute for the signaling role of the salary, because doing so would imply giving away too much profit.

By offering a salary equal to f_1^h, the principal credibly communicates her private information to the agent, which changes the agent's belief and motivates him to exert more effort. This channel is different from the incentive effect provided by bonus b_1. The salary affects the agent's effort by convincing the agent of his ability to achieve higher output when pay per unit of effort is held constant. The incentive channel affects the agent's effort level by raising b_1 when the agent's belief regarding matching quality is held constant. One direct implication of these two forces is that the salary paid by the principal with high matching quality is increasing in the matching quality of her agent θ_h. This is because the principal with low matching quality has greater potential to gain from mimicking the principal of higher matching quality. The principal with higher matching quality thus has to pay a higher salary to separate herself.

Corollary 1 ***Bonus Providing a Signal***

- *If* $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} > 0$*, then* $b_1^h > b_1^{s,h}$*, and*
  $f_1^h = (P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l)))(1 - b_1^h) - P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)$*;*
- *If* $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} < 0$*, then* $b_1^h < b_1^{s,h}$*, and*
  $f_1^h = (P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l)))(1 - b_1^h) - P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)$.

If the IIC condition is not satisfied, Corollary 1 shows that the bonus can also provide a signal. If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} > 0$, better matching quality improves the marginal productivity of effort in terms of the log-likelihood of high output. To prevent a principal with low matching quality from mimicking a principal with high matching quality, the principal with high matching quality pays a bonus that is higher than the level under symmetric information. This is because when the production function exhibits high complementarity between matching quality and effort, signaling through the bonus is cheaper for the principal with high matching quality, as mimicking would involve sharing too much profit with the agent for the principal of low matching quality. This result also features the overprovision of effort compared to the level under symmetric information due to the higher bonus offered. Similar to the case under the IIC condition, the salary can still provide a signal 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 role of salary in signaling is undermined.

If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} < 0$, Corollary 1 shows that the bonus is lower than the level under symmetric information. This result characterizes the condition under which there is underprovision of effort compared to level under symmetric information. In this case, complementarity between matching quality and effort is so insignificant that the firm finds it less costly to use underprovision of effort as a signal. This is because mimicking would induce too little effort for the principal of low matching quality.

3 A Two-period Model

I choose a specific production technology that satisfies the IIC condition of zero log-supermodularity between matching quality and effort in the one-period model: $P(\theta, e_t) = \theta e$. I further assume that the agent has a quadratic disutility function $\psi(e) = \frac{1}{2}e^2$.

Based on Proposition 1 and Proposition 2, one can easily verify that the optimal bonus offered with such a production function is information insensitive in the one-period model under both symmetric and asymmetric information. This production function greatly simplifies the interpretation of the analysis for a two-period model, because any subsequent changes that lead to a bonus that is different from the level under a one-period model is not due to a change in the matching quality but rather to long-term contracting. I will elaborate on this point later.

Lemma 2 *Contract* $\mathcal{C} = \{w_{00}^{hh}, w_{10}^{hh}, w_{01}^{hh}, w_{11}^{hh}; w_{00}^{hl}, w_{10}^{hl}, w_{01}^{hl}, w_{11}^{hl}\}$ *can be alternatively characterized by* $\mathcal{C} = \{f^{hh}, b_1^{hh}(y_1), b_2^{hh}(y_2), b_3^{hh}(y_1, y_2); f^{hl}, b_1^{hl}(y_1), b_2^{hl}(y_2), b_3^{hl}(y_1, y_2)\}$, *where the superscripts denote the matching quality over two periods. Here,* f *is the fixed compensation regardless of performance, and* $b_1(1) + f$ *is paid if* $y_1 = 1$ *and* $y_2 = 0$*;* $b_2(1) + f$ *is paid if* $y_1 = 0$ *and* $y_2 = 1$*;* $b_1(1) + b_2(1) + b_3(1, 1) + f$ *is paid if* $y_1 = y_2 = 1$*; and* f *is paid if* $y_1 = y_2 = 0$*.*

According to Lemma 2, a general contract can be characterized by a specification that consists of fixed pay and variable pay. Fixed pay does not depend on performance.

Variable pay is contracted upon different combinations of realized output measures. This specification offers a convenient interpretation of the compensation structure.

3.1 Without Cross-pledging

Long-term contracts are beneficial to the principal in two ways. First, she could signal her private information by using a bonus based on the second-period output measure. In other words, the principal's contracting space that can be used for signaling is expanded. Second, the principal could use cross-pledging to alleviate the incentive problem. Because the focus of the paper is on the signaling role of compensation, I first proceed by considering the case in which cross-pledging based on two period payoffs is not allowed, from which I obtain the main mechanism. No cross-pledging is perhaps an extreme case, but it corresponds to scenarios in which the firm cannot offer equity, for example, when shareholders are not satisfied with CEO compensation. For completeness, I then allow cross-pledging and characterize the optimal contract. Note that the main mechanism is robust to cross-pledging.

Now consider a third benchmark case in which information is symmetric in a two-period model. The optimal contract is obvious to derive. Because there is no need to provide signals due to symmetric information, salary is zero. Also, because cross-pledging is not allowed, the bonus for each type of matching quality in each period is the same as the amount provided in the one-period model.

I then analyze the fourth benchmark case in which information is asymmetric but committing to a long-term contract is impossible. The following lemma analyzes the contract in this benchmark case; it shows that the optimal contract is stationary in the sense that the bonus does not depend on underlying matching quality.

Lemma 3 *If committing to a long-term contract is impossible, the optimal contract can be characterized by two one-period contracts.*

- *The first one-period contracts for* $m_1 = h$ *and* $m_1 = l$ *at* $t = 1$ *are:*

For $m_1 = h$, $\{f_1 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_1(1) = \frac{1}{2}\}$.

For $m_1 = l$, $\{f_1 = 0, b_1(1) = \frac{1}{2}\}$.

- *The second one-period contracts for* $m_2 = h$ *and* $m_2 = l$ *at* $t = 2$ *are:*

  For $m_2 = h$, $\{f_2 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_2(1) = \frac{1}{2}\}$.

  For $m_2 = l$, $\{f_2 = 0, b_2(1) = \frac{1}{2}\}$.

According to Lemma 3, when commitment is impossible, private information in the first period does not affect the equilibrium contract in the second period. When new information arrives in the second period, the principal makes the same offer irrespective of the matching quality in the first period. Specifically, the contract for a low type in the second period who was a low type in the first period is the same as the contract for a low type who was a high type in the first period.

What if commitment is possible? Long-term contracts in this case depart from the short-term contracts in the sense that commitment allows the principal to reallocate signaling cost over two periods. In Lemma 4, when cross-pledging is not allowed, the principal cannot use equity compensation that is contracted upon both y_1 and y_2. Bonuses have to be contracted upon y_1 and y_2 separately and are positive in order to offer incentives. The agent's incentive problems in the two periods are tied only through the principal's truth-telling constraint, not through the incentive constraints. To satisfy the limited liability constraints, salaries are non-negative and may be positive to provide signals.

