

## 1 Model Setup

To study moral hazard, we can use the same basic model introduced in the previous lecture notes but make one crucial change: effort is no longer verifiable. This could be for one of two reasons. First, the principal may simply not observe the effort the agent exerts on the project. Second, the principal could be well aware of how much effort the agent exerts, but be unable to prove what effort level the agent exerts to a court. For whatever reason, when effort is non-verifiable, a contract can no longer stipulate an effort level.

**Definition 1** *When effort is non-verifiable, a contract is a list of payments  $\{w_i\}_{i=1}^N$  from the principal to the agent corresponding to states of the world  $\{x_i\}_{i=1}^N$ .*

Because effort is no longer contractible, effort is a choice variable of the agent. Figure 1 is analogous to Figure 1 from the previous lecture notes. Whereas before, the principal offered a contract to the agent that the agent either rejected or accepted, the agent now has two choices. First, he accepts or rejects the contract. Then, after accepting the contract, he chooses what effort level to exert.

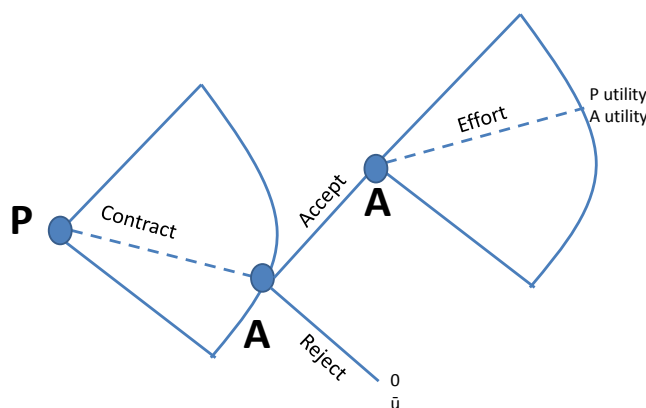


Figure 1: Moral Hazard Game Tree

Since effort is now a choice variable of the agent, the question becomes: what effort level will the agent choose if he accepts the contract? We assume that, after accepting the contract, the agent chooses an effort level to maximize his expected utility. In other words, the agent solves the problem

$$\max_e \sum_{i=1}^N p_i(e) u(w_i) - v(e),$$

---

<sup>1</sup>by Stephen Hansen. Feel free to send any comments or suggestions to [stephen.hansen@upf.edu](mailto:stephen.hansen@upf.edu).

where the payments  $\{w_i\}_{i=1}^N$  are given. We will denote the agent's utility maximizing effort level as<sup>2</sup>

$$e = \arg \max_{\tilde{e}} \sum_{i=1}^N p_i(\tilde{e}) u(w_i) - v(\tilde{e}). \quad (1)$$

## 2 General Optimization Problem

When the principal designs a contract, there are now two constraints. First, just as in the basic model, he must get the agent to participate in the contract. Second, he must take into account the fact that the agent is going to choose effort to maximize expected utility. We can write the problem as

$$\max_{\{w_i\}_{i=1}^N, e} \sum_{i=1}^N p_i(e) B(x_i - w_i) \quad (2)$$

$$\text{such that } \sum_{i=1}^N p_i(e) u(w_i) - v(e) \geq \bar{u} \quad (3)$$

$$e = \arg \max_{\tilde{e}} \sum_{i=1}^N p_i(\tilde{e}) u(w_i) - v(\tilde{e}) \quad (4)$$

Constraint (3) is the participation constraint, just as it appeared in the basic model. Constraint (4) is the *incentive compatibility constraint*, and demands that the agent's effort choice is the solution to his utility maximization problem. In other words, the effort level that the principal anticipates has to be compatible with the agent's actual effort level.

Note that although a contract only specifies a list of payments to the agent, the principal maximizes over these payments as well as the effort level. Although the agent's effort is not contractible, it is still relevant for the principal's expected utility, so it must be included in the optimization problem. The wage payments that the principal selects have two effects: they directly impact  $B(x_i - w_i)$  and they induce an effort level  $e$ . The incentive compatibility constraint captures this last effect.

The optimization problem written above is in general very complex to solve, and we are not going to try. Instead, we will solve two simpler versions of the problem: one with only two effort levels and one with continuous effort and specific functional forms for the output distribution and the agent's utility. Another simplification we will make is to assume that the principal is risk neutral and the agent risk averse. We do this because under these conditions the optimal contract with verifiable effort is to pay the agent the same wage in all states of the world. This simple contract provides a useful benchmark against which to compare the optimal contract with moral hazard.

