# Toward a General Theory of Commitment, Renegotiation and Contract Incompleteness (2) : Commitment Problem and Optimal Incentive Schemes in Agency with Bilateral Moral Hazard

鈴木, 豊 / SUZUKI, Yutaka

(出版者 / Publisher)
法政大学経済学部学会
(雑誌名 / Journal or Publication Title)
経済志林 / The Hosei University Economic Review
(巻 / Volume)
66
(号 / Number)
1
(開始ページ / Start Page)
339
(終了ページ / End Page)
358
(発行年 / Year)
1998-07-30
(URL)
https://doi.org/10.15002/00002599

# (II) Commitment Problem and Optimal Incentive Schemes in Agency with Bilateral Moral Hazard*

*by* Yutaka Suzuki

**Abstract**

This paper investigates the characteristics of the optimal incentive contracts when the principal is also a productive agent. In this bilateral moral hazard framework, the two requirements should be satisfied in designing an incentive scheme. One is the agent's incentive provision and the other is the principal's incentive provision. Because of the trade off between these two incentive provisions, only the second best is obtainable if the incentive contract should be based only on the total output as in the small profit sharing firms. We show that the simple linear sharing rule often observed in the real world achieves such a second best efficiency. Last, we present an example and provide a simple but interesting comparative statics results on the incentive schemes.

## 1. Introduction

In the traditional principal-agent problem, it is assumed that the principal delegates the agent to act on behalf of her interest, which

* This is an outgrowth from one part of my lecture notes on Informational Economics. The readers who are interested in the updated manuscript can request it from the author.
Please send the e-mail to the following address: yutaka@mt.tama.hosei.ac.jp

cannot be observed by her. That brings about an incentive problem. In order to mitigate it, the principal delegates the agent all the functions as productive inputs and designs an incentive contract that is based on common observables[1], such as outputs, realized costs, profits, which are correlated with the agent's hidden action choice.

A well-known fact is that there are many incentive contracts that give a socially efficient outcome when the principal and the agent are risk neutral. When both parties are risk neutral, risk sharing is not an issue but only the incentive provision for the agent's action choice is an issue in designing a sharing rule. One simple example that generates the first best solution is giving the agent all the authorities and risk of production and just charging a fixed rent[2]. When the agent is risk averse, not only the incentive provision but also the optimal risk sharing between the principal and the agent is an important issue in designing an optimal incentive contract. There is a trade off between two alignments. Thus, the socially efficient outcome cannot be obtained in this case. This is also a well-known fact.

Rather, in this model, we shall assume that the principal also participates in the production process and invests her own effort that is not observable to the agent. The assumption that the principal accesses to the production process looks uncommon but more realistic. Usually, the principal provides the workers with productive machinery, the quality of which is related to the principal's effort input and sometimes not observable to the workers. The principal's advertising investment or lobbying effort is another example. Advertising investment is not usually observable to the worker.
Further, even though the advertisement itself may be observable, it may not be verifiable without costs. In this framework, not only the incentive provision for the agent's action choice but also the principal's credible commitment on her effort level is an issue in designing an incentive scheme.

This framework is rather similar to the literature that discuss moral-hazard problem in team production in the sense that *after* the principal sets up the incentive contract as a Stackelberg leader, she herself participates in the production as a team member. So, the principal here can be viewed as *a principal agent.* In such situations, in the case where only the contract based on the total output is feasible, there will be a trade off in designing an incentive contract that the principal should provide the agent with the incentive to take an optimal action and at the same time she should *credibly commit herself* to take her own best action. Giving incentive to the agent to work hard *always hurts* her own credible commitment to some extent. Thus, only the second best efficiency can be attainable. In section2, we show that there always exists a simple linear sharing rule that obtains such a second best. Also, it is discussed that each party's sharing proportion under this sharing rule crucially depends upon each party's marginal contribution to total production.

In section 3, with a simple example, it is shown that the agent's sharing proportion increases as his relative contribution to production gets larger. When the agent's contribution to total production gets more important, more stress should be given on the agent's action in the action combination. Thus, the agent's sharing portion should increase to impose more weight on the agent's action incentive. However, the essential point is roughly discussed that the additional requirement for the principal's credible commitment about her optimal effort restricts her degree of freedom in designing a contract.

In section 4, we discuss the mitigation of the commitment problem under the repeated relationships. In section 5, we conclude the paper.