Lemma 4 *When cross-pledging is impossible, for an equilibrium contract* $\mathcal{C} = \{b\{.\}, f\{.\}\}$, $b_3^{hh}(1,1) = b_3^{hl}(1,1) = b_3^{ll}(1,1) = 0$. *To induce effort, the following components are greater than zero:* $b_1^{hh} = b_1^{hl} > 0$, $b_2^{hh}(1) > 0$, $b_2^{hl}(1) > 0$, $b_1^{ll}(1) > 0$ *and* $b_2^{ll}(1) > 0$. *In order to satisfy limited liability,* $f^{hl} \geq 0$ *and* $f^{hh} \geq 0$.

The principal can send a positive signal at the beginning of the first period in two ways. She could either pay a higher salary or promise more profit sharing through bonuses, even if matching quality deteriorates in the next period. The more profit sharing (the higher the bonus) the principal offers, the greater the aggregate welfare the contract could achieve due to greater incentive effect. The proposition below presents the optimal contracts under a separating equilibrium that survives the Intuitive Criterion, each of which is unique in a parameter range.

Proposition 3

- ***Low Separating Profit*** *($\theta_h < 2\theta_l$). The principal of matching quality θ_h commits to a contract in the first period: $\{b_1^{hh} = \frac{1}{2},\ b_2^{hh} = \frac{1}{2},\ f^{hh} = \frac{1}{4}\theta_l\theta_h - \frac{1}{4}\theta_l^2(2 - \frac{\theta_h}{\theta_l});\ b_1^{hl} = \frac{1}{2},\ b_2^{hl} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1}),\ f^{hl} = 0\}$. The principal will offer two one-period contracts to an agent of matching quality θ_l in the first period, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.*

- ***High Separating Profit*** *($\theta_h \geq 2\theta_l$). The principal of matching quality θ_h commits to a contract in the first period: $\{b_1^{hh} = \frac{1}{2},\ b_2^{hh} = \frac{1}{2},\ f^{hh} = \frac{1}{2}\theta_l(\theta_h - \theta_l);\ b_1^{hl} = \frac{1}{2},\ b_2^{hl} = 1,\ f^{hl} = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)\}$. The principal will offer two one-period contracts to an agent of matching quality θ_l in the first period, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.*

According to Proposition 3, if $\theta_h < 2\theta_l$, $b_2^{hl} > \frac{1}{2}$. In fact, paying a salary at the end of the second period when matching quality deteriorates is not renegotiation proof. When good private information arrives in the second period, the principal wants to further signal by paying a high salary. However, when bad information arrives, a long-term contract with $f^{hl} > 0$ is subject to renegotiation, as salary does not have either incentive or signaling value. Anticipating this, the principal offering a renegotiation-proof contract in the first period would commit to a contract that consists of two compensation plans: one offers a high salary f^{hh}, the other substitutes the salary ($f^{hl} = 0$) with a higher bonus b_2^{hl} in the second period. The bonus is one if $\theta_h = 2\theta_l$. Depending on the signal in the second period, the principal chooses one of the two compensation plans.

If the matching quality continues to be good, the principal chooses the contract that offers a high salary, as the salary has signaling value. Figure 3 depicts the equilibrium

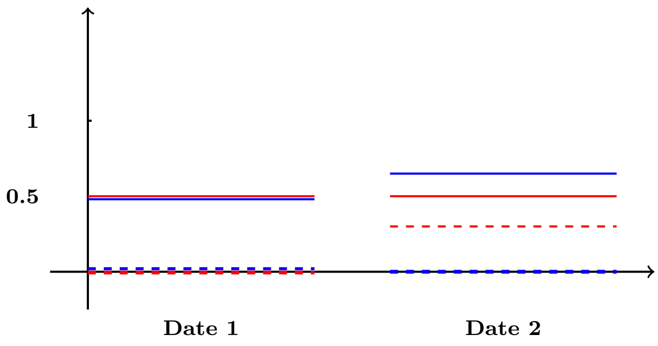

Figure 3: Equilibrium Contracts under Low Separating Profit

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

contracts. One can see that the contract is downwardly rigid in the bonus for both types, but the salary is zero if matching quality deteriorates.

If $\theta_h \geq 2\theta_l$, the mimicking incentive for the agent is greater than if $b_2^{hl} > \frac{1}{2}$; thus, the principal would have a stronger incentive to separate. The principal pays the agent a salary if she receives bad private information in the second period. As a result, if the principal receives good information in the first period, she would not want to renegotiate the positive salary and offer a higher bonus to reduce the signal cost as opposed to the case of low separating profit ($b_2^{hl} > \frac{1}{2}$). Instead, the principal's salary provides a signal when matching quality remains good and when it deteriorates.

Specifically, the renegotiation-proof contract offers a salary $f^{hl} = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)$ when matching quality deteriorates. Intuitively, the greater the mimicking incentive (i.e., the higher $(\theta_h - 2\theta_l)$), the greater the salary f^l. The salary can be paid in the following way: $\frac{1}{8}\theta_l(\theta_h - 2\theta_l)$ in the first period and the same amount in the second period. The salary is zero if $\theta_h = 2\theta_l$. The bonus is set to the optimal level 1, which maximizes effort.

If matching quality continues to be good, the principal chooses a contract that offers a higher salary to send a stronger signal. Figure depicts this implementation. One can see that the contract is downwardly rigid in both salary and bonus. Proposition 3 further implies that such a downwardly rigid contract exists only when there is enough

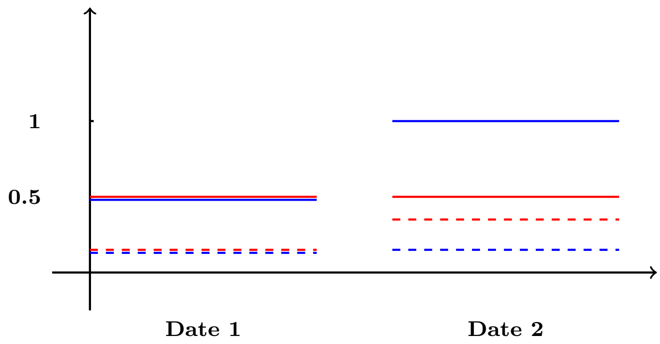

Figure 4: Equilibrium Contracts under High Separating Profit

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

variation in matching quality or when the mimicking incentive is substantial.

In both cases, the principal pays a higher rent to the agent in the form of a higher bonus based on y_2 compared to the case in which commitment is impossible. Recall Lemma 3. The signaling cost in each period is $\frac{1}{4}\theta_l(\theta_h - \theta_l)$, and the bonus is $\frac{1}{2}$. With long-term contracts, the principal is able to reallocate the cost of signaling from the first-period salary to the second-period bonus. Specifically, under low separating profit, the principal pays zero f_2^{hl}, and under high separating profit, the principal pays $f^{hl} = \frac{1}{4}\theta_l(\theta_h - 2\theta_l) < \frac{1}{4}\theta_l(\theta_h - \theta_l)$, both of which are smaller than $\frac{1}{4}\theta_l(\theta_h - \theta_l)$. In both cases, $b_2^{hl} > \frac{1}{2}$. This signaling approach requires greater profit sharing with the agent in the second period and thus a lower salary, explaining downward rigidity in both salary and bonus.

