Incentive Contracts with Signaling

Daisuke Hirata*
Hitotsubashi University

This Version: March 12, 2017

Abstract

This paper studies a simple model of incentive-contracting where (i) a principal learns an agent's ability before the agent himself, and (ii) both the agent's productivity with the principal as well as his outside option depends on his ability. I characterize the optimal contracts for the principal, defined to be the most profitable equilibrium outcomes among those satisfying the D1 criterion: Pooling at an earlier date is strictly optimal if the agent's outside option is sufficiently sensitive to the principal's private information, whereas separation at a later date is (weakly) optimal otherwise. Further, the principal's profit is shown to be neither continuous nor monotone with respect to the agent's outside option. Implications for unraveling in entry-level labor markets are also discussed.

*d.hirata@r.hit-u.ac.jp
This paper is based on a chapter of my thesis submitted to Harvard University. I am grateful to Oliver Hart for his guidance and encouragement. I also thank seminar participants at Harvard University, the University of Tokyo, and the 9th Japan-Taiwan-Hong Kong Contract Theory Conference for helpful comments and discussions. The usual disclaimer applies.

1 Introduction

This paper studies the incentive-contracting problem when a principal is better informed of an agent's ability than the agent himself.[1] While economists often assume the contrary, such a situation would naturally arise particularly in entry-level job markets. On the one hand, employers would know from their past experiences what attributes are important in determining one's prospects in their jobs, and their recruitment process (e.g., interview questions) must be designed to best infer an applicant's productivity by inspecting those important characteristics. On the other hand, although new workers might know well what attributes they do or do not have, they would be much more uncertain about how those attributes are converted into their productivity and/or how they are evaluated by potential employers. Then, it would be possible that at the contracting stage (after some screening process) an employer has a better estimate of a worker's ability or productivity than the worker himself.

Furthermore, if job-hunting is a sequential process, such informational asymmetry is not necessarily resolved by the time when a worker decides to accept or reject an offer. That is, even though he would be better informed of his (expected) ability after he observes the offers made by multiple employers, he may not be allowed to postpone his decision until he collects sufficient information. If so, he can optimally accept an offer and quit the search and learning process while he is still uncertain about his prospects. Actually, in many entry-level job markets such as the ones for new MBAs, law graduates, and clinical psychologists, employers often make an *exploding offer*, which expires unless a worker accepts within a very short time window (Roth and Xing, 1994, 1997), so as to prevent the worker from comparing multiple offers.

If an agent is uncertain about his own ability, his subjective belief could have two

[1] I follow the convention of referring to a principal/employer as she and to an agent/employee as he.

opposite effects. On the one hand, the higher his belief, the more effort he will exert conditional on the acceptance of an offer, because he believes his effort is likely to yield a better outcome and thus to lead a higher bonus payment. On the other hand, however, the higher his belief, the more rent he will require to accept a contract offered by the principal, since he believes he can get a better outside option by continuing his job search. Accordingly, the principal could also have two opposite incentives to manipulate the agent's belief through her offer. That is, she might want to raise his belief in order to induce more effort, by offering a contract which is "appropriate" for a high-ability agent (e.g., high-powered incentives). Or she might want to lower it so as to let him accept an unfavorable offer, by offering a contract for a low-ability agent.

The purpose of this paper is to investigate how those forces are balanced at an equilibrium and shape the optimal contracts and timing for the principal. Specifically, I study a simple model of moral hazard with two dates, where a principal learns the agent's ability at date 1 but the agent can do so only at date 2. I analyze equilibria satisfying the D1 criterion (Cho and Kreps, 1986; Cho and Sobel, 1990), which is a standard equilibrium refinement in the signaling-game literature, and characterize the most profitable outcomes to the principal among those equilibria. As a result, it turns out that the principal optimally makes a same exploding offer to any type of agent at date 1, if the agent's outside option is sufficiently sensitive to the agent's ability; otherwise, it is weakly optimal to delay an offer until date 2. Interestingly, the principal's highest profit is not monotonically decreasing in the agent's outside option. If the outside option for a high-ability agent increases, it decreases the incentive for the principal to wait until date 2 because she will need to give a larger rent to him after the information reveals. Hence, pooling at date 1 becomes easier to sustain as an equilibrium, which is profitable for the principal because she can induce a higher level of effort from a low-ability agent. Conversely, if the incentive compatibility binds with the high-ability agent, a marginal decrease in his outside option forces the principal to always wait until date 2, even though pooling is strictly profitable when she face the low-type agent. Consequently, her profit may discontinuously decrease by such a marginal change in the agent's reservation wage.

The rest of this paper is organized as follows. Section 1.1 briefly discusses the related literature. Section 2 sets up the model and Section 3 presents the analysis. Section 4 makes a few concluding remarks. Appendix A contains omitted proofs.