## 2. The Model

A risk-neutral principal owns a firm. She hires a risk-neutral agent and designs a wage contract. $a \in A = [0, \infty)$ and $e \in E = [0, \infty)$ denote the efforts provided by the agent and the principal, respectively. $Q: A \times E \rightarrow R$ represents a composite function of both efforts, i.e., $q = Q(a, e)$. $Q(a, e)$ is continuous and differentiable in $a$ and $e$. It is assumed that both efforts are productive, I.e., $Q_a > 0$ and $Q_e > 0$, where subscripts mean partial derivatives. The output $y$ that is commonly observable is a function of the effort composition $q$ and the state of nature $\theta \in \Theta$, that is, $y = Y(Q(a, e), \theta)$, where $Y$: $Q \times \Theta \rightarrow R$[3]. Randomness of $y$ can be suppressed by parameterization.

$F(y \mid q(a, e))$ denotes the output distribution conditional on the effort composition and $f(y \mid q(a, e))$ the output probability density function.

The agent's wage scheme $w$ is based on the observable (verifiable) $y$. The price of output is normalized by one thus $y$ can be thought as cash flow. $\Psi(e)$ and $C(a)$ denote the disutilities of the principal's and the agent's efforts, respectively, with $\Psi'(e) \geq 0$, $\Psi''(e) > 0$ and $C'(a) \geq 0$, $C''(a) > 0$. Further, it is assumed that $\Psi(0) = C(0) = 0$ and $\Psi'(0) = C'(0) = 0$, $\Psi'(\infty) = C'(\infty) = \infty$.

Assumption 1

$f$ is twice differentiable in $q$.

**Assumption 2**

$F(y \mid q_1) \geq F(y \mid q_2)$, $\forall y$, $q_1 \leq q_2$

**Assumption 3**

$F(y \mid Q(a, e))$ is convex in $a$ and $e$.

Assumption 1 is needed to guarantee the existence of the agent's and the principal's optimal effort choices. Assumption2 denotes the effort composition $q$ is productive in the sense of first order stochastic dominance. Assumption 3 is given for the uniqueness of the optimal action combination $(a, e)$ and means decreasing marginal productivity of $a$ and $e$ in probability sense[4].

### 2.1 The First Best

Let us temporarily assume that the principal can observe the agent's effort choice directly. The principal's optimization problem is

$$\underset{w(y),\, a,\, e}{\text{Max}} \quad \int (y-w(y))f(y \mid Q(a, e))dy - \Psi(e)$$

$$\text{s.t} \quad \int w(y)f(y \mid Q(a, e))dy - C(a) \geq \bar{U}$$

The constraint is the agent's participation constraint that requires his expected utility obtained from this production process not less than that of other alternatives denoted by $\bar{U}$. Since the principal can observe the agent's action choice, a forcing contract is possible.

It is obvious that at the optimum

$$\int w(y)f(y \mid Q(a, e))dy - C(e) = \bar{U} \tag{1}$$

Thus, the principal's optimizing problem is

$$\underset{w(y),\, a,\, e}{\text{Max}} \quad E[y \mid Q(a, e)] - \Psi(e) - C(a) - \bar{U}$$

where $E[y \mid Q(a, e)] \equiv \int yf(y \mid Q(a, e))dy$

Assumption 3 and the convexity of $\Psi(e)$ and $C(a)$ guarantees the uniqueness of the optimal effort combination $(a^*, e^*)$. At the optimum, the principal chooses her effort level $e^*$ that satisfies $\partial E[y \mid Q(a^*, e^*)]/\partial e = \Psi'(e^*)$ and forces the agent to choose $a^*$ such that $\partial E[y \mid Q(a^*, e^*)]/\partial a = C'(a^*)$ where the primes denote the first derivatives. Only the requirement imposed on $w^*(y)$ is to satisfy (1). One of the simple example of $w^*(y)$ is a fixed fee, that is, $w^*(y) = \bar{U} + C(a^*)$, $\forall y$. With a fixed wage scheme, the principal remains *a residual claimant* and obviously has an incentive to choose $e^*$.

### 2.2 The Second Best

Now assume that the principal cannot observe the agent's effort choice and also monitoring it is impossibly costly. A forcing contract is impossibly costly. Only what she can do is to give the agent incentive to choose an optimal action through a wage contract $w(y)$. In a traditional principal-agent model, where the agent is the only party who provides the effort inputs, it is well known that charging the agent a fixed rent and giving all the reminder of the output generates the first best solution when the agent is risk-neutral. However, when both the principal and the agent invest their own productive inputs together, such a wage contract cannot implement the first best outcome, since the principal should not only give the agent incentive to do optimally but also commit herself to do optimally through the wage contract in this bilateral moral-hazard setting. Under the contract that is strategically equivalent to selling the firm to the agent, the principal cannot convince the agent that she does her best because she (the principal!) only gets a fixed rent. Thus, in this bilateral moral hazard setting, the principal's designing of an incentive scheme is substantially constrained by two incentive problems. One is the agent's incentive constraint, and the other is her own incentive constraint. Hence, the principal's problem becomes