In addition, the principal with constant good matching quality has to send a stronger signal than the principal with deteriorating matching quality. The former offers a high salary f^{hh} but the latter offers a high bonus b_2^{hl}. The intuition is that a salary is a more costly signal, because it has zero incentive value. The agent will still have to exert effort in order to obtain the bonus, which allows the firm to recoup at least some of the profit.

For completeness of the analysis, I prove in the following corollary that the pooling equilibrium does not survive the Intuitive Criterion. The intuition is that a principal with high matching quality would always benefit from a deviation by offering a certain salary that is equilibrium dominated for the principal of low matching quality.

Corollary 2 *If cross-pledging is not allowed, the pooling equilibrium does not survive the Intuitive Criterion.*

3.2 With Cross-pledging

The previous section analyzes the optimal contract when cross-pledging is not allowed. The contract features downward rigidity in salary and bonus. In this section, I complete the analysis by considering cross-pledging under which incentives are provided via equity compensation. The principal would use b_3^{hh}, b_3^{hl} and b_3^{ll} to alleviate the incentive problem. In other words, the principal could use equity compensation that vests at the end of the second period. By shirking in one period, the agent reduces the probability of full success and, consequently, the reward for the effort exerted in the other period.

As in the previous sections, I first consider the contract under symmetric information. Because signaling is not needed, a salary is not paid.

Lemma 5 *If information is symmetric, the principal offers the following contracts in equilibrium:*

- *Principal with matching quality hh offers* $\{b_1^{hh} = 0, b_2^{hh} = 0, b_3^{hh} = \frac{1}{\theta_h^2}, f^{hh} = 0\}$;
- *Principal with matching quality hl offers* $\{b_1^{hl} = 0, b_2^{hl} = 0, b_3^{hl} = \frac{1}{\theta_l \theta_h}, f^{hl} = 0\}$;
- *Principal with matching quality ll offers:* $\{b_1^{ll} = 0, b_2^{ll} = 0, b_3^{ll} = \frac{1}{\theta_l^2}, f^{ll} = 0\}$.

Lemma 5 shows that when cross-pledging is allowed, the principal uses performance-based pay contracted on two output measures to induce effort. In this way, the principal minimizes the rent the agent extracts due to limited liability. In Proposition 4, I characterize the optimal contracts under information asymmetry.

Proposition 4 • *The principal of high matching quality commits to the following contract in the first period:* $\{b_1^{hh} = 0, b_2^{hh} = 0, b_3^{hh} = \sqrt{\frac{1-(1-q)\theta_l^2}{q\theta_h^4}}, f^{hh} = f^{hl} + \theta_l\theta_h^2 b_3^{hh}(1-\theta_h b_3^{hh}); b_1^{hl} = 0, b_2^{hl} = \frac{3q\theta_l^2-2q^2\theta_l^2-q\theta_h+q^2\theta_h-\theta_l^2+1}{2q(1-q)\theta_l^2}, b_3^{hl} = \frac{1}{2\theta_l^2} - \frac{\theta_h^2\beta^{hh}b_3^{hh}}{2(1-q)\theta_l^2}, f^{hl} = 1-\theta_l+(\theta_h-\theta_l)\theta_l^2 b_3^{hl}\}$.

- *The principal of low matching quality offers the following contract in the first period:* $\{b_1^{ll} = 0, b_2^{ll} = 0, b_3^{ll} = \frac{1}{\theta_l^2}, f^{ll} = 0\}$.

In Proposition 4, compared to the optimal contract without cross-pledging, the bonus based on the first-period measure is zero and that based on the first- and second-period measure is positive, which reduces rent extraction by the agent. The principal with high matching quality only uses b_3^{hh} and f^{hh} to induce effort and signal her private information. The equilibrium contract is also downwardly rigid in salary and incentive pay, which is implied by the features below.

First, due to information asymmetry, the principal with deteriorating matching quality will offer the agent a larger bonus based on the second-period measure. This can be seen easily by comparing the contract in Proposition 4 with Lemma 5. The intuition is the same as in the case without cross-pledging. If $b_2^{hl} = 0$, the principal would 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_3^{hl} down and increase b_2^{hl} as a signal for the first period. Consequently, the compensation offered by a principal with deteriorating matching quality pays more compensation based on the second performance measure.

Second, unlike in the case of symmetric information, b_3^{hl} does not enter into the the principal's maximization function linearly. On the one hand, high b_3^{hl} leads to less rent extracted by the agent because of cross-pledging. On the other hand, it increases the mimicking profit of the principal with low matching quality in the first period, as it induces greater effort. Due to this trade-off, b_3^{hl} is smaller than under symmetric information.

Third, $b_3^{hh} > \frac{1}{\theta_h^2}$. In order to induce sufficient effort in the first period, the principal with constant high matching quality will offer higher long-term equity compensation to induce first-period effort, because the agent knows that if the matching quality declines, the principal will offer a bonus based on the second-period output measure. Expecting this, the first-period incentive of an agent with constantly high matching quality would be weakened if the principal did not raise b_3^{hh}. Thus, long-term equity compensation is also used to provide a signal.

These features imply that the equilibrium contract when cross-pledging is allowed is also downwardly rigid in salary and incentive pay. The principal with high matching quality in the first period can promise to pay at least $\frac{1}{2}f^{hl}$ in each period. If the matching quality improves, she pays more by $f^{hh} - f^{hl}$. The incentive pay offered by the principal with high matching quality in the first period is back loaded and puts more weight on the second-period output measure.

For completeness of the analysis, I also verify that the pooling equilibrium does not survive the Intuitive Criterion under cross-pledging.

Lemma 6 *If cross-pledging is allowed, the pooling equilibrium does not survive the Intuitive Criterion.*

4 Transferable Skills and Disclosure Policies

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 that other firms perceive the agent's skill to be firm-specific. Studies on executive compensation have long been interested in pay-performance sensitivity and its relation with human capital (Murphy and Zábojník, 2007; Dutta, 2008). Executives may possess firm-specific skills as well as transferable skills. This section extends the model by assuming a type-dependent reservation utility. If the agent's reservation utility is type contingent, his participation constraint may become binding. One implication is that the principal needs to provide higher compensation in order to retain the agent.

This extension also merits consideration in light of mandatory compensation disclosure policy. Current executive compensation disclosure requirements are applicable to most US domestic issuers and to non-US companies that do not qualify as foreign private issuers. These requirements were adopted by the US Securities and Exchange Commission (SEC) in 1992. The recent Dodd-Frank Wall Street Reform and Consumer Protection Act also contain new disclosure policies that affect the governance of issuers.[9] In this paper, the value of the outside option of the agent relies on whether the market believes that the agent has better skills. If firms are not required to disclose compensation, then the market will not know the agent's skills or will find it more costly to assess these skills. Zero reservation utility represents an extreme case in which the agent has only firm-specific skills or the market has no way through which to infer the agent's general skills. Thus, mandatory compensation disclosure may affect the compensation level and structure.