---

<sup>2</sup>The "arg max" stands for the "argument that maximizes." Equation (1) says that  $e$  is the value of  $\tilde{e}$  that maximizes  $\sum_{i=1}^N p_i(\tilde{e}) u(w_i) - v(\tilde{e})$ . We will assume that  $e$  is unique. The book does not make this assumption and writes  $\in$  rather than  $=$  for equation (1). This notation means that  $e$  is in the set of maximizers of  $\sum_{i=1}^N p_i(\tilde{e}) u(w_i) - v(\tilde{e})$ .

### 3 Two Effort Levels

Suppose that effort can either be  $e_L$  or  $e_H$  where  $e_L < e_H$ . To simplify notation, let  $p_i(e_L) = p_L^i$  and  $p_i(e_H) = p_H^i$ . Within this simplified framework, we want to ask what is the optimal contract when the principal is risk neutral and the agent is risk averse. This question depends on what effort level the principal wants to implement. Suppose the principal wants to implement  $e_L$ . The optimal contract with verifiable effort is to pay the agent a constant wage that makes the participation constraint bind. With non-verifiable effort, when the wage is constant in all states of the world, the agent will choose effort level  $e_L$  since the expected utility of wealth is the same regardless of what effort he chooses, while the disutility of effort is less. So, the optimal contract for implementing  $e_L$  with verifiable effort is also the optimal contract for implementing  $e_L$  with non-verifiable effort since it is incentive compatible.

The situation changes when the principal wants to implement  $e_H$ . A contract that pays the agent the same wage in all states of the world cannot implement  $e_H$  since the agent will get a higher utility from choosing  $e_L$  than  $e_H$ . To induce the agent to choose  $e_H$ , the principal must introduce incentive pay into the contract that rewards the agent for choosing the higher effort level. However, this exposes the agent to wage uncertainty, which he does not like given that he is risk averse. This in turn means that the principal has to raise the agent's expected wage, compared to that in the optimal verifiable effort contract, in order to keep his expected utility at  $\bar{u}$ . This in a nutshell is the core issue in moral hazard contracts: the trade-off between incentives and risk-sharing.

#### 3.1 Optimization problem and solution

To analyze these issue formally, we will explicitly solve the principal's problem for implementing  $e_H$ :

$$\max_{\{w_i\}_{i=1}^N} \sum_{i=1}^N p_i^H (x_i - w_i) \quad (5)$$

$$\text{such that } \sum_{i=1}^N p_i^H u(w_i) - v(e_H) \geq \bar{u} \quad (6)$$

$$\sum_{i=1}^N p_i^H u(w_i) - v(e_H) \geq \sum_{i=1}^N p_i^L u(w_i) - v(e_L) \quad (7)$$

Unlike the problem in the previous section, here we are not optimizing over the wage payments and effort level simultaneously. Instead, we are fixing an effort level and optimizing over the wage payments only. The incentive compatibility constraint (7) therefore has a slightly different interpretation. It demands that the agent receive a higher utility from choosing  $e_H$  than  $e_L$  given a set of wage payments.

The Lagrangian for the problem is

$$\begin{aligned} \sum_{i=1}^N p_i^H (x_i - w_i) + \lambda \left[ \sum_{i=1}^N p_i^H u(w_i) - v(e_H) - \bar{u} \right] \\ + \mu \left[ \sum_{i=1}^N (p_i^H - p_i^L) u(w_i) - (v(e_H) - v(e_L)) \right]. \end{aligned}$$

Now, before we proceed to write down the Kuhn-Tucker conditions, note that the incentive compatibility constraint is not necessarily concave in the wage payments since we have made no assumption on the sign of  $p_i^H - p_i^L$ . In fact, the book is a little sloppy in applying the Kuhn-Tucker theorem not only in this instance but also in others. The footnote on page 43 actually shows a method for proving global concavity, but rather than get into the details, I will just ask you to accept the fact that the KT conditions define the solution. These are:

$$-p_i^H + \lambda^* p_i^H u'(w_i^*) + \mu^* (p_i^H - p_i^L) u'(w_i^*) = 0 \quad \forall i \quad (8)$$