$$\max_{w(y),\, a,\, e} \int (y-w(y))f(y \mid Q(a, e))dy - \Psi(e)$$

$$\text{s.t(i)} \quad \int w(y)f(y \mid Q(a, e))dy - C(a) \geq \bar{U}$$

$$\text{(ii)} \quad (a, e) \text{ satisfies}$$

$$a \in \arg\max \int w(y)f(y \mid Q(a, e))dy - C(a)$$

$$e \in \arg\max \int (y-w(y))f(y \mid Q(a, e))dy - \Psi(e)$$

The second constraint requires the effort combination $(a, e)$ be a Nash equilibrium that satisfies two incentive constraints, while the first constraint denotes the agent's participation constraint as before.

It is easily shown that at the (second best) optimum the participation constraint is also binding, I.e.,

$$\int w(y)f(y \mid Q(a, e))dy - C(a) = \bar{U} \qquad (2)$$

Let us denote $W_{\bar{U}}$ as the set of $w(y)$ satisfying (2), then the principal's problem is to choose $w(y) \in W_{\bar{U}}$ that maximizes the following problem:

$$\underset{w(y) \in W_{\bar{U}},\, a,\, e}{\text{Max}} \quad E[y \mid Q(a, e)] - \Psi(e) - C(a) - \bar{U}$$

$$\text{s.t} \quad (a, e) \text{ satisfies}$$

$$a \in \arg\max \int w(y) f(y \mid Q(a, e)) dy - C(a)$$

$$e \in \arg\max \int (y - w(y)) f(y \mid Q(a, e)) dy - \Psi(e)$$

Since there is no $w(y)$ in the maximand, we can easily conjecture that a specific contractual form of $w(y)$ does not matter in deciding the principal's expected utility only if it induces the same effort combination as a Nash equilibrium and in $W_{\bar{U}}$. Further, by Assumption 3 and $\Psi''(e) > 0$, $C''(a) > 0$, the constraints that guarantee the Nash equilibrium action can be replaced by

$(a, e)$ satisfies

$$\int w(y) f_Q(y \mid Q(a, e)) \frac{dQ}{da} dy - C'(a) = 0 \qquad (3)$$

and

$$\int (y - w(y)) f_Q(y \mid Q(a, e)) \frac{dQ}{da} dy - \Psi'(e) = 0 \qquad (4)$$

Let us denote $C^0$ is the set of $(a^0, e^0)$ that there exists at least one wage contract $w(y) \in W_{\bar{U}}$ under which $(a^0, e^0)$ is induced as a Nash equilibrium action satisfying (3) and (4).

**Lemma 1:**
$C^0$ is not an empty set.

**Proof:**
Let $w^0(y) = \bar{U}$, $\forall y$. It is obvious that $w^0(y) \in W_{\bar{U}}$ since $C(0) = 0$. Since $w^0(y)$ is a fixed wage contract, it does not give the agent incentive to work hard, so the agent chooses. $a = 0$. But the principal has

full incentive to work hard under this contract, that is, $e = e^*(0)$, where $e^*(0)$ denotes the principal's effort choice of full incentive when the agent chooses $a = 0$. Obviously, $(0, e^*(0))$ satisfies ( 3 ) and ( 4 ). QED

Another possible example is that $w^0(y) = y - k$, where $k = E[y \mid Q(a^*(0), 0)] - [\bar{U} + C(a^*(0))]$. Under this contract, the agent has full incentive to work hard, that is, $a = a^*(0)$, where $a^*(0)$ denotes the full incentive of effort choice when the principal chooses $e = 0$. But since the principal is given a fixed rent $k$, he does not have any incentive to work hard, that is, $e = 0$. Since $w^0(y) \in W_{\bar{U}}$ and $(a^*(0), 0)$ satisfies ( 3 ) and ( 4 ), $(a^*(0), 0)$ is a Nash equilibrium action combination under wage contract $w^0(y)$.

**Lemma 2:**
$C^0$ is a bounded set.

To $C^0$ see is a bounded set, it is enough to show that $0 \leq a^0 \leq a^*(e^0) < \infty$ and $0 \leq e^0 \leq e^*(a^0) < \infty$, $\forall (a^0, e^0) \in C^0$, where $a^*(e^0)$ denotes the agent's full incentive action when the principal can credibly commit herself to choose $e^0$ and similarly $e^*(a^0)$ denotes the principal's full incentive one.