The timeline in this extension is the same as that in the one-period baseline model but differs from it in the reservation utility, which is R for the agent of high matching quality and 0 for the low matching quality. As previously, the agent exerts effort $e \in [0,1]$ with disutility $\psi(e) = \frac{1}{2}e^2$. At the end of date 1, the probability of obtaining $y = 1$ is $p = P(\theta, e) = \theta e$. It can be easily verified that the optimal effort of an agent with matching quality 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:

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

$$s.t. \quad \theta_l e^{l*}(1 - b_1^l) - f_1^l \geq \theta_l e^{h*}(1 - b_1^h) - f_1^h \qquad IC_p$$

$$\theta_h e_h^* b_1^h + f_1^h - \frac{1}{2} e_h^{*2} \geq R \qquad IR_a$$

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

If $R = 0$, as in the one-period benchmark model, Constraint IR_a is not strictly binding because the agent is protected by limited liability. However, if R is sufficiently large, the surplus the agent extracts due to limited liability may not be large enough to satisfy the constraint. Denote λ as the Lagrangian multiplier of Constraint IR_a. I consider the binding $\lambda > 0$. Under the least costly separating equilibrium, the principal pays $f_1^l = 0$ to an agent of low matching quality, since there is no signaling gain from motivating such an 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 Constraint IC_p binding. Thus, I substitute b_1^h obtained from Constraint IC_p into the objective function and Constraint IR_a.

Proposition 5 ***Compensation and Managerial Skills***

Assume $b_1^{o,h}$ is the bonus paid to an agent with zero reservation utility and high matching quality, and f_1^h and b_1^h with positive reservation utility and high matching quality.

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

- *If $\underline{R} < R \leq \bar{R}$, Constraint IR_a 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 a separating equilibrium exists.

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

Proposition 5 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 IIC 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 provide a signal to the agent. Instead, she chooses to pool with the principal with low matching quality.

When the reservation utility is at an intermediate level, the IR_a constraint binds. Having general skills implies higher performance sensitivity. One might set the bonus at $\frac{1}{2}$ and increase the salary so that Constraint IR_a binds. However, this is not optimal. To see the intuition, 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) implies that the marginal benefit relative to the marginal cost of setting the bonus to $\frac{1}{2}$ increases. Hence, increasing the bonus makes the IR_a constraint bind more easily. In other words, when the agent's skills become sufficiently transferable, compensation disclosure may result in high-powered incentives.

Mandatory disclosure policies thus have two effects. First, when skills are not sufficiently transferable, firms have to use higher performance sensitivity to provide signals in addition to a positive salary. Second, when skills are sufficiently transferable, the principal may choose not to provide a signal due to the excessive signaling cost. One implication of disclosure policy, therefore, is that they could moderate executive pay; this helps explain why the institution of disclosure of the value of CEO option grants was followed by a moderation in executive pay in the late 2000s (Shue and Townsend, 2017; Frydman and Jenter, 2010).

5 Bonus Caps and Efficiency Implications

A banker bonus cap was passed by the European Parliament (EP) in April 2013 and was to go into effect in January 2014. In February 2014, the EP and the European Council (the Council), agreed to restrict the bonuses of retail asset managers.[10]

In this section, I study the impact of bonus caps on the compensation structure

[10] The Council 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 the bonus amount must be paid in shares of the fund under management, and at least 40% of the bonus amount must be deferred for three years.

based on the baseline model. The timeline is the same as that of the one-period benchmark model. After observing the contract, the agent exerts 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 $y \in \{0,1\}$. In this section, I consider a special form of the production technology that does not conform to the IIC condition: $P(\theta, e) = \theta_i e(\theta_i + \frac{1}{2}ke)$. I use a production function with negative log-supermodularity to illustrate policy implications that may not be intended.

	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 1: A Comparison of Efficiency

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

In Table 1, I present a numerical example using this production function: $k = 1$, $\theta_h = 0.9$, $\theta_l = 0.4$. The ratio of bonus to salary cannot exceed 2.5939.[11] It provides a simple analysis of compensation and the welfare of the board and the CEO in three cases. The first column is the contract that a principal of high matching quality offers without a bonus cap. The second and third columns represent the contracts that the principal of high matching quality offers under a cap on the bonus-to-salary ratio. The second column presents 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.

By imposing a limit on the ratio of the bonus to salary, the regulator forces the board to adjust the bonus and, consequently, increase the salary, achieving a lower bonus-to-salary ratio. This can be seen by comparing the bonuses and salaries in the

[11] In this example, the bonus under symmetric information is 0.8969.

first and second columns. In other words, the principal has to pay a higher signaling cost. The table also shows that the contract in the second column yields greater efficiency at the expense of the principal. The board has to pay a greater signaling cost in order to abide by the rule. The agent benefits from the bonus cap.

One might argue that the board could consider increasing only the salary. As shown in the third column, such an approach may lead to greater profit destruction for the principal compared to the second approach, because signaling via salary alone is more costly, as explained in Corollary 1. In this case, the principal simply reallocates profit without an efficiency enhancement.

As shown in Figure 5, the ratio of bonus to salary is not monotonic in the bonus. The intuition is that salary equals the profit of the principal with low matching quality if she mimics minus the profit if she does not. When the bonus is high, an increase in the bonus implies more profit sharing than effort provision, leading to low mimicking profit and an increased bonus-to-salary ratio. When the bonus is low, an increase in the bonus implies greater effort provision than profit sharing, leading to high mimicking profit and a decreased bonus-to-salary ratio. To meet the bonus cap, the principal thus has to increase the bonus, because an increase in the bonus implies an even greater increase in the salary (or a decrease in the bonus implies an even greater decrease in the salary).

My model explains some of the incentives behind the measures taken by some banks. Some banks have restructured their CEO compensation by increasing the base salary, resulting in higher estimated total pay. For instance, HSBC's chief executive, Stuart Gulliver, received a salary increase from £1.2 million to £2.9 million thanks to a £32,000 weekly shares of "fixed pay allowance" in 2013.

Another possible consequence of this policy, especially for firms that operate in volatile environments, is that it may impose a high signaling cost. Bonus caps could exacerbate the information problem by making truth-telling more costly or even impos-

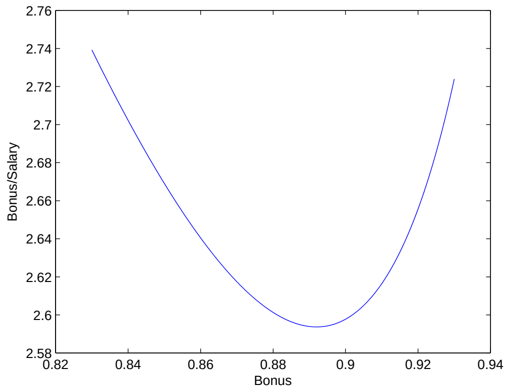

Figure 5: Bonus over Salary

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

sible. A principal of high matching quality may find it more profitable not to provide a signal and to pool with the other type of principal. In order to achieve efficiency improvement, determining the appropriate bonus-to-salary ratio is important. If it is too low, firms would find it difficult to motivate and retain talented executives.