1.1 Related Literature

The most related to this paper are the studies in the contract theory literature that investigate the effect of a principal's private information about the profitability or production technology of her business (see, e.g., Bénabou and Tirole, 2003; Inderst, 2001; Silvers, 2012; Spier, 1992). This is because both the profitability in those papers and the agent's ability in this paper are modeled as a parameter of a production function. However, a key distinction is that the agent's ability in this paper is also correlated with his outside option while the principal's information is purely firm-specific in those papers.[2] The correlation between the principal's knowledge and the agent's outside options creates additional effects of the agent's belief, and thereby makes it optimal for the principal to make early, exploding offers. Relatedly, De la Rosa (2011) and Santos-Pinto (2008) study how the agent's over-confidence or positive self-image affects the optimal incentive scheme, but they also assume that the agent's reservation utility is independent of the agent's belief.

More broadly, this paper is related to the literature on the mechanism design problem with an informed principal (e.g., Myerson, 1983; Maskin and Tirole, 1992). Compared to this literature, it should be noted that the analysis in this paper (as well as the studies mentioned above) implicitly assumes that the principal cannot force the agent to commit

[2]Put differently, the model in this paper could be alternatively viewed as the situation where a principal privately knows about non-firm-specific profitability (e.g., the forecasts of business conditions) that affects both the production function inside the firm and the outside option for the agent.

to work for her, without fully specifying the contract. In contrast, the mechanism design literature typically assume that the principal can delay to reveal her private information, as long as a mechanism satisfies participation constraints in expectation. In the current context, this means that the principal can offer a contract saying, e.g., "if you sign this contract you must work for me, and I will choose the incentive scheme from this pre-specified set after you sign." In the environment of this paper, allowing such contracts may or may not be beneficial to the principal depending on parameter values, but it generally increases the incentive for early contracting.

As this paper implicitly assumes that the principal learns the agent's type before other employers make competitive offers, it also loosely relates to the personnel economics literature that examine the asymmetric information about a worker's ability between the current and other potential employers (e.g., Waldman, 1984; Gibbons and Katz, 1991). However, it should be noted that these papers study the competition for mid-career workers, while this paper is intended to model entry-level labor markets. Note also that workers' beliefs are irrelevant in those papers, for they are assumed to simply take the offer with the highest (fixed) wage.

2 Model

Suppose that a principal hires an agent to run a project. If the agent decides to work for the principal, the project will either succeed ($Y = 1$) or fail ($Y = 0$). The probability of success is given by

$$\text{Prob}[Y = 1] = \min \{e \cdot \theta + 2F, 1\},$$

where $e \geq 0$ is the agent's effort and $\theta \in \{H, L\} \subset \mathbb{R}_+$ is his ability. Without any loss, we assume $H > L$.[3] The prior probability of $\theta = H$ (resp. $\theta = L$) is denoted by p (resp. $1 - p$) and is common knowledge between the principal and agent. The prior mean of θ is denoted by $M = pH + (1 - p)L$. To induce effort, the principal can offer a contingent bonus $b \in [0, 1]$, which is paid if and only if $Y = 1$.[4] Taking the incentive bonus b as fixed, the principal's profit and the agent's utility are $\Pi(b, Y) = (1 - b)Y$ and $U(b, Y, e) = bY - \frac{c}{2}e^2$, where c is a cost parameter. To guarantee an internal solution to the effort choice problem, we assume c is sufficiently high:

Assumption 1. The cost parameter is sufficiently high: $c > H^2/(1 - 2F)$.

If the agent decides not to work for the principal, he will return to an outside labor market and find another employer. The principal's profit in this case is assumed to be zero.[5] The value of the outside option for the agent, contingent on θ, is denoted by u_θ, with $u_H \geq u_L$. The prior expectation of u_θ is $u_M = pu_H + (1 - p)u_L$. To simplify the analysis and focus on the main insights, we also assume the following:

Assumption 2. The expected value of the agent's outside option is sufficiently low: $u_M \leq 0$.[6]

As motivated in the introduction, the principal could better estimate θ during the hiring process, whereas the agent would be uncertain what outside options he could expect until he indeed goes through the process with other employers. To model such a situation,

[3] However, L should not be interpreted as the lowest productivity among a pool of qualified agents rather than among all potential workers. See Section 4 for details.

[4] An implicit assumption here is that the agent is wealth-constrained.

[5] This does not necessarily mean that the principal cannot find another employee. As mentioned in footnote 3, the agent in this model should be seen as sufficiently qualified even when $\theta = L$, and it could be very costly to find an equally qualified worker even if it is possible.