$$\lambda^* \geq 0 \quad (9)$$

$$\mu^* \geq 0 \quad (10)$$

$$\sum_{i=1}^N p_i^H u(w_i^*) - v(e_H) - \bar{u} \geq 0 \quad (11)$$

$$\sum_{i=1}^N (p_i^H - p_i^L) u(w_i^*) - (v(e_H) - v(e_L)) \geq 0 \quad (12)$$

$$\lambda^* \left[ \sum_{i=1}^N p_i^H u(w_i^*) - v(e_H) - \bar{u} \right] = 0 \quad (13)$$

$$\mu^* \left[ \sum_{i=1}^N (p_i^H - p_i^L) u(w_i^*) - (v(e_H) - v(e_L)) \right] = 0 \quad (14)$$

As always, we need to figure out which constraints are binding. From (8) we have, for all  $i$ ,

$$\frac{p_i^H}{u'(w_i^*)} = \lambda^* p_i^H + \mu^* (p_i^H - p_i^L). \quad (15)$$

If we add up these  $N$  equations and use the fact that  $\sum p_i^H = \sum p_i^L = 1$ , we obtain

$$\frac{1}{u'(w_i^*)} = \lambda^* > 0.$$

Since  $\lambda^*$  is positive, we can conclude that the participation constraint is binding. If we divide through (15) by  $p_i^H$  we obtain

$$\frac{1}{u'(w_i^*)} = \lambda^* + \mu^* \left(1 - \frac{p_i^L}{p_i^H}\right). \quad (16)$$

Now, suppose that  $\mu^* = 0$ . Then, (16) tells us that  $w_i^*$  is constant. But, this is impossible since the constant-wage contract does not satisfy the incentive compatibility constraint. So, it must be the case that  $\mu^* > 0$ , implying that the incentive compatibility constraint is binding.

### 3.2 Properties of the optimal contract

Since the incentive compatibility constraint imposes a positive cost on the principal, we can conclude that the non-verifiability of effort reduces his profit compared to the verifiable effort case. The logic is the following. By equation (16) there is wage variability across states of the world in the optimal contract since in general  $\frac{p_i^L}{p_i^H} \neq \frac{p_j^L}{p_j^H}$  for  $i \neq j$ . At the same time, the participation constraint is binding, meaning that the optimal contract provides an expected utility of  $\bar{u}$  to the agent, just like the optimal contract with verifiable effort. But this means that the principal must be paying a higher expected wage to the agent with effort unverifiable than effort verifiable since the agent is risk averse,<sup>3</sup> thus reducing profit. In other words, the incentive pay associated with implementing  $e_H$  harms the principal because of the extra money he pays the agent to compensate him for the associated risk.

Equation (16) also provides an interesting relationship between pay and states of the world.<sup>4</sup>

**Result 1**  $w_i^*$  is increasing in  $\frac{p_i^H}{p_i^L}$ .

$\frac{p_i^H}{p_i^L}$  is a number between 0 and  $\infty$  that measures the likelihood that the agent exerted the high effort level given that the principal observes outcome  $x_i$ . For example, when  $p_i^L$  is near 0,  $\frac{p_i^H}{p_i^L}$  is very large, indicating a high likelihood that the agent exerted high effort. When  $p_i^H$  is near 0,  $\frac{p_i^H}{p_i^L}$  is very low, indicating a low likelihood that the agent exerted high effort.<sup>5</sup> So, the result tells us that the agent is paid a higher wage for outcomes that are more likely to be the result of high effort. Intuitively speaking, to satisfy the incentive compatibility constraint, the agent must be paid more when he exerts high effort than when he exerts low effort. Therefore, the outcomes associated with high effort must be rewarded and those associated with low effort punished.

The result also allows us to ask an interesting question: if the outcomes are such that  $x_1 < \dots < x_N$ , must it be the case that  $w_1 < \dots < w_N$ ? In other words, does the optimal contract for implementing  $e_H$  pay a higher wage for higher outcomes? Clearly this is true if and only if  $\frac{p_1^H}{p_1^L} < \dots < \frac{p_N^H}{p_N^L}$ . While some reasonable distributions of output have this property,<sup>6</sup> there is no reason why it *must* be true. Suppose there are three outcomes  $x_1 < x_2 < x_3$  with the interpretation that  $x_2$  is mediocre success,  $x_3$  is outstanding success, and  $x_1$  is outright failure. In some production processes (for example in creative industries), low effort will almost certainly generate a mediocre outcome while high effort may result in outstanding success or dismal failure. In this case  $\frac{p_1^H}{p_1^L} > \frac{p_2^H}{p_2^L}$  and  $\frac{p_3^H}{p_3^L} > \frac{p_2^H}{p_2^L}$ ,

<sup>3</sup>Let  $w$  be the constant wealth level that satisfies  $u(w) = \bar{u}$ . If  $\sum p_i^H u(w_i^*) = \bar{u}$ , then it must be the case that  $\sum p_i^H w_i^* > w$  if  $u$  is concave.