Another implication of this extension is that the heterogeneous effects of bonus caps on firms with different technologies need to be taken into account. The overall effect of bonus caps on societal welfare depends on the distribution of different types of production functions. In some firms, managerial talent contributes considerably to the output function, while in other firms, it does not. Imposing bonus caps, however, may not improve the efficiency of the former.

6 Conclusion

This paper characterizes the optimal contract offered by a principal who knows more than the agent about matching quality. Contracts have two roles: providing signals and incentives to the agent. I show the conditions under which the principal solely relies on salary to signal her private information to the agent. Bonuses can also be information sensitive if the conditions are not met. Thus, bonuses could play a dual role by providing signals as well as incentives. When the bonus is used to provide a signal, the principal either uses profit sharing (high bonus) or underprovision of effort (low bonus) to signal her private information. I choose a specific production function that features zero log-supermodularity in matching quality and effort. Such a technology separates the signaling role and the incentive role; the contract under the separating equilibrium could thus assign the former role to the salary and the latter to the bonus in a one-period model.

In a two-period model, I first analyze a benchmark case in which the principal cannot commit to long-term contracts. Because the agent anticipates that the principal will make a new take-it-or-leave-it offer when new information arrives, he will not agree to an arrangement that promises a high bonus in future as a signaling device. The equilibrium contract is thus stationary in the sense that the second-period contract does not depend on first-period private information. If commitment is possible, the principal could promise a higher bonus based on the second-period performance measure as a way of providing a signal of high first-period matching quality. In other words, the principal pays more rent to the agent in the second period in exchange for less paid to the agent in the first period. If matching quality continues to be high, the principal wants to provide an even higher salary to provide a signal. Such a contract is non-decreasing in both salary and bonus. It achieves greater efficiency by giving the agent more profit-sharing opportunities and inducing more effort in the second period.

This paper also sheds light on disclosure policies and regulations targeted at managerial compensation. I first consider an extension in which the manager possesses general skills. It suggests that when managerial skills are sufficiently transferable, the principal may choose not to provide a signal due to the excessive signaling cost. I then consider regulations capping the bonus-to-salary ratio. I find that under some production functions, the principal may have to increase the bonus to meet the requirement, because an increase in the bonus implies an even greater increase in the salary. Bonus caps could exacerbate the information problem by making truth-telling more costly or even impossible.

Several important conclusions can be drawn from this paper. First, the bonus is sensitive not only to publicly observable information but also to private information. The mapping from objective performance measures to the bonus may contain the private information of the principal that is not observable to econometricians. Neglecting this channel might lead to overestimation of the incentive effect of performance-based pay. Second, salaries play a crucial role in facilitating communication in organizations, especially in cases where bonuses do not provide feedback.

References

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

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

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

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

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

Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist, 2013, Monitoring managers: Does it matter?, *The Journal of Finance* 68, 431–481.

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

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

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

———, 2015, Subjective evaluations: Discretionary bonuses and feedback credibility, *American Economic Journal: Microeconomics* 7, 99–108.

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

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

He, Zhiguo, 2012, Dynamic compensation contracts with private savings, *The Review of Financial Studies* 25, 1494–1549.

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

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

Larcker, David F., Scott Saslow, and Brian Tayan, 2014, 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.

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

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

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

Maskin, Eric, and Jean Tirole, 1992, The principal-agent relationship with an informed principal, ii: Common values, *Econometrica* 60, 1–42.

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

Myerson, Roger B, 1983, Mechanism design by an informed principal, *Econometrica* 51, 1767–1797.

Rogerson, William P, 1985, Repeated moral hazard, *Econometrica: Journal of the Econometric Society* pp. 69–76.

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

Shue, Kelly, and Richard R Townsend, 2017, Growth through rigidity: An explanation for the rise in ceo pay, *Journal of Financial Economics* 123, 1–21.

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

Appendix Proof

Proposition 1. 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 sets an optimal level of performance-based pay to maximize the expected profit, which is a function of the matching quality (θ), the performance-based pay ($b_1(\theta)$), and the optimal effort that the agent chooses given the matching quality and the performance-based pay ($e^*(\theta, b_1(\theta))$).

$$\max_{b_1} \quad u = U(\theta, e^*(\theta, b_1(\theta)), b_1(\theta))$$

The first order derivative is thus:

$$F.O.C. \quad \frac{\partial u}{\partial b_1} = \frac{\partial U}{\partial e}\frac{\partial e}{\partial b_1} + \frac{\partial U}{\partial b_1} = 0 \tag{7.1}$$

Take first order derivative of Equation 7.1 w.r.t. θ:

$$\frac{\partial^2 U}{\partial b_1 \partial\theta} + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial\theta} + (\frac{\partial^2 U}{\partial b_1\partial e}\frac{\partial e}{\partial b_1} + \frac{\partial^2 U}{\partial^2 b_1} + \frac{\partial^2 U}{\partial^2 e}(\frac{\partial e}{\partial b_1})^2 + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial b_1} + \frac{\partial U}{\partial e}\frac{\partial^2 e}{\partial^2 b_1}) \times \frac{db_1}{d\theta} = 0$$

Rearrange the equation, I obtain:

$$\frac{db}{d\theta} =(\frac{\partial^2 U}{\partial b_1 \partial\theta} + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial\theta})/(\frac{\partial^2 U}{\partial b_1\partial e}\frac{\partial e}{\partial b_1} + \frac{\partial^2 U}{\partial^2 b_1} + \frac{\partial^2 U}{\partial^2 e}(\frac{\partial e}{\partial b_1})^2 + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial b_1} + \frac{\partial U}{\partial e}\frac{\partial^2 e}{\partial^2 b_1}) \tag{7.2}$$

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

$$\frac{\partial^2 U}{\partial b_1\partial e}\frac{\partial e}{\partial b_1} + \frac{\partial^2 U}{\partial^2 b_1} + \frac{\partial^2 U}{\partial^2 e}(\frac{\partial e}{\partial b_1})^2 + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial b_1} + \frac{\partial U}{\partial e}\frac{\partial^2 e}{\partial^2 b_1} < 0$$

As a result,

$$\frac{db}{d\theta} = 0 \Leftrightarrow \frac{\partial^2 U}{\partial b_1 \partial\theta} + \frac{\partial e}{\partial b_1}\frac{\partial^2 U}{\partial e\partial\theta} = 0$$

From Equation 7.1, I obtain:

$$\frac{\partial e}{\partial b_1} = -\frac{\partial U}{\partial b_1}/\frac{\partial U}{\partial e}$$

Also because $U(\theta, e^*(\theta, b_1(\theta)), b_1(\theta)) = P(\theta, e)(1 - b_1(\theta))$, I obtain:

$$\frac{\partial U}{\partial b_1} = -P(\theta, e)$$
$$\frac{\partial U}{\partial e} = (1 - b)\frac{\partial P}{\partial e}$$

I therefore obtain:

$$P\frac{\partial^2 P(\theta,e)}{\partial e\partial\theta} - \frac{\partial P(\theta,e)}{\partial e}\frac{\partial P(\theta,e)}{\partial\theta} = 0$$
$$\Leftrightarrow \quad \frac{\partial^2 lnP(\theta,e)}{\partial e\partial\theta} = 0$$

One example of solutions to the above PDE is $P(\theta,e) = h(\theta)f(e)$.