[6] All of the results will remain qualitatively the same under a much weaker assumption that $\beta_L^{PC} < \beta_H^*$ (see equations (1)–(2) for the definitions of these variables). However, minor changes will be required in both the statements and proofs, because we need to consider more subcases when the participation constraint is biding for the L-type.

suppose that there are two dates, 1 and 2, and that the principal and agent can directly learn θ at the beginning of date 1 and date 2, respectively. Further, the principal can make a contract offer either at date 1 or date 2, and in the case of an early offer, she can also force the agent to take it or leave it by the end of date 1 (i.e., the offer is exploding). Although the agent cannot directly observe θ at date 1, he rationally updates his belief if an early offer is made. Without any loss, I identify his subjective belief with the expectation of θ according to that belief. In what follows, let $\vartheta(b) \in [L, H]$ denote the agent's belief after observing an exploding offer of $b \in [0, 1]$, and $u_\vartheta := \frac{\vartheta - L}{H-L} u_H + \frac{H-\vartheta}{H-L} u_L$ be the subjective expectation of the outside option when the belief is ϑ.[7] To summarize, the timeline of the game is as follows:

0. The nature draws θ from the prior distribution.

1. After the principal observes θ through the screening process, she decides whether to make an exploding offer $b \in [0, 1]$ to the agent, or wait until date 2 (denoted by $b = \varnothing$). If she makes an offer, the agent decides to accept or reject it after updating his belief.

   (a) If the agent accepts the offer, he choose his effort level, the output Y realizes, and the payoffs are finalized.

   (b) If the agent rejects, the agent gets the outside option u_θ and the principal's profit is zero.

2. If the principal did not make an exploding offer in date 1, she offers $b \in [0, 1]$ after θ becomes publicly observable.

   (a) If the agent accepts the offer, he choose his effort level, the output Y realizes, and the payoffs are finalized.

[7] It is unnecessary to specify the belief after no offer at date 1, because he will know θ for sure at the beginning of date 2.

(b) If the agent rejects, the agent gets the outside option u_θ and the principal's profit is zero.

3 Analysis

Preliminaries

To analyze the model backwardly, first consider the agent's effort choice problem. If the agent accepts a contract b with belief $\vartheta \in [L, H]$, he will solve

$$\max_e \left[b\vartheta e - \frac{c}{2}e^2 \right],$$

which yields $e^*(b, \vartheta) = \frac{b\vartheta}{c}$ under Assumption 1. Hence, the principal's expected profit will be

$$\widetilde{\Pi}(b, \vartheta, \theta) := (2\Delta - b) \left[\frac{b\vartheta\theta}{c} + 2F \right].$$

Given belief ϑ, the agent optimally accepts an offer b if and only if

$$b \in \mathrm{Ac}(\vartheta) := \left\{ b \in [0, 1] : \frac{1}{2} \left(\frac{b\vartheta}{c} \right)^2 \geq u_\vartheta \right\}.$$

When θ becomes publicly observable at date 2, the principal's problem is to maximize $\widetilde{\Pi}(b, \theta, \theta)$ subject to $b \in \mathrm{Ac}(\theta)$. The optimal offer is $b = \beta_\theta^F := \max\left\{\beta_\theta^*, \beta_\theta^{PC}\right\}$, where

$$\beta_\theta^* := \arg\max_{b \in [0,1]} \widetilde{\Pi}(b, \theta, \theta) = \frac{1}{2} - \frac{cF}{\theta^2}, \text{ and} \tag{1}$$

$$\beta_\theta^{PC} := \min\left(\mathrm{Ac}(\theta) \cup \{1\}\right), \tag{2}$$

and the principal's unique equilibrium profit with full information is $\widetilde{\Pi}(\beta_\theta^F, \theta, \theta)$.[8]

In what follows, we characterize the most profitable (pure strategy) perfect Bayesian equilibria to the principal among those satisfying the D1 criterion (Cho and Kreps, 1986; Cho and Sobel, 1990). Since the principal can always wait until date 2, each θ-type principal must also earn non-negative expected profits in any equilibrium. Therefore, the principal can never profitably deviate from an equilibrium by offering b that will be rejected. Taking this observation into account, we can define the D1 criterion in this model as follows.