(i) $0 \leq a^0 \leq a^*(e^0) < \infty$

Since $a^*(e^0)$ is the agent's full incentive action given $e^0$, at $a^*(e^0)$

$$E_Q[y \mid Q(a^*(e^0), e^0)] \frac{\partial Q(a^*(e^0), e^0)}{\partial a} = C'(a^*(e^0))$$

where $E_Q \equiv dE[y \mid Q]/dQ$. It is obvious that $a^*(e^0) < \infty$ since $E[y \mid Q]$ is concave in $a$ and $e$ and $C'(\infty) = \infty$.

However, when the principal cannot credibly commit $e^0$, then she can only commit herself to choose $e^0$ through the wage contract that satisfies ( 4 ). Using ( 3 ), ( 4 ) reduces to

$$E_Q[y \mid Q(a^0, e^0)]\frac{\partial Q(a^0, e^0)}{\partial e} - C'(a^0)\frac{\dfrac{\partial Q(a^0, e^0)}{\partial e}}{\dfrac{\partial Q(a^0, e^0)}{\partial a}} = \Psi'(e^0)$$

Since $\Psi'(e^0) \geq 0$, $E_Q[y \mid Q(a^0, e^0)]\partial Q(a^0, e^0)/\partial a - C'(a^0) \geq 0$.
By the concavity of $E(\cdot)$ and the convexity of $C(a)$ in $a$, we obtain $0 \leq a^0 \leq a^*(e^0) < \infty$.

(ii) $0 \leq e^0 \leq e^*(a^0) < \infty$
This can be shown just the same.

One thing to note from the proof of Lemma 2 is that any contract that induces positive incentive on one side in Nash equilibrium suffers incentive loss on the other side.

**Proposition 1:**
The First best action $(a^*, e^*)$ is not in $C^0$

**Proof:**
Given the principal's first best action $e^*$, let us assume that $w^0(y)$ gives the agent full incentive to choose $a^*$, that is, $a^*$ satisfies

$$\int w^0(y) f_Q(y \mid Q(a^*, e^*))dy \frac{\partial Q(a^*, e^*)}{\partial a} = C'(a^*)$$

Then,

$$\int w^0(y) f_Q(y \mid Q(a^*, e^*))dy = \frac{C'(e^*)}{\dfrac{\partial Q(a^*, e^*)}{\partial a}} \qquad (5)$$

If $(a^*, e^*)$ is a Nash equilibrium action under $w^0(y)$, by plugging (5) into (4), we should obtain

$$\frac{\partial E[y \mid Q(a^*, e^*)]}{\partial e} - \frac{\dfrac{\partial Q(a^*, e^*)}{\partial e}}{\dfrac{\partial Q(a^*, e^*)}{\partial a}} C'(a^*) = \Psi'(e^*) \qquad (6)$$

Since $a^* > 0$, the second term in the left- hand side in ( 6 ) is strictly positive and ( 6 ) cannot be true because by the definition of $e^*$

$$\frac{\partial E[y \mid Q(a^*, e^*)]}{\partial e} = \Psi'(e^*)$$

Proposition 1 says that in this moral hazard framework, there is no wage contract that implements the first best outcome. Intuitively, this bilateral moral hazard setting is *team production*, even though one of the team members is entitled to a principal who designs a contract, and extracts all the surplus. In this point, the essence of Proposition 1 is nothing but that founded in Holmstrom (1982) which deals with moral hazard problem in team production. He shows that any sharing rule that satisfies budget balancing constraint cannot achieve the first best result in team production[5]. Indeed, the budget balancing constraint is also imposed here and the wage contract $w(y)$ is a dividing rule of a given budget $y = Y(Q(a, e), \theta)$. What we can derive from ( 3 ) and ( 4 ) is that any wage contract that gives the agent strictly positive incentive also hurts the principal's ability of commitment to some extent if the budget balancing constraint is imposed.

**Proposition. 2:**

If $(a^0, e^0) \in C^0$, then any contract $w^0(y) \in W_{\bar{U}}$ that satisfies either ( 3 ) or ( 4 ) satisfies the other constraint and finally guarantees $(a^0, e^0)$ as a Nash equilibrium action.

**Proof:**

If $(a^0, e^0)$ is in $C^0$, then there exists at least one $w^0(y) \in W_{\bar{U}}$ such that

$$\int w^0(y) f_Q(y \mid Q(a, e))dy \frac{\partial Q(a^0, e^0)}{\partial a} - C'(a^0) = 0 \qquad (7)$$

and

$$\int (y - w^0(y)) f_Q(y \mid Q(a^0, e^0))dy \frac{\partial Q(a^0, e^0)}{\partial a} - \Psi'(e^0) = 0 \ (8)$$

Using ( 7 ), ( 8 ) changes to

$$\frac{\partial E[y \mid Q(a^0, e^0)]}{\partial e} - \frac{\dfrac{\partial Q(a^0, e^0)}{\partial e}}{\dfrac{\partial Q(a^0, e^0)}{\partial a}} C'(a^0) = \Psi'(e^0) \qquad (9)$$

and (9) is independent of $w^0(y)$.