<sup>4</sup>To derive this result, note that  $u'(w_i^*)$  is decreasing in  $w_i^*$  since  $u$  is concave. This means that  $\frac{1}{u'(w_i^*)}$  is increasing in  $w_i^*$ . Clearly, the right hand side of (16) is decreasing in  $\frac{p_i^L}{p_i^H}$  and therefore increasing in  $\frac{p_i^H}{p_i^L}$ .

<sup>5</sup>Note that  $\frac{p_i^H}{p_i^L}$  is not a probability since it is not between 0 and 1.

<sup>6</sup>For example, suppose that output has a Binomial distribution  $(n, p^H)$  when  $e = e_H$  and a Binomial

meaning that the optimal contract pays a wage that first decreases and then increases in the outcome.

So far, we have argued that the moral hazard problem costs the principal (if he wants to implement  $e_H$ ) since he pays a higher expected wage than with verifiable effort. However, we can go further and say that the optimal contract for implementing  $e_H$  is *Pareto inefficient*. The reason is that the marginal rates of substitution of wealth in any two states of the world are not equal for the principal and agent. The MRS between wealth in states  $i$  and  $j$  for the principal is  $\frac{p_i^H}{p_j^H}$  and for the agent is  $\frac{p_i^H u'(w_i^*)}{p_j^H u'(w_j^*)}$  which does not equal  $\frac{p_i^H}{p_j^H}$  since  $w_i^* \neq w_j^*$ . Thus, there is a contract that would make *both* the principal and agent better off compared to the moral hazard outcome.

## 4 Moral Hazard with Continuous Effort

We will now turn to solving a particular moral hazard problem with continuous effort that is commonly used in applied papers in contract theory. Suppose that output is  $x = a + \varepsilon$  where  $a$  is effort and  $\varepsilon \sim N(0, \sigma^2)$ . Thus, not only is effort now continuous, but output is as well. The agent's preferences over wages and effort are  $u(w, a) = -e^{-r(w - \frac{1}{2}a^2)}$ . This is a constant absolute risk aversion (CARA) utility function, where the coefficient of absolute risk aversion is given by  $r$ . We will continue to assume that the principal is risk neutral, and restrict him to offering linear contracts of the form  $w = \alpha + \beta x$ , where  $\alpha$  is a fixed payment and  $\beta$  is performance pay that depends on output.<sup>7</sup> The principal now chooses only  $\alpha$  and  $\beta$  rather than the whole wage schedule.

### 4.1 Optimal contract with verifiable effort

We will first establish the optimal contract with verifiable effort in order to set a benchmark against which to compare the optimal contract with non-verifiable effort. The principal's expected wealth from a particular contract is given by

$$\begin{aligned} E[x - w] &= E[a + \varepsilon - \alpha - \beta(a + \varepsilon)] \\ &= (1 - \beta)a - \alpha. \end{aligned}$$

In order to compute the agent's expected utility from a contract, we need to make use of the fact that  $E[e^{k\varepsilon}] = e^{\frac{k^2\sigma^2}{2}}$  for any constant  $k$  when  $\varepsilon \sim N(0, \sigma^2)$ . Using this property,

distribution  $(n, p^L)$  when  $e = e_L$ , where  $p^H > p^L$ . Then

$$\frac{p_i^H}{p_i^L} = \frac{\binom{n}{x_i} (p^H)^{x_i} (1 - p^H)^{n-x_i}}{\binom{n}{x_i} (p^L)^{x_i} (1 - p^L)^{n-x_i}} = \left(\frac{p^H}{p^L}\right)^{x_i} \left(\frac{1 - p^H}{1 - p^L}\right)^{n-x_i},$$

which is increasing in  $x_i$ .