Definition 1. Fix an equilibrium and let Π_θ^* denote the θ-type principal's equilibrium profit, for each $\theta \in \{H, L\}$. A pair (b, θ) is said to be *deleted by the D1 criterion* if (i) for all $\vartheta \in [L, H]$, $\widetilde{\Pi}(b, \vartheta, \theta) \geq \Pi_\theta^*$ implies $\widetilde{\Pi}(b, \vartheta, \theta') > \Pi_{\theta'}^*$, and (ii) there exists $\vartheta \in [L, H]$ such that $b \in \mathrm{Ac}(\vartheta)$ and $\widetilde{\Pi}(b, \vartheta, \theta') > \Pi_{\theta'}^*$, where $\theta' \in \{H, L\} - \{\theta\}$.[9] A belief system $\vartheta^*(\cdot) : [0, 1] \to [L, H]$ is said to *satisfy the D1 criterion* if $\vartheta^*(b) = \theta'$ whenever (b, θ) is deleted. The equilibrium is said to *satisfy the D1 criterion* if its associated belief system satisfies the D1 criterion. □

The following Lemma is useful in the subsequent analysis to restrict the belief systems that satisfy the D1 criterion. Although this corresponds to the standard sorting condition in the signaling game literature, it is effective only when the contracts are accepted by the agents. As a consequence, it does not always eliminate the possibility of pooling equilibria.[10]

[8]Notice that β_θ^{PC} is defined to be 1 when $\mathrm{Ac}(\theta)$ is empty, and $\widetilde{\Pi}(1, \theta, \theta) = 0$. Hence, the principal's full-information profit is still $\widetilde{\Pi}(\beta_\theta^F, \theta, \theta)$ even in the case where no contract is signed at an equilibrium.

[9]Notice that this definition is based only on the agent's pure best replies while the standard definition uses mixed best replies. However, this difference is irrelevant here because the pure best reply is generically unique.

[10]Technically, a key property of the present model is that the principal's profit is not continuous with respect to the agent's belief, because a marginal increase in ϑ can make the agent to switch from accepting to rejecting an offer and thereby discontinuously lower the profit. This is why the results in Cho and Sobel (1990) do not apply here.

Lemma 1. *Suppose that $b' > b$. If $\widetilde{\Pi}(b', \vartheta', L) \geq \widetilde{\Pi}(b', \vartheta, L)$, then $\widetilde{\Pi}(b', \vartheta', H) > \widetilde{\Pi}(b', \vartheta, H)$.*

Proof. See Appendix A. ■

Optimal Equilibria

Now we are ready to characterize the most profitable equilibrium outcomes among those satisfying the D1 criterion. The first result establishes that the full-information outcome is always supportable by an equilibrium.

Proposition 1. *There always exists an equilibrium satisfying the D1 criterion such that the principal never offers an exploding offer at date 1.*

Proof. See Appendix A. ■

Notice that the equilibrium outcome in Proposition 1 is always (weakly) more profitable than any separating equilibria, because the agent must correctly and certainly know θ on the equilibrium-path of such equilibria. The question is, therefore, whether and when pooling equilibria exist. The next proposition fully characterizes the condition for pooling at date 1 to be supportable. This condition immediately implies that whenever a pooling equilibrium exists, it yields a weakly higher profit than any separating equilibria.[11]

Proposition 2. *There exists a pooling equilibrium satisfying the D1 criterion, where the principal offers a same $b \in [0, 1]$ to both types of the agent at date 1, if and only if $\widetilde{\Pi}(b, M, \theta) \geq \widetilde{\Pi}(b^F_\theta, \theta, \theta)$ for each $\theta \in \{L, H\}$.*

Proof. See Appendix A. ■

[11] See Section 4 for a brief discussion regarding semi-pooling equilibria.

Corollary 3. *Suppose that some pooling equilibrium at date 1 satisfies the D1 criterion. Then, there exists a pooling equilibrium at date 1 that yields a (weakly) higher profit for the principal than any other pure strategy equilibrium surviving the D1 criterion.*

Proof. The is an immediate corollary of Proposition 2. ■

Comparative Statics

Since Propositions 1 and 2 pin down the principal's highest equilibrium profit for a fixed set of parameters, the next question should be how it varies with changes in the parameters. Actually, the principal's profit has a few interesting comparative statics properties with respect to changes in the agent's outside option(s). So as to keep Assumption 2 intact against such variations, now take u_M as fixed and let $D := u_H - u_L \geq 0$ be a free parameter. That is, higher D means both higher u_H and lower u_L. Note that D would increase when the principal's assessment of θ becomes more accurate or more correlated with the assessments by other potential employers.

As D increases from 0, the participation constraint for the H-type becomes more and more binding where as that for the L-type remains slack. Therefore, the principal's profit from separating at date 2 is (weakly) decreasing in D. In contrast, the constraint $\widetilde{\Pi}(b, M, H) \geq \widetilde{\Pi}(\beta_L^F, H, H)$ becomes less demanding when D rises, for the right-hand side decreases in u_H. Consequently, the set of supportable pooling equilibria at date 1 is (weakly) increasing in D, and so is the profit at the optimal equilibrium as long as it is non-empty.