As a result, Proposition 2 tells that the contractual form of the wage schemes does not matter in making $(a^0, e^0)$ as a Nash equilibrium if it induces either the principal $e^0$ with the agent's given $a^0$, or the agent $a^0$ with the principal's given $e^0$, when $(a^0, e^0)$ already belongs to $C^0$[6].

**Proposition 3:**
If $(a^0, e^0)$ is already in $C^0$, then there always exists a linear wage contract $w(y) = r \cdot y + s \in W_{\bar{U}}$, where $0 \leq r \leq 1$, under which $(a^0, e^0)$ is induced as a Nash equilibrium.

**Proof:**
Since $(a^0, e^0)$ is already in $C^0$, to show that $w(y) = r \cdot y + s \in W_{\bar{U}}$ induces $(a^0, e^0)$ as a Nash equilibrium action combination, it is enough to show that $w(y) = r \cdot y + s \in W_{\bar{U}}$ satisfies (7) by Proposition 2. In order for $w(y) = r \cdot y + s$ to be in $W_{\bar{U}}$

$$\int (r \cdot y + s) f(y \mid Q(a^0, e^0)) dy - C(a^0) = \bar{U} \qquad (10)$$

From (10), we obtain the fixed payment part of the wage

$$s = \bar{U} + C(a) - rE[y \mid Q(a^0, e^0)] \qquad (11)$$

Equation (11) tells that $s$ will be determined when $r$ is determined. Thus the only requirement to be shown is that there exists $0 \leq r \leq 1$ satisfying

$$\int (r \cdot y + s) f_Q(y \mid Q(a^0, e^0)) dy \frac{\partial Q(a^0, e^0)}{\partial a} = C'(a^0) \qquad (12)$$

that is, under $w(y) = r \cdot y + s$, the agent has an incentive to choose $a^0$ when the principal surely chooses $e^0$. (12) reduces to

$$r\frac{\partial E[y \mid Q(a^0, e^0)]}{\partial a} = C'(a^0) \tag{13}$$

From Lemma 2, $0 \leq a^0 \leq a^*(e^0)$ and we can easily see from (13) that as $r$ increases from zero to one, the agent's incentive compatible action choice also continuously increases from zero to $a^*(e^0)$ given the principal's $e^0$. Thus there exists $r$ that induces the agent to take $a^0$ given the principal's $e^0$.

Since, as mentioned above, the principal will be indifferent with any contract only if it induces the same action combination as a Nash equilibrium, Proposition 3 tells that if $(a^0(*), e^0(*))$ is the second best action combination in $C^0$ that maximizes the principal's expected utility, then we always find *a linear wage contract* $w^*(y) = r^* \cdot y + s^*$ with which such optimality can be achieved[7].

**Proposition 4:**
At the second best optimum, $a^0(*) > 0$ and $e^0(*) > 0$.

**Proof:**
Assume that $e^0(*) = 0$ at the optimum, that is, the principal prefers to design a contract that induces $(a^*(0), 0)$ as a second best Nash equilibrium action. From Proposition 3, one of such contracts is $w(y) = r \cdot y + s$ where $r = 1$ and $s = \bar{U} + C(a^*(0)) - E[y \mid Q(a^*(0), 0)]$ to guarantee $w(y) \in W_{\bar{U}}$. When the principal changes the contract *by marginally decreasing* the agent's sharing proportion (the piece-rate) $r$, the expected utility increases by the amount of

$$\left[\frac{\partial E[y \mid Q(a^*(0), 0)]}{\partial a} - C'(a^*(0))\right]\frac{da}{dr} + \left[\frac{\partial E[y \mid Q(a^*(0), 0)]}{\partial e} - \Psi'(0)\right]\frac{de}{dr}$$

The first bracket is zero by the definition of $a^*(0)$ and the second bracket is obviously positive. But, when $r = 1$, from (8), we obtain

$$\frac{de}{dr} = -\frac{\dfrac{\partial E[y \mid Q(a^*(0),\, 0)]}{\partial a}}{\Psi''(0)} < 0$$

Thus *by marginally decreasing* $r$, the principal can obtain more expected utility than she does with $(a^*(0),\, 0)$. So, this is contradiction. The proof of $e^0(*) > 0$ can be shown by the same way.