It easily follows that if $\frac{\partial^2 lnP(\theta,e)}{\partial e\partial\theta} > 0$, $b_1^{s,h} > b_1^{s,l}$. And vice versa. Q.E.D.

Lemma 1. The principal's problem $P1$ is equivalent to the following maximization problem $P1'$:

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

Proof Substituting Constraint IC_a and Constraint IC_l into the objective function, problem $P1$ with two constraints then is simplified to problem $P1'$. Q.E.D.

Proposition 2. If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} = 0$, then bonus is information insensitive, $b_1^h = b_1^l = b_1^{s,h} = b_1^{s,l}$. Only salary provides signal, $f_1^h = (P(\theta_l, e(\theta_h, b_1^h) - P(\theta_l, e(\theta_l, b_1^l)))(1-b_1^l)$, $f_1^l = 0$.

Proof Following Lemma 1, set the principal's maximization objective as

$$\max_{b_1^h} \quad u = U(\theta_h, e^*(\theta_h, b_1^h), b_1^h) = P(\theta_h, e(\theta_h, b_1^h))(1-b_1^h) - P(\theta_l, e(\theta_h, b_1^h))(1-b_1^h)$$

Similar to the proof in Proposition 1, I obtain the derivative of bonus with respect to θ_h:

$$\frac{db_1^h}{d\theta_h} = 0 \Leftrightarrow \frac{\partial^2 U}{\partial b_1^h\partial\theta_h} + \frac{\partial e}{\partial b_1^h}\frac{\partial^2 U}{\partial e\partial\theta_h} = 0$$

Because from Equation 7.1, I obtain:

$$\frac{\partial e}{\partial b_1^h} = -\frac{\partial U}{\partial b_1^h}/\frac{\partial U}{\partial e}$$

Also because $U(\theta, e^*(\theta, b_1(\theta)), b_1(\theta)) = P(\theta_h, e(\theta_h, b_1^h))(1-b_1^h) - P(\theta_l, e(\theta_h, b_1^h))(1-b_1^h)$, I obtain:

$$\frac{\partial U}{\partial b_1^h} = -P(\theta_h, e) + P(\theta_l, e)$$
$$\frac{\partial U}{\partial e} = (1-b_1^h)\frac{\partial(P(\theta_h,e) - P(\theta_l,e))}{\partial e}$$

I therefore obtain:

$$(P(\theta_h,e) - P(\theta_l,e))\frac{\partial^2(P(\theta_h,e) - P(\theta_l,e))}{\partial e\partial\theta_h} = \frac{\partial(P(\theta_h,e) - P(\theta_l,e))}{\partial e}\frac{\partial(P(\theta_h,e) - P(\theta_l,e))}{\partial\theta_h}$$
$$\Leftrightarrow \quad \frac{\partial^2 ln(P(\theta_h,e) - P(\theta_l,e))}{\partial e\partial\theta_h} = 0$$

$P(\theta,e) = h(\theta)f(e)$ again is an example of solutions to the above PDE. Q.E.D.

Corollary 1.

- If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} > 0$, then $b_1^h > b_1^{s,h}$, and $f_1^h = (P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l)))(1 - b_1^h) - P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)$;
- If $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} < 0$, then $b_1^h < b_1^{s,h}$, and $f_1^h = (P(\theta_l, e(\theta_h, b_1^h)) - P(\theta_l, e(\theta_l, b_1^l)))(1 - b_1^h) - P(\theta_l, e(\theta_l, b_1^l))(b_1^h - b_1^l)$.

Proof Following the proof in Proposition 2, it can be easily shown that if $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} > 0$, then $\frac{db_1^h}{d\theta_h} > 0$, thus $b_1^h > b_1^l = b_1^{s,l} = b_1^{s,h}$. Similarly, if $\frac{\partial^2 ln(P(\theta,e)-P(\theta_l,e))}{\partial\theta\partial e} < 0$, then $\frac{db_1^h}{d\theta_h} < 0$, thus $b_1^h < b_1^l = b_1^{s,l} = b_1^{s,h}$.

I then obtain the amount of salary from Constraint IC_l. Q.E.D.

Lemma 2. Contract $\mathcal{C} = \{w_{00}^{hh}, w_{10}^{hh}, w_{01}^{hh}, w_{11}^{hh}; w_{00}^{hl}, w_{10}^{hl}, w_{01}^{hl}, w_{11}^{hl}\}$ can be alternatively characterized by $\mathcal{C} = \{f^{hh}, b_1^{hh}(y_1), b_2^{hh}(y_2), b_3^{hh}(y_1, y_2); f^{hl}, b_1^{hl}(y_1), b_2^{hl}(y_2), b_3^{hl}(y_1, y_2)\}$ where superscripts denote the matching quality over two periods. f is the fixed compensation regardless of the performance. $b_1(1) + f$ is paid if $y_1 = 1$ and $y_2 = 0$, $b_2(1) + f$ is paid if $y_1 = 0$ and $y_2 = 1$, $b_1(1) + b_2(1) + b_3(1, 1) + f$ is paid if $y_1 = y_2 = 1$, and f is paid if $y_1 = y_2 = 0$.

Proof At the second period, the agent of matching quality θ_h at date 1 and θ_h at date 2 maximizes effort e_2^{hh} given the contract and first period effort e_1^h.

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

$$\Longleftrightarrow$$

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

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}. I obtain the follwing:

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

Likewise, at the second period, the agent of matching quality θ_h at date 1 and θ_l at date 2 maximizes effort e_2^{hl} given the contract and first period effort e_1^h. 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 makes an effort as follows:

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

When the agent makes effort e_1^h in the first period, neither the principal nor the agent knows the private information in period 2, thus $b_1^{hl} = b_1^{hh}$. The agent with high matching quality thus maximizes the first-period compensation plus the second-period expected compensation.

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

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. Given the agent's effort, the principal maximizes her profit w.r.t. the eight parameters in $\mathcal{C}$.

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

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

Recall we have set the following variables previously $w_{00}^{hh} = f^h$, $w_{10}^{hh}-w_{00}^{hh} = b_1^{hh}$, $w_{01}^{hh}-w_{00}^{hh} = b_2^{hh}$, and $w_{11}^{hh}-w_{00}^{hh}-(w_{01}^{hh}-w_{00}^{hh})-(w_{10}^{hh}-w_{00}^{hh}) = b_3^{hh}$; $w_{00}^{hl} = f^{hl}$, $w_{10}^{hl}-w_{00}^{hl} = b_1^{hl}$, $w_{01}^{hl}-w_{00}^{hl} = b_2^{hl}$, and $w_{11}^{hl}-w_{00}^{hl}-(w_{01}^{hl}-w_{00}^{hl})-(w_{10}^{hl}-w_{00}^{hl}) = b_3^{hl}$ It is obvious to see that the above system can be transferred to the following one under the specification of $\{b\{.\}, f\{.\}\}$:

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

Q.E.D.

Lemma 3 If committing to a long term contract is impossible, the optimal contract can be characterized by two one-period contracts.