To further look into the behavior of optimal pooling equilibria, let

$$\beta_M^1 := \arg\max_b \widetilde{\Pi}(b, M, H)$$

and β_M^2 be such that $\beta_M^2 > \beta_L^F$ and $\widetilde{\Pi}(\beta_M^2, M, L) = \widetilde{\Pi}(\beta_L^F, L, L)$. That is, β_M^1 is the point that minimizes the principal's incentive to reveal $\theta = H$, and β_M^2 is the maximal point at which the principal has no incentive to reveal $\theta = L$. Also define $\beta_M^* := \arg\max_b \widetilde{\Pi}(b, M, M)$ to denote the best possible pooling outcome. Observe that $\beta_M^* < \beta_H^1$ must hold by definition. Based on these variables, then, we need to consider three cases.

- First, suppose that $\beta_M^* < \beta_M^1 < \beta_M^2$: The set of supportable pooling contract becomes non-empty when D hits D^*, which is defined to be the point such that $\widetilde{\Pi}(\beta_M^1, M, H) \geq \widetilde{\Pi}(\beta_H^F, H, H)$ holds with equality. At this point, the unique supportable pooling equilibrium yields the expected profits of $\widetilde{\Pi}(\beta_M^1, M, M)$. Furthermore, the assumption of $\beta_M^1 < \beta_M^2$ implies $\widetilde{\Pi}(\beta_M^1, M, L) > \widetilde{\Pi}(\beta_L^F, L, L)$ and hence,

$$\begin{aligned}\widetilde{\Pi}(\beta_M^1, M, M) &= p \cdot \widetilde{\Pi}(\beta_M^1, M, H) + (1-p)\, \widetilde{\Pi}(\beta_M^1, M, L) \\ &> \; p \cdot \widetilde{\Pi}(\beta_H^F, H, H) + (1-p)\, \widetilde{\Pi}(\beta_L^F, L, L),\end{aligned}$$

  i.e., the principal's profit at this unique pooling equilibrium is *strictly* higher than her full-information profits. Since no pooling equilibrium is supportable at any $D < D^*$, the profits at the optimal equilibrium jumps up at $D = D^*$. When D further increases to $D^* + \delta$, pooling becomes supportable on a larger interval of b, containing β_M^1 in its interior. As $\beta_M^1 > \beta_M^*$ by definitions, we can conclude that the principal's profit is strictly right-increasing at $D = D^*$.

- Second, suppose that $\beta_M^* < \beta_M^2 \leq \beta_M^1$: Then, the first supportable pooling is $b_H = b_L = \beta_M^2$, which becomes supportable at $D = D^*$, which is now defined by $\widetilde{\Pi}(\beta_M^2, M, H) = \widetilde{\Pi}(\beta_H^F, H, H)$. When D further increases, pooling at $\beta_M^2 - \varepsilon$ becomes supportable, and it is strictly more profitable because $\widetilde{\Pi}(\cdot, M, M)$ is decreasing at β_M^2 by the assumption of $\beta_M^* < \beta_M^2$. That is, the principal's profit is strictly right-increasing at $D = D^*$. In this case, however, it is continuous because $\widetilde{\Pi}(\beta_M^2, M, L) = \widetilde{\Pi}(\beta_L^F, L, L)$ holds by definition, as well as $\widetilde{\Pi}(\beta_M^2, M, H) = \widetilde{\Pi}(\beta_H^F, H, H)$ at $D = D^*$.

- Lastly, suppose that $\beta_M^2 \leq \beta_M^*$: As in the second case, the first supportable pooling is at β_M^2 and the principal's profit is continuous at $D = D^*$. Yet, the profit is not right-increasing in this case, because $\widetilde{\Pi}(\cdot, M, M)$ is increasing on $[0, \beta_M^2]$, and pooling at $b > \beta_M^2$ becomes never supportable.

For each of these three cases, the trajectory of the principal's profit at the optimal equilibrium is illustrated in Figure 1. To summarize, we have established the following proposition.

Proposition 4. *Fix u_M as given and take D to be the free parameter. There exists a cutoff D^* such that in the optimal equilibrium, the principal offers a same exploding offer to both types of agents if and only if $D \geq D^*$. With respect to D, the principal's profit at the optimal equilibrium is weakly (resp. strictly) decreasing on $[0, D^*)$ (resp. in a left neighborhood of D^*), and weakly increasing on $[D^*, \infty]$. Further, it is strictly right-increasing at D^* if $\beta_M^2 > \beta_M^*$, and discontinuously jumps up at D^* if $\beta_M^2 > \beta_M^1$.*

Proof. See the above arguments. ■

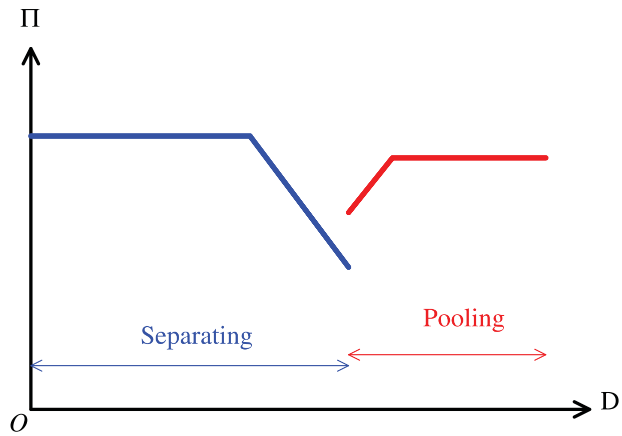
Since the decrease in u_L is relevant only by keeping u_M constant, the principal's profit has similar comparative statics properties with respect u_H as long as u_M is sufficiently low. In particular, we have the following corollary.