Proposition 4 together with Proposition 3 leads to the optimality of a linear contract in a bilateral moral hazard setting. In a traditional agency model, where the principal is excluded from the production process, a linear contract that treats the agent as a residual claimant has been suggested as one of the first best contract. The principal who is not engaged in production line can get the first best solution by charging the agent a fixed rent, combined with the piece-rate of 100 %. However, if the principal gets involved in the production process, by investing his own effort, such a contract does not work efficiently. When the principal also participates in the production process, she herself behaves as an agent after the wage contract is settled and in order to induce the agent to take the right action, she must convince him about her action strategy. But receiving a fixed rent cannot perform this job effectively, resulting in the incentive loss on the side of herself. Consequently, in such a team production between the principal and the agent, there is no way for the principal to establish full incentive without hurting her own commitment. Only the second best will be obtainable and simple sharing rules we frequently observe in small profit sharing firms perform well.

**The possibility of the corner solution: $r^* = 1$**

The result of the proposition 4 depends crucially upon the assumption of $\Psi'(0) = 0$. If we relax this assumption and assume that $\partial E[y \mid Q(a^*(0),\, 0)]/\partial e \leq \Psi'(0)$, we will get the corner solution $r^* = 1$. This is because the principal cannot increase her expected utility by marginally decreasing the piece-rate $r$ from 1, when the marginal cost of the principal *at zero effort* is sufficiently high. In this case, the optimal linear scheme is $r^* = 1$ and $s^* = \bar{U} + C(a^*(0)) - E[y \mid Q(a^*(0), 0)]$, which induces $a^0(*) = a^*$ and $e^0(*) = 0$ at the second best[8]. This is just the traditional principal-agent situation, where the principal does not work.

Finally, specifying actual proportions to be shared between the principal and the agent depends on the production function, that is, on the degree that each player marginally contributes to the outcome. Next section clarifies the argument and provides comparative statics with a simple example.

## 3. An example

The principal installs his effort $e$ and the agent provides his effort $a$. Both investments of inputs are assumed to be unobservable by the other party. True amount (or the quality) of products is a function of both efforts, $Q(e, a) = \alpha \ln a + (1-\alpha)\ln e$, but the measured products is the true amount plus error term, that is, $y = Q(e, a) + \theta$ and $E(\theta) = 0$, where $E$ denotes the expectation operator. Since the price of products is normalized to one, $F(y \mid Q(e, a))$ denotes the revenue distribution with given effort choices of the principal and the agent. We specify the effort costs of the principal and the agent as follows; $\Psi(e) = \phi \cdot e$, and $C(a) = c \cdot a$,

The principal maximizes her expected profit denoted by

$$\alpha \ln a + (1-\alpha)\ln e - Ew(y) - \phi \cdot e \qquad (14)$$

When the principal can enforce the worker the amount of effort, the optimal combination of inputs is $(a^*, e^*) = (\alpha/c, 1-\alpha/\phi)$ that generates the marginal benefit of each input equals to the marginal cost, and the principal pays $c \cdot a^*$ as a lump sum payment to the agent[9].

However, when the principal cannot force the amount of effort directly or indirectly, the principal should consider the incentive constraint of the agent. From proposition 3, any combination of Nash equilibrium actions can be found by a linear sharing rule $w(y) = r \cdot y + s$. Thus, if $(a^0, e^0) \in C^0$, then $(a^0, e^0)$ should satisfy

$$r\frac{\alpha}{a^0} = c \tag{15}$$

$$(1-r)\frac{1-\alpha}{e^0} = \phi \tag{16}$$

In Figure 1, the principal's indifferent curves are drawn on $(a, e)$ space and the set of $(a^0, e^0) \in C^0$ is drawn is drawn as a straight line, since $e^0 = (1-\alpha)/\phi \cdot (1-(ca^0/\alpha))$ from (15) and (16). Point F denotes the first best action combination $(\alpha/c, 1-\alpha/\phi)$ and Point E denotes the second best $(a^0(*), e^0(*))$, which gives the maximum level of expected utility to the principal among $(a^0, e^0) \in C^0$. In order to find the second best action combination $(a^0(*), e^0(*))$, it is enough to find the corresponding $r^*$ since each action combination in the straight line corresponds to each value of $0 \leq r \leq 1$ in $w(y) = r \cdot y + s$. For example, as $r$ increases from zero to one, the action combination moves down along the line from A to B.