- The first one-period contracts for $m_1 = h$ and $m_1 = l$ at $t = 1$ are:
  For $m_1 = h$, $\{f_1 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_1(1) = \frac{1}{2}\}$.
  For $m_1 = l$, $\{f_1 = 0, b_1(1) = \frac{1}{2}\}$.
- The second one-period contracts for $m_2 = h$ and $m_2 = l$ at $t = 2$ are:
  For $m_2 = h$, $\{f_2 = \frac{1}{4}\theta_l(\theta_h - \theta_l), b_2(1) = \frac{1}{2}\}$.
  For $m_2 = l$, $\{f_2 = 0, b_2(1) = \frac{1}{2}\}$.

Proof The derivation is straightforward given the proof in Proposition 2. One only needs to substitute the production function $P(\theta, e_t) = \theta e$ and the disutility function $\psi(e) = \frac{1}{2}e^2$ into the principal and the agent's maximization programs and solve for the optimal contracts. Q.E.D.

Lemma 4 When cross-pledging is impossible, for an equilibrium contract $\mathcal{C} = \{b\{.\}, f\{.\}\}$, $b_3^{hh}(1,1) = b_3^{hl}(1,1) = b_3^{ll}(1,1) = 0$. To induce effort, the following components are greater than zero: $b_1^{hh} = b_1^{hl} > 0$, $b_2^{hh}(1) > 0$, $b_2^{hl}(1) > 0$, $b_1^{ll}(1) > 0$ and $b_2^{ll}(1) > 0$. In order to satisfy the limited liability, $f^{hl} \geq 0$ and $f^{hh} \geq 0$.

Proof 1. Assume the compensation paid to the agent of low matching quality 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 in the separating equilibrium. I first prove that if b_3^{ll} is set to zero for the agent of type θ_l at date 0, $b_1^{ll} > 0$, $b_2^{ll} > 0$. To simplify notation, $f^{hl} = f^l$ and $f^{hh} = f^h$.

Composite bonuses β^{hh} and β^{hl} defined in the proof Lemma 2 are important auxiliary variables. Here define $\beta^{ll} = (b_2^{ll} + \theta_l e_1^l b_3^{ll})$, so $b_2^{ll} = \beta^{ll} - \theta_l e_1^{ll} b_3^{ll}$. Following the proof in Lemma 2, it is easy to prove that $e_2^{ll} = \beta^{ll}\theta_l$ and $e_1^{ll} = \theta_l b_1^{ll} + \theta_l^3 \beta^{ll} b_3^{ll}$. The principal's maximization program is thus:

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

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 $b_1^{hh} = b_1^{hl} > 0$, $b_2^{hh} > 0$, and $f^l \geq 0$ and $f^h \geq 0$. Because in the first period, neither the agent nor the principal knows the second-period matching quality, thus $b_1^{hh} = b_1^{hl} = b_1^h$, and $e_1^{hh} = e_1^{hl} = e_1^h$. A principal who hires an agent of matching quality θ_h at date 0 maximizes the profit subject to two truth-telling constraints. The principal solves the following maximization problem P^h:

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

$$s.t. \quad \theta_l e_1^h(1 - b_1^h) + \theta_l e_2^{hl}(1 - b_2^{hl}) - \theta_l e_2^{hl} \theta_l e_1^h b_3^{hl} - f^l \leq \frac{1}{4}\theta_l^2 + \frac{1}{4}\theta_l^2 \tag{7.3}$$

$$\theta_l e_2^{hh}(1 - b_2^{hh}) - \theta_l e_2^{hh} \theta_h e_1^h b_3^{hh} - f^h \leq \theta_l e_2^{hl}(1 - b_2^{hl}) - \theta_l e_2^{hl} \theta_h e_1^h b_3^{hl} - f^l \tag{7.4}$$

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} (see Lemma 2), the above two constraints are equivalent to:

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

In order to satisfy the limited liability, $f^l \geq 0$ and $f^h \geq 0$. 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 Lemma 2, 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 $(\theta_h - \theta_l)(1-b_1^h)q\theta_h^3\beta^{hh}b_3^{hh}$. Without cross-pledging, it is zero. Thus $b_1^{hh} = b_1^{hl} = b_1^h > 0$ to induce first period effort. Similarly, one can find that $b_2^{hh} > 0$ to induce second period effort.

3. The last step is to prove that $b_2^{hl} > 0$. The principal deteriorating matching quality 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}:

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

$$s.t. \quad \theta_l e_2'^{hl}(b_2'^{hl} + \theta_h e_1^h b_3'^{hl}) + f'^l - \frac{1}{2}e_2'^{hl2} \geq \theta_l e_2^{hl}(b_2^{hl} + \theta_h e_1^h b_3^{hl}) + f^l - \frac{1}{2}e_2^{hl2}$$

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}:

$$\{\beta^{hl}, b_3^{hl}, f^l\} \in arg \max_{\{\beta'^{hl}, b_3'^{hl}, f'^l\}} \theta_l^2\beta'^{hl}(1-\beta'^{hl}) - f'^l$$

$$s.t. \quad \frac{1}{2}\theta_l^2\beta'^{hl2} - f'^l \geq u$$

b_3^{hl} thus does not enter the renegotiation-proof contract. To induce second period effort, b_2^{hl} must be greater than zero. Q.E.D.

Proposition 3.

- **Low Separating Profit** ($\theta_h < 2\theta_l$). The principal of matching quality θ_h commits to a contract at date 0: $\{b_1^{hh} = \frac{1}{2}$, $b_2^{hh} = \frac{1}{2}$, $f^{hh} = \frac{1}{4}\theta_l\theta_h - \frac{1}{4}\theta_l^2(2 - \frac{\theta_h}{\theta_l})$; $b_1^{hl} = \frac{1}{2}$, $b_2^{hl} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1})$, $f^{hl} = 0\}$. The principal will offer two one-period contracts to the agent of matching quality θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.
- **High Separating Profit** ($\theta_h \geq 2\theta_l$). The principal of matching quality θ_h commits to a contract at date 0: $\{b_1^{hh} = \frac{1}{2}$, $b_2^{hh} = \frac{1}{2}$, $f^{hh} = \frac{1}{2}\theta_l(\theta_h - \theta_l)$; $b_1^{hl} = \frac{1}{2}$, $b_2^{hl} = 1$, $f^{hl} = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)\}$. The principal will offer two one-period contracts to the agent of matching quality θ_l at date 0, $b_1^{ll} = b_2^{ll} = \frac{1}{2}$.