Corollary 5. *Given u_L is sufficiently low, the principal's profit at the optimal equilibrium may be discontinuous and increasing in u_H.*

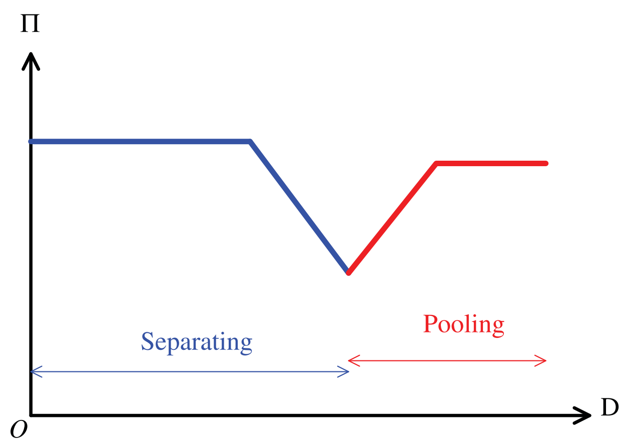
Proof. The is an immediate corollary of Proposition 4. ■

4 Discussions

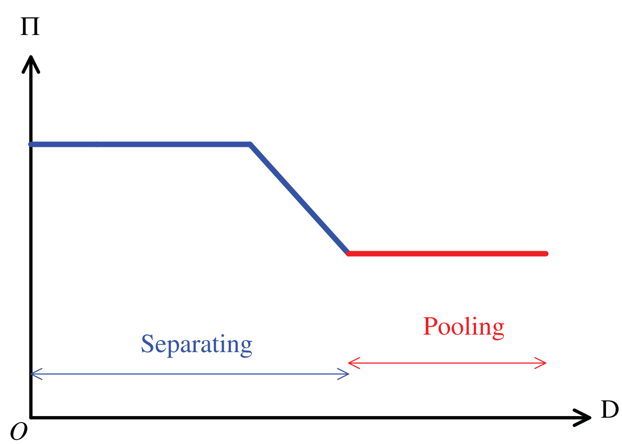
This paper studies an incentive-contracting problem with the assumption that the principal can learn about the agent's productivity before the agent himself. I characterize the



(a) $\beta^*_M < \beta^1_M < \beta^2_M$.



(b) $\beta^*_M < \beta^2_M \le \beta^1_M$.



(c) $\beta^2_M \le \beta^*_M$.

Figure 1: Principal's highest profit as a function of D.

most profitable equilibria to the principal and illustrate how such (adversely) asymmetric information can cause the principal to make early, exploding offers. It should be noted that in the present model, the time length between date 1 and 2 could be very short, as is often observed in reality, if the hiring schedule in the outside market is sufficiently tight. This could be seen as a possible advantage compared with the existing models of unraveling (e.g., Halaburda, 2010; Li and Rosen, 1998; Li and Suen, 2000), where the time window is defined by the evolution of public information on match qualities. One might wonder, however, if the principal's information is really essential in the above analysis, because when the optimal equilibrium is pooling, she would be (weakly) better off by offering β^*_M before learning θ herself. This argument is completely valid if the above model is taken literally, but not necessarily in general. Suppose, as briefly mentioned in footnotes 3 and 5, that there exists a third type of agents and the principal prefers vacancy to hiring such agents. If the cost of hiring those unqualified agents is sufficiently high, the principal would strictly prefer to make offers after she learns θ but before the agent does.

In terms of comparative statics, this study suggests, unlike the standard moral hazard models, an increase in the (H-type) agent's reservation wage may benefit the principal. An interesting implication of this result is that wage competitions at later dates may actually enhance unraveling. If an employer wants to hire more qualified agents but such agents are scarce at date 2, she might consider to raise the wage she offers, hoping it would let more agents to remain until date 2. However, our model would suggest that such a wage increase would actually lead more competitors to make an exploding offer, and consequently, the qualified agents could become even more scarce at date 2. This argument is of course informal, because the present model abstracts away from such strategic interactions among employers. It would be an interesting avenue for future research to extend the model and study market equilibria, endogenizing the agent's outside options.