By plugging (15) and (16) into (14) and taking the derivatives with respect to $r$, we get the first order condition on the optimal shar-

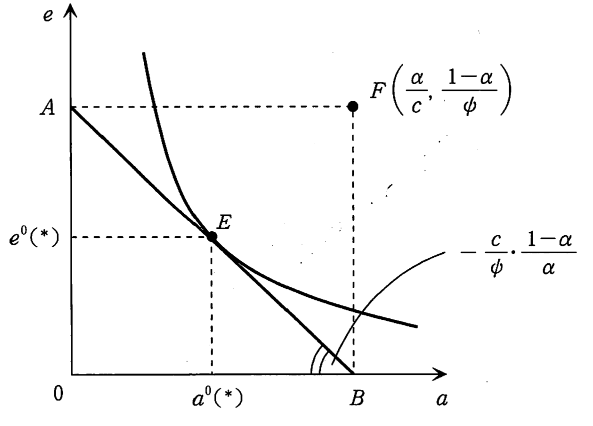


**Figure 1.** An example: The second best solution

ing portion (piece-rate) $r^*$

$$\frac{\alpha}{r^*}-\frac{1-\alpha}{1-r^*}-\alpha+(1-\alpha)=0 \tag{17}$$

Solving (17), we get

$$r^*=\frac{\alpha-\sqrt{\alpha(1-\alpha)}}{2\alpha-1} \tag{18}$$

Now computing the marginal rate of substitution between $a$ and $e$ on the indifference curve of the principal, we get $MRS_{ae}=(\alpha/a^0-c)/((1-\alpha)/e^0-\phi)$. Similarly, we get $MRT_{ae}=(1-\alpha)c/\alpha\phi$ as the marginal rate of transformation between $a$ and $e$ on $C^0$. We can easily check that the marginal rate of substitution equals to the marginal rate of transformation at $(a^0(*), e^0(*))$.
It is easy to check $0<r^*<1$ and from (18) we obtain

$$\frac{dr^*}{d\alpha}=\frac{1-2\sqrt{\alpha(1-\alpha)}}{2\sqrt{\alpha(1-\alpha)}(2\alpha-1)^2}>0 \tag{19}$$

since both the numerator and the denominator are positive for $0<\alpha<1$. (19) shows that as the relative proportion the worker's labor input contributes to the production increases, his sharing proportion also increases. Some readers may think that an intuitive ex-

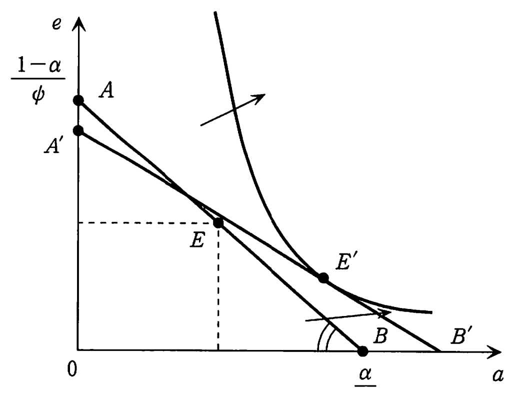


**Figure 2.** The effect of increasing $\alpha$ on the optimal solution

planation for this result may be that it is *fair* that the worker receives more when he contributes more. But this intuition ignores the *incentive aspect* of the wage contract that lies behind. The right interpretation is that *the incentive weight* should move to the worker when he relatively contributes more to the output. To do so, the worker's sharing proportion (the piece-rate in the linear contract) should increase. As a result, his effort provision increases in equilibrium.

If $\alpha=1$, which represents the traditional principal-agent model, $r^*=1$ and the first best will be achieved. Thus, this model includes the traditional principal agent model as a special case. Finally, the fixed wage: $s^*$ will be decided by satisfying the agent's individual rationality constraint, that is,

$$\int (r^*y+s^*)f(y\mid Q(e^0(*),\ a^0(*)))dy = c\cdot a^0(*) \qquad (20)$$

## 4. Commitment problem and its mitigation under Repeated Relationships

If we investigate this problem in the repeated game framework, the principal will be able to solve the commitment problem under certain conditions. Then, the only observable variable is the total output and so, the model is an example of repeated games with imperfect monitoring. In this model, as Green and Porter (1984) studied, deviations cannot be detected perfectly: rather than observing the other parties' effort choices, each party observes only the total output, which is buffeted by an measurement error each period. In this setting, the agent cannot tell whether a low total output occurred because the principal cheated or because there was an adverse error. Green and Porter examine trigger -output equilibria, in which any output below a critical level triggers a punishment period during which the principal and the agent play their static Cournot Nash efforts under the static optimal incentive schemes. In equilibrium, neither the principal nor the agent ever deviates. Nonetheless, an especially bad shock can cause the output to fall below the critical level, triggering a punishment period. Since punishments happen by accident, *infinite punishments* in the trigger-strategy analysis *are not optimal. Two-phase strategies* (*stick and carrot strategies*) of the kind analyzed by Abreu (1986) and Abreu, Pearce, and Stacchetti (1986) show that they can be optimal.