<sup>7</sup>Although linear contracts are optimal with verifiable effort and CARA utility, they are not optimal with non-verifiable effort. A contract that severely punishes the agent for low output realizations comes arbitrarily close to approximating the optimal contract with verifiable effort. We assume linear contracts only for simplicity.

we can write:

$$\begin{aligned}
E[u(w, a)] &= E \left[ -e^{-r(w - \frac{1}{2}a^2)} \right] \\
&= E \left[ -e^{-r(\alpha + \beta x - \frac{1}{2}a^2)} \right] \\
&= E \left[ -e^{-r(\alpha + \beta(a + \varepsilon) - \frac{1}{2}a^2)} \right] \\
&= -e^{-r(\alpha + \beta a - \frac{1}{2}a^2)} E[e^{-r\beta \varepsilon}] \\
&= -e^{-r(\alpha + \beta a - \frac{r\beta^2 \sigma^2}{2} - \frac{1}{2}a^2)}.
\end{aligned}$$

We will call the quantity

$$\alpha + \beta a - \frac{r\beta^2 \sigma^2}{2} - \frac{1}{2}a^2 \quad (17)$$

the agent's *certainty equivalent income*. The is because (17) is equal to the amount of money that, if received by the agent with certainty, gives him the same expected utility as the contract. It is equal to wealth (net of the cost of effort) minus a risk premium term  $\frac{r\beta^2 \sigma^2}{2}$ .

Using these observations, we can write the principal's optimization problem as

$$\max_{\alpha, \beta} (1 - \beta)a - \alpha \quad (18)$$

$$\text{such that } \alpha + \beta a - \frac{r\beta^2 \sigma^2}{2} - \frac{1}{2}a^2 \geq \bar{w} \quad (19)$$

Here we have written the participation constraint (19) in terms of certainty equivalent income.  $\bar{w}$  is the amount of income that satisfies  $-e^{-r\bar{w}} = \bar{u}$ . As long as the agent receives a certainty equivalent income of  $\bar{w}$  he will get his reservation utility  $\bar{u}$ .

Rather than explicitly write down the Lagrangian, we can simply appeal to economic theory to claim that  $\beta = 0$  in the optimal contract: since the agent is risk averse, he should be paid the same wage in all states of the world.  $\alpha$  is then set to make the participation constraint bind. To choose the optimal effort level, the principal maximizes<sup>8</sup>

$$a - \frac{1}{2}a^2 - \bar{w}$$

and so chooses effort level  $a = 1$ . This effort level equates the expected marginal product of effort with the marginal cost of effort.

## 4.2 Optimal contract with non-verifiable effort

When effort is non-verifiable, we need to add an incentive compatibility constraint to the principal's maximization problem:

$$\max_{\alpha, \beta} (1 - \beta)a - \alpha \quad (20)$$

$$\text{such that } \alpha + \beta a - \frac{r\beta^2 \sigma^2}{2} - \frac{1}{2}a^2 \geq \bar{w} \quad (21)$$

$$a = \arg \max_{\tilde{a}} \alpha + \beta \tilde{a} - \frac{r\beta^2 \sigma^2}{2} - \frac{1}{2}\tilde{a}^2. \quad (22)$$

---

<sup>8</sup>To derive this expression we have plugged  $\beta = 0$  into the objective function and participation constraint, and then plugged in for  $\alpha$  in the objective function from the binding participation constraint.

Again, we have written the incentive compatibility constraint in terms of certainty equivalent income since if the agent chooses effort to maximize his certainty equivalent income he will also be maximizing utility. Since (17) is concave in  $a$ , the first order condition for optimal effort is also sufficient, and the agent optimally chooses effort  $a^* = \beta$ . Thus, we can simply replace the incentive compatibility constraint (22) above with  $a = \beta$ .

The principal's problem thus becomes

$$\max_{\alpha, \beta} \beta(1 - \beta) - \alpha \quad (23)$$

$$\text{such that } \alpha + \beta^2 - \frac{r\beta^2\sigma^2}{2} - \frac{1}{2}\beta^2 \geq \bar{w}. \quad (24)$$