To conclude, let us briefly discuss the possibility of semi-pooling equilibria at date

1. Indeed, semi-pooling can be an equilibrium and moreover, strictly optimal for the principal. To see this, suppose that $\beta_M^* < \beta_M^2 < \beta_M^1$ and $D = D^*$. Then, pooling is profitable to the principal but $\widetilde{\Pi}(\beta_M^2, M, H) = \widetilde{\Pi}(\beta_H^F, H, H)$ is a binding constraint. If the L-type principal randomize between $b_L = \beta_L^*$ and $b^L = \beta_M^2 + \varepsilon$, it is both strictly profitable and incentive compatible for the H-type principal to offer $b_H = \beta_M^2 + \varepsilon$ with probability one, because $\widetilde{\Pi}(\beta_M^2 + \varepsilon, \vartheta, H) > \widetilde{\Pi}(\beta_H^F, H, H)$, where $\vartheta > M$ is the conditional belief at $b = \beta_M^2 + \varepsilon$. Further, randomization is also incentive compatible for the L-type principal if ε is sufficiently small and ϑ is sufficiently close to M. Thus, such a mixed strategy forms an equilibrium and is more profitable than full pooling at β_M^2. However, semi-pooling equilibria may not necessarily fill the "gap" in the principal's profit, because randomization by the L-type principal cannot increase her own profit, whereas it is the jump in the L-type's profit that gives rise to the discontinuity in Proposition 4.[12]

A Proofs

Proof of Lemma 1. By definition,

$$\widetilde{\Pi}(b', \vartheta', L) - \widetilde{\Pi}(b', \vartheta, L) = 2F(b - b') + \left[(2\Delta - b')\frac{b'\vartheta'}{c} - (2\Delta - b)\frac{b\vartheta}{c}\right] L.$$

Since $b' > b$, the first term on the RHS is negative and hence, the second term must be strictly positive if the LHS is non-negative. Then, the statement immediately follows. ■

Proof of Proposition 1. To begin, define $\beta_L^{IC} > \beta_L^F$ by $\widetilde{\Pi}(\beta_L^{IC}, H, L) = \widetilde{\Pi}(\beta_L^F, L, L)$, i.e., β_L^{IC} is the highest possible offer that a L-type principal can have an incentive to offer at date 1. Note that by definition, the D1 criterion never deletes (b, H) for $b \geq \beta_L^{IC}$.

[12] It is easy to check randomization by the H-type principal is never profitable, because it will decrease the conditional expectation at the point where both types pool.

First, suppose that $\beta_L^{IC} \geq \beta_H^F$. Define a belief system $\vartheta^*(\cdot)$ by $\vartheta^*(b) = L$ if $b < \beta_L^{IC}$ and $\vartheta^*(b) = H$ if $b \geq \beta_L^{IC}$. Associated with this $\vartheta^*(\cdot)$, it is apparent that $(b_H, b_L) = (\varnothing, \varnothing)$ forms a PBE. For any $b < \beta_L^{IC}$, $\widetilde{\Pi}(b, \vartheta, H) \geq \widetilde{\Pi}(\beta_H^F, H, H)$ implies $\widetilde{\Pi}(b, \vartheta, H) \geq \widetilde{\Pi}(\beta_L^{IC}, H, H)$ and hence by Lemma 1, $\widetilde{\Pi}(b, \vartheta, L) > \widetilde{\Pi}(\beta_L^{IC}, H, L)$. Therefore, (b, L) is not deleted for $b < \beta_L^{IC}$ and $\vartheta^*(\cdot)$ survives the D1 criterion.

Second, suppose that $\beta_L^{IC} < \beta_H^F = \beta_H^*$, and let $\vartheta^*(\cdot)$ be the same belief system as in the previous case. Again, it is immediate to check $(b_H, b_L) = (\varnothing, \varnothing)$ with this $\vartheta^*(\cdot)$ is an equilibrium. In this case, $b < \beta_L^{IC}$ directly implies $\widetilde{\Pi}(b, \vartheta, H) \leq \widetilde{\Pi}(b, H, H) < \widetilde{\Pi}(\beta^*, H, H)$ and hence, (b, L) is not deleted for any $b < \beta_L^{IC}$.