Proof To simplify notation, $f^{hl} = f^l$ and $f^{hh} = f^h$. The principal of high matching quality in the first period offers an optimal contract that is renegotiation-proof. A renegotiation-proof contract must satisfy the maximization program P^{hh} for a principal of matching quality hh listed in step 1 and program P^{hl} for a principal of matching quality hl listed in step 2:

1. Program P^{hh}

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

$$s.t.\quad \theta_l e_2^{\prime hh}(1-b_2^{\prime hh}) - \theta_l e_2^{\prime hh}\theta_h e_1^{\prime h} b_3^{\prime hh} - f^{\prime h} \le \theta_l e_2^{hl}(1-b_2^{hl}) - \theta_l e_2^{hl}\theta_h e_1^h b_3^{hl} - f^l \quad (7.5)$$

$$\theta_h e_2^{\prime hh}(b_2^{\prime hh} + \theta_h e_1^h b_3^{\prime hh}) + f^{\prime h} - \frac{1}{2}(e_2^{\prime hh})^2 \ge \theta_h e_2^{hh}(b_2^{hh} + \theta_h e_1^h b_3^{hh}) + f^h - \frac{1}{2}(e_2^{hh})^2 \quad (7.6)$$

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:

$$\{\beta^{hh}, b_3^{hh}, f^h\} \quad \in arg \max_{\{\beta^{\prime hh}, b_3^{\prime hh}, f^{\prime h}\}} \theta_h^2 \beta^{\prime hh}(1-\beta^{\prime hh}) - f^{\prime h}$$
$$s.t.\quad \frac{1}{2}\theta_h^2 \beta_h^{\prime 2} - f^{\prime h} \ge u$$

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

If the Constraint 7.5 is satisfied, the Constraint 7.6 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.

$$\{b_2^{hh}, f^h\} \quad \in arg \max_{\{b_2^{\prime hh}, f^{\prime h}\}} \theta_h e_2^{\prime hh}(1-b_2^{\prime hh}) - f^{\prime h}$$
$$s.t.\quad \theta_l e_1^{\prime h}(1-b_2^{\prime hh}) - f^{\prime h} \le \theta_l e_1^h(1-b_2^{hl}) - f^l$$

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

$$\theta_h^2 b_2^{\prime hh}(1-b_2^{\prime hh}) - f^{\prime h} > \theta_h^2 b_2^{hh}(1-b_2^{hh}) - f^h$$
$$\Leftrightarrow f^h - f^{\prime h} > \theta_h^2 b_2^{hh}(1-b_2^{hh}) - \theta_h^2 b_2^{\prime hh}(1-b_2^{\prime hh})$$
$$\Leftrightarrow f^h - f^{\prime h} > \theta_l\theta_h b_2^{hh}(1-b_2^{hh}) - \theta_l\theta_h b_2^{\prime hh}(1-b_2^{\prime hh})$$
$$\Leftrightarrow \theta_l\theta_h b_2^{\prime hh}(1-b_2^{\prime hh}) - f^{\prime h} > \theta_l\theta_h b_2^{hh}(1-b_2^{hh}) - f^h$$
$$\Leftrightarrow \theta_l\theta_h b_2^{\prime hh}(1-b_2^{\prime hh}) - f^{\prime h} > \pi^{hl}$$

The above inequality conflicts with the principal's truth-reporting constraint 7.5.

2. The second step is to show whether the principal of matching quality hl renegotiates the salary to zero depends on the separating profit. It is easy to see that b_1^h does not depend on the renegotiation as it is already sunk. Define $b_1^{hh} = b_1^{hl} = b_1^h$. Because b_1^h does not depend on the renegotiation as it is already sunk, the principal sets $b_1^h = \frac{1}{2}$. It is easy to verify that without cross-pledging, $b_1^l = \frac{1}{2}$. Following the proof in Lemma 4, Program $P^{\prime hl}$ can be simplified to the following program:

$$\{b_2^{hl}, f^l\} \quad \in arg \max_{b_2^{\prime hl}, f^{\prime l}} \theta_l^2 b_2^{\prime hl}(1-b_2^{\prime hl}) - f^{\prime l}$$
$$s.t.\quad \frac{1}{2}\theta_l^2 (b_2^{\prime hl})^2 - f^{\prime l} \ge u$$

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 $\theta_h \ge 2\theta_l$. From Constraint 7.3, one could verify that if $b_2^{hl} = 1$, $f^l = \frac{1}{4}\theta_l(\theta_h - 2\theta_l)$.

If $\theta_h < 2\theta_l$. 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 7.3, one could find that $b_2^{hl} = \frac{1}{2}(1 + \sqrt{\frac{\theta_h}{\theta_l} - 1})$.

Q.E.D.

Corollary 2 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:

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

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.

Lemma 5 If information is symmetric, the principal offers the following contracts in equilibrium:

- Principal with matching quality hh offers $\{b_1^{hh} = 0, b_2^{hh} = 0, b_3^{hh} = \frac{1}{\theta_h^2}, f^{hh} = 0\}$;
- Principal with matching quality hl offers $\{b_1^{hl} = 0, b_2^{hl} = 0, b_3^{hl} = \frac{1}{\theta_l\theta_h}, f^{hl} = 0\}$;
- Principal with matching quality ll offers: $\{b_1^{ll} = 0, b_2^{ll} = 0, b_3^{ll} = \frac{1}{\theta_l^2}, f^{ll} = 0\}$.

Proof I first prove an auxiliary result: a contract that 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} but only b_3^{hh}.

1. I first show that a contract which contains b_1^h offered by a principal of high matching quality 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 that contains b_2^{hh} offered by a principal of high matching quality 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.

I then prove Lemma 5.

1. Under contract $\{b\{.\}, f\{.\}\}$, the agent of constantly low matching quality could obtain utility level following Lemma 4:

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

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

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

   As a result, $b_3^{ll} = \frac{1}{\theta_l^2}$, $e_1^l = e_2^{ll} = 1$.

2. Following the same argument in the previous step, principal of high matching quality at date 0 only uses b_3^{hl} and b_3^{hh}. 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:

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

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:

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

Because $e_2^{hh}, e_2^{hl} \leq 1$, the principal 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 constant high matching quality 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 and sets $\beta^{hh} = \frac{1}{2}$, the agent's utility would be $\frac{1}{8}\theta_h^2$. The agent thus will not want to renegotiate. Q.E.D.

Proposition 4

- The principal of high matching quality commits to the following contract in the first period: $\{b_1^{hh} = 0, b_2^{hh} = 0, b_3^{hh} = \sqrt{\frac{1-(1-q)\theta_l^2}{q\theta_h^4}}, f^{hh} = f^{hl} + \theta_l \theta_h^2 b_3^{hh}(1 - \theta_h b_3^{hh}); b_1^{hl} = 0, b_2^{hl} = \frac{3q\theta_l^2 - 2q^2\theta_l^2 - q\theta_h + q^2\theta_h - \theta_l^2 + 1}{2q(1-q)\theta_l^2}, b_3^{hl} = \frac{1}{2\theta_l^2} - \frac{\theta_h^2 \beta^{hh} b_3^{hh}}{2(1-q)\theta_l^2}, f^{hl} = 1 - \theta_l + (\theta_h - \theta_l)\theta_l^2 b_3^{hl}\}$.
- The principal of low matching quality offers the following contract in the first period: $\{b_1^{ll} = 0, b_2^{ll} = 0, b_3^{ll} = \frac{1}{\theta_l^2}, f^{ll} = 0\}$.

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:

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

If $f^l > 0$, then $\beta^{hl} = 1$. The principal of constant high matching quality will not want to renegotiate the contract as proved in Proposition 3. 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 high matching quality 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.

Lemma 6 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:

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

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.