The Lagrangian for the above problem is

$$\beta(1 - \beta) - \alpha + \lambda \left( \alpha + \beta^2 - \frac{r\beta^2\sigma^2}{2} - \frac{1}{2}\beta^2 - \bar{w} \right)$$

with associated Kuhn-Tucker conditions

$$1 - 2\beta^* + \lambda^*\beta^*(1 - r\sigma^2) = 0 \quad (25)$$

$$-1 + \lambda^* = 0 \quad (26)$$

$$\alpha^* + (\beta^*)^2 - \frac{r(\beta^*)^2\sigma^2}{2} - \frac{1}{2}(\beta^*)^2 - \bar{w} \geq 0 \quad (27)$$

$$\lambda^* \geq 0 \quad (28)$$

$$\lambda^* \left[ \alpha^* + (\beta^*)^2 - \frac{r(\beta^*)^2\sigma^2}{2} - \frac{1}{2}(\beta^*)^2 - \bar{w} \right] = 0 \quad (29)$$

From (26) we obtain  $\lambda^* = 1$ , which implies the participation constraint binds. Since  $\lambda^* = 1$  we can use (25) to obtain

$$\beta^* = \frac{1}{1 + r\sigma^2}. \quad (30)$$

$\alpha^*$  is then chosen to make the agent's participation constraint bind.

### 4.3 Properties of the optimal contract

There are several interesting properties of the optimal contract:

1. There is inefficient risk sharing since  $\beta^* > 0$ . The introduction of incentive pay to deal with the moral hazard problem exposes the agent to wage uncertainty.
2. There is inefficient effort since  $\beta^* < 1$ . The optimal contract does not provide enough incentives for the agent to exert efficient effort. Thus, the optimal contract has both inefficient risk sharing and inefficient effort.
3.  $\beta^*$  is decreasing in  $r$ . As the agent becomes more risk averse, it becomes more expensive for the principal to provide incentives, so pay becomes less sensitive to output. Also, when  $r = 0$  so that the agent is risk neutral,  $\beta^* = 1$ , and we are back to the "sell the firm" contract in which the agent keeps all the output in exchange for paying the principal a fixed payment.



4.  $\beta^*$  is decreasing in  $\sigma^2$ . As output becomes more variable, incentive pay becomes more risky for the agent since the relationship between effort and output becomes more uncertain. The principal responds by making pay less sensitive to output. The general message is that tasks that are hard to measure should feature less incentive pay.

## 5 The Value of Information

We have argued that the non-verifiability of effort leads to an inefficient outcome, and that if the principal could contract on effort he would make himself better off while making the agent no worse off. One might wonder, then, whether the principal should include all available information about effort in the optimal contract. Suppose there are two output levels  $x_1$  and  $x_2$ , two effort levels  $e_H$  and  $e_L$ , and two signals of effort  $y_1$  and  $y_2$  that are *independent* of output. For example, if we think about a lawyer working in an office, the number and quality of legal cases she writes is her output and this is obviously informative about her effort. On the other hand, the number of trips she makes each day to the office café is also informative about her effort, but not necessarily correlated with output. While hard working lawyers may take less trips to the café, going to the café in itself does not have to be correlated with poor performance (once we fix an effort level) since the lawyer could be working on her cases there.

There are now four states of the world  $(x_1, y_1)$ ,  $(x_1, y_2)$ ,  $(x_2, y_1)$ , and  $(x_2, y_2)$ . The question is: should the agent's wage depend on the realization of  $y$  as well as the realization of  $x$ ? Let  $w_{ij}$  be the agent's wage when  $x = x_i$  and  $y = y_j$ . We have seen that we can write the optimal contract in this setting (see equation (16) above), as

$$\frac{1}{u'(w_{ij})} = \lambda + \mu \left[ 1 - \frac{\Pr[x = x_i, y = y_j \mid e = e_L]}{\Pr[x = x_i, y = y_j \mid e = e_H]} \right]$$

which, by the independence of  $x$  and  $y$  can be written as

$$\frac{1}{u'(w_{ij})} = \lambda + \mu \left[ 1 - \frac{\Pr[x = x_i \mid e = e_L] \Pr[y = y_j \mid e = e_L]}{\Pr[x = x_i \mid e = e_H] \Pr[y = y_j \mid e = e_H]} \right].$$

So, the optimal contract should depend on  $y$  whenever

$$\Pr[y = y_j \mid e = e_L] \neq \Pr[y = y_j \mid e = e_H]$$

that is, whenever  $y$  provides information about the agent's effort level.

This result is actually true much more generally, and is known as the Informativeness Principle. It is one of the most important insights in contract theory.