## 5. Concluding remarks

This paper investigates the characteristics of a wage contract and the impossibility of the first best, when the principal and the agent, after the wage contract is settled, constitute a team and produce the output together by investing their own efforts that are not observable with each other. In a traditional principal-agent model, where the principal does not play a role of another productive agent, the only issue is how the incentive can be induced from the agent. However, in a bilateral moral hazard framework like team production, another issue is added, that is, how the principal can convince the agent her own incentive (she does best for the agent). Both incentives are trade-off and only the second best is possible even with the risk neutral agent. It is shown that a simple linear contract we observe frequently in the real world can implement such a second best efficiency.

Actual sharing proportion of each party crucially depends on the characteristics of production function, especially, the ratio that each party contributes to the production. A simple example shows that the agent's sharing proportion increases as the relative contribution to production gets larger to impose more weight on the agent's effort incentive.

Last, we assumed the verifiability of the total output: $y$. If we assume the unverifiability of $y$, the solution of the underinvestment problem will become more difficult. Renegotiation design literature (for example, Aghion, Dewartripont, and Ray (1994)) investigates such kind of problem. In the part (I), we examined the incentive inducement through an *endogenous discontinuity generated in the renegotiation process* in an incomplete contract model with unverifiable information. Based on these results, we will pursue a more integrated and generalized theory of 'Commitment, Renegotiation and Contract Incompleteness'.

**Notes**

1 In the main text, we assume that these common observables can be verified also to the court without cost. In other words, there is a verifiable information, on which the principal can write an incentive contract. In this simpler framework, we examine the commitment problem.

2 This incentive contract performs well unless the agent's limited liability matters.

3 Note that $q$ and $\theta$ independently affect $y$. If $\theta$ affects $q$, then the analysis will be more complicated. A possible example is $q = a\theta + e$. As for it, see Nalebuff and Stiglitz (1983).

4 Assumption 3 is stronger than the assumption that $E[y \mid Q(a, e)]$ is concave in $a$ and $e$. In the no uncertainty case, it is enough for the uniqueness of optimal effort combination.

5 Holmstrom (1982) breaks the budget balancing constraint and by giving group incentives with a penalty scheme achieves the first best outcome.

6 Proposition 2 critically hinges on the assumption that the principal and the agent are risk neutral. If the agent is risk averse, for example, the contractual form does matter in achieving optimality.

7 We are not restricting the principal's contract spaces to include only linear contracts. Rather, we allow the principal to choose arbitrary contracts but ask whether there is an equilibrium that is linear. Many other equilibria exist besides this, but the linear equilibrium has both the fascinating efficiency property and the simplicity such that it can often be observed in reality. Theoretically, this point is the same as the derivation of a *linear* Bayesian Nash equilibrium of the double auction, as in Chatterjee and Samuelson(1983), and Hall and Lazear(1984)

8 This type of optimal corner solution can be obtained also in other more advanced models. For example, Suzuki, Y (1998) derives a corner solution as an optimal solution in the context of control of the dynamic competition among the agents.

9 The individual rationality level of the agent is normalized to zero.

**References**

Aghion, P., Dewartripont, M., and P. Ray (1994) "Renegotiation Design with Unverifiable Information", *Econometrica,* 62: 257–282.

Green, E. and R.Porter, 1984. "Noncooperative Collusion Under Imperfect Price Information". *Econometrica* 52: 87–100.

Chaterjee, K. and W. Samuelson, 1983. "Bargaining under Incomplete Information", *Operations Research.* 31: 835–51.

Gibbons, R. (1992) *Game Theory for Applied Economists.* Princeton University Press. Princeton.

Hall, R. and E. Lazear (1984) "The Excessive Sensitivity of Layoffs and Quits to Demand". *Journal of Labor Economics.* 2: 233–57.

Holmstrom, B. (1982) "Moral Hazard in Teams." *Bell Journal of Economics,* 13, 324–340.

Macmillan, J. (1992) *Games, Strategies, and Management: How Managers Can Use Game Theory to Make Better Business Decisions.* New York and Oxford: Oxford University Press.

Nalebuff, B. and J. Stiglitz. (1983) "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell Journal of Economics.* 14. 21–43.

Suzuki, Y. (1998) "Capital Accumulation through 'Controlled Competition': Two Stage Tournaments as a Strategy for Firm Growth" I.C.E.S working paper No. 60 presented at the Econometric Society 1998 North America Summer Meeting, held at Montreal, Canada, June 25–28.