Finally, suppose that $\beta_L^{IC} < \beta_H^F = \beta_H^{PC}$. Let $\vartheta^*(\cdot)$ be a belief system satisfying (i) $\vartheta^*(b) = H$ for all $b \geq \beta_L^{IC}$ and all $b \in \mathcal{B}_H^{IC}$, and (ii) $\vartheta^*(b) \in \{H, L\}$ for all b, where

$$\mathcal{B}_H^{IC} := \left\{ b \in [0,1] : \widetilde{\Pi}(b, L, H) \geq \widetilde{\Pi}(\beta_H^F, H, H) \right\}.$$

Notice that $\sup \mathcal{B}_H^{IC} < \beta_H^F$ always holds (as long as $\mathcal{B}_H^{IC}$ is non-empty). Then, we can check any such belief system can support $(b_H, b_L) = (\varnothing, \varnothing)$ as an equilibrium. Since either (b, H) or (b, L) must survive the D1 criterion for each $b \in [0, 1]$, we can always pick $\vartheta^*(\cdot)$ that satisfies the D1 criterion. ■

Proof of Proposition 2. The "only if" part is immediate because the principal can always secure the profits of $\widetilde{\Pi}(\beta_\theta^F, \theta, \theta)$ by delay her offer. To show the "if" part, suppose that an arbitrary $b \in [0, 1]$ satisfies the condition, and let $\vartheta^*(\cdot)$ be the belief system such that $\vartheta(b') = L$ if $b' < b$, $\vartheta(b') = M$ if $b' = b$, and to $\vartheta(b') = H$ if $b' > b$. Then, we can easily check that it is optimal for each θ-type principal to make an exploding offer $b_\theta = b$, given the agent's belief system $\vartheta^*(\cdot)$: First, if $\theta = L$ and the principal offers $b_L = b' < b$, her profit will be $\widetilde{\Pi}(b', L, L) \leq \widetilde{\Pi}(\beta_L^F, L, L) \leq \widetilde{\Pi}(b, M, L)$, where the first and second inequalities hold by the definition of $\beta_L^F \equiv \beta_L^*$ and by assumption, respectively. Hence the L-type principal has no incentive to deviate by $b_L = b' < b$. Similarly, it is never profitable for the H-type principal to offer $b_H = b' > b$, because by definitions, either $b' \notin \mathrm{Ac}(H)$ or $\widetilde{\Pi}(b', H, H) \leq \widetilde{\Pi}(\beta^F_H, H, H) \leq \widetilde{\Pi}(b, M, H)$. Next, suppose that the principal offers $b_L = b' > b$. If this is strictly profitable, i.e., if $b' \in \mathrm{Ac}(H)$ and $\widetilde{\Pi}(b', H, L) > \widetilde{\Pi}(b, M, L)$, then Lemma 1 implies $\widetilde{\Pi}(b', H, H) > \widetilde{\Pi}(b, M, H)$, but this is a contradiction to the previous argument. The last case, $\theta = H$ and $b_H = b' < b$, is also guaranteed to be never profitable by Lemma 1. In sum, $(b_H, b_L) = (b, b)$ is an equilibrium outcome with the belief system $\vartheta^*(\cdot)$. Since Lemma 1 directly implies $\vartheta^*(\cdot)$ satisfy the D1 criterion, the proof is complete. ■

References

BÉNABOU, R. AND J. TIROLE (2003): "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70, 489–520.

CHO, I.-K. AND D. M. KREPS (1986): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–222.

CHO, I.-K. AND J. SOBEL (1990): "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50, 381–413.

DE LA ROSA, L. E. (2011): "Overconfidence and Moral Hazard," *Games and Economic Behavior*, 73, 429–451.

GIBBONS, R. AND L. F. KATZ (1991): "Layoffs and Lemons," *Journal of Labor Economics*, 9, 351–380.

HALABURDA, H. (2010): "Unravelling in two-sided matching markets and similarity of preferences," *Games and Economic Behavior*, 69, 365–393.

Inderst, R. (2001): "Incentive Schemes as a Signaling Device," *Journal of Economic Behavior & Organization*, 44, 455–465.

Li, H. and S. Rosen (1998): "Unraveling in Matching Markets," *American Economic Review*, 88, 371–387.

Li, H. and W. Suen (2000): "Risk Sharing, Sorting, and Early Contracting," *Journal of Political Economy*, 108, 1058–1091.

Maskin, E. and J. Tirole (1992): "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60, 1–42.

Myerson, R. B. (1983): "Mechanism Design by an Informed Principal," *Econometrica*, 51, 1767–1797.

Roth, A. E. and X. Xing (1994): "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84, 992–1044.

——— (1997): "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," *Journal of Political Economy*, 105, 284–329.

Santos-Pinto, L. (2008): "Positive Self-image and Incentives in Organisations," *Economic Journal*, 118, 1315–1332.

Silvers, R. (2012): "The Value of Information in a Principal-Agent Model with Moral Hazard: The Ex Post Contracting Case," *Games and Economic Behavior*, 74, 352–365.

Spier, K. E. (1992): "Incomplete Contracts and Signaling," *RAND Journal of Economics*, 23, 432–443.

WALDMAN, M. (1984): "Job Assignments, Signalling, and Efficiency," *RAND Journal of Economics*, 15, 255–267.