**Result 2** *If there is a variable that contains information about the agent's effort that is not already contained in the agent's output, then the optimal contract should include this variable.*

The Informativeness Principle does not say that the contract should contain all information that the principal sees. If in our example above,  $y$  had been perfectly correlated with  $x$ , then there would be no need to include  $y$  in the contract. Going back to our lawyer

example, if  $x$  is the number of pages of legal briefings written and  $y$  is the number of pages the lawyer has printed out from her computer, there is no need to include  $y$  in the contract since it contains almost no additional information about effort given that the principal already observes  $x$ .

## 6 Is Incentive Pay Always Good?

The Informativeness Principle might lead one to believe that contracting on all observable variables cannot harm the principal. This conclusion is in fact misleading. We will now explore the idea that, in some situations, contracting on observable variables may harm the principal because the agent will divert effort away from valuable but hard to measure outputs. Economists refer to this a situation of *multitasking*.

We will adapt the CARA utility, normal output setup to model the multitasking problem. Suppose there are two different outputs—or tasks—to which the agent can devote effort:  $x_1 = a_1 + \varepsilon_1$  where  $a_1$  is the agent's effort on the first task and  $\varepsilon_1 \sim N(0, \sigma_1^2)$  is a random shock to  $x_1$ ; and  $x_2 = a_2 + \varepsilon_2$  where  $a_2$  is the agent's effort on the second task and  $\varepsilon_2 \sim N(0, \sigma_2^2)$  is a random shock to  $x_2$ .  $\sigma_1^2$  measures the difficulty of measuring task 1 while  $\sigma_2^2$  measures the difficulty of measuring task 2. The principal is risk neutral and offers contracts of the form  $w = \alpha + \beta_1 x_1 + \beta_2 x_2$ .

The agent has a cost of effort function  $c(a_1, a_2) = \frac{a_1^2}{2} + \frac{a_2^2}{2} + \delta a_1 a_2$  where  $0 < \delta < 1$ . Thus the marginal cost of effort on task 1 is  $a_1 + \delta a_2$  and the marginal cost of effort on task 2 is  $a_2 + \delta a_1$ . What is important is that the marginal cost of effort on each task is increasing in the amount of effort that the agent exerts on the other task. This means the agent will tend to substitute effort on the task that is rewarded less with effort on the task that is rewarded more. To explore this idea more formally, recall that when the agent has a CARA utility function we can express his utility from a contract in terms of certainty equivalent income, which in this case is  $\alpha + \beta_1 a_1 - \frac{r\beta_1^2\sigma_1^2}{2} + \beta_2 a_2 - \frac{r\beta_2^2\sigma_2^2}{2} - \frac{a_1^2}{2} - \frac{a_2^2}{2} - \delta a_1 a_2$ . The equations that define the agent's optimal effort levels are

$$\begin{aligned}\beta_1 - a_1^* - \delta a_2^* &= 0 \\ \beta_2 - a_2^* - \delta a_1^* &= 0\end{aligned}$$

from which we obtain

$$\begin{aligned}a_1^* &= \frac{\beta_1 - \delta\beta_2}{1 - \delta^2} \\ a_2^* &= \frac{\beta_2 - \delta\beta_1}{1 - \delta^2}\end{aligned}$$

Here one can see that there are two ways to motivate each task. The principal can raise the incentive pay associated with the task, or the principal can reduce the incentive pay associated with the other task, thereby decreasing the opportunity cost of working on the task.

We will skip the mathematical details and simply state that the optimal incentive

payments are given by

$$\beta_1^* = \frac{1 - (1 - \delta)r\sigma_2^2}{1 + r\sigma_1^2 + r\sigma_2^2 + r^2(1 - \delta^2)\sigma_1^2\sigma_2^2} \quad (31)$$

$$\beta_2^* = \frac{1 - (1 - \delta)r\sigma_1^2}{1 + r\sigma_1^2 + r\sigma_2^2 + r^2(1 - \delta^2)\sigma_1^2\sigma_2^2} \quad (32)$$

When  $\delta = 0$  so that the marginal cost of effort on each task is independent of the effort the agent exerts on the other task, these expressions become

$$\beta_1^* = \frac{1}{1 + r\sigma_1^2} \quad (33)$$

$$\beta_2^* = \frac{1}{1 + r\sigma_2^2} \quad (34)$$

which are simply the optimal incentive payments for each task treated separately. Thus, the interaction between the two efforts in the cost of effort function is the reason that multitasking is relevant in this model.<sup>9</sup> One can also show that both (31) and (32) are decreasing in  $\delta$ : as the agent's incentive to substitute one effort for another become more pronounced, the principal provides less incentive pay.

Even more interesting is the relationship between incentive pay and the ease with which tasks can be measured. (31) and (32) are decreasing in both  $\sigma_1^2$  and  $\sigma_2^2$ . We have already discussed why incentive pay for a task should decrease with the difficulty of measuring it, but what is less clear is why it should also decrease with the difficulty of measuring the other task. The reason is that if task 2, for example, becomes harder to measure, the principal will reduce the direct incentive pay for task 2 in order to expose the agent to less risk, but then also lower the incentive pay for task 1 in order to ensure that the agent still exerts effort on task 2. Interestingly, as  $\delta \rightarrow 1$  and  $\sigma_2^2 \rightarrow \infty$ ,  $\beta_1^* \rightarrow 0$ . If effort substitutability is high and task 2 is extremely difficult to measure, then the principal should provide no incentive pay at all on task 1 or task 2.

While multi-tasking is one reason why contracting on all observable information might harm the principal, there are others as well. Several empirical studies have argued that incentive pay can replace altruism as a guide of human behavior. In the 1960's Richard Titmuss documented the response of the British public to the introduction of payments for donating blood. It turned out that after payments for blood were introduced, the quality of blood donors decreased substantially. One hypothesis is that the people donating blood became those that needed money, not those that were acting out of a need to help society. More recent studies have found similar results: in Israel the introduction of fines for parents that were late for picking up their children from school *increased* the portion of parents that were late, and the introduction of compensation payment to residents of Swiss districts in which nuclear power plants were to be located increased opposition to them.

---

<sup>9</sup>Another situation in which multitasking concerns are relevant is when the outputs on both tasks are correlated.

## 7 Moral Hazard and Limited Liability

So far we have stressed the trade-off between risk and incentives as the cause of the moral hazard problem. We conclude our discussion by showing that non-verifiable effort can also lead to contractual inefficiency when the agent has limited liability. Limited liability arises when there is a limit to the amount of wealth that the principal can extract from the agent. This may be because the agent does not have sufficient wealth to pay the principal, or because he is legally protected in the event of project failure (as is the case with corporations).

Suppose that the principal and the agent are both risk neutral. This assumption eliminates the trade-off between risk and incentives we have emphasized up until now. There are two output levels  $x_1 = 0$  and  $x_2 = 1$ .  $e$  is the agent's effort level, and  $\Pr[x = x_2 | e] = e$ . The agent's disutility of effort is  $\frac{e^2}{2}$  and his reservation utility is  $\bar{u}$ . Without limited liability, the optimal contract for the principal is to sell the project to the agent. The agent will choose effort to maximize surplus ( $e = 1$ ), and the principal will extract all the resulting surplus (net of the reservation utility  $\bar{u}$ ) in the form of the sale price.

We now ask how the optimal contract changes when the agent must have non-negative wealth in both states of the world. Let  $w$  be the wage that the principal pays when  $x = x_1$  and let  $w + b$  be the wage he pays when  $x = x_2$ . The principal's expected utility from the contract is  $e(1 - b) - w$  and the agent's is  $eb + w - \frac{e^2}{2}$ . Thus, the agent will choose effort level  $e = b$ , which means the principal's expected utility becomes  $b(1 - b) - w$  and the agent's becomes  $\frac{b^2}{2} + w$ . Thus, the principal solves

$$\begin{aligned} & \max_{b,w} b - b^2 - w \\ \text{such that } & \frac{b^2}{2} + w \geq \bar{u} \\ & w \geq 0 \\ & w + b \geq 0. \end{aligned}$$

Rather than solve the full Lagrangian, we appeal to economic arguments. Here, the principal's goal is to extract as much wealth from the agent as possible while preserving work incentives. Since  $w$  is simply a wealth transfer from the principal to the agent and has no role to play in incentive provision, the optimal contract sets  $w = 0$ .<sup>10</sup> This makes the maximization problem become

$$\begin{aligned} & \max_{b,w} b - b^2 \\ \text{such that } & \frac{b^2}{2} \geq \bar{u}. \end{aligned}$$

$b = \frac{1}{2}$  maximizes the principal's objective function, so as long as  $\bar{u} \leq \frac{1}{8}$  this is the optimal contract. Otherwise, the principal sets  $b = \sqrt{2\bar{u}}$ . As long as  $\bar{u} \leq \frac{1}{2}$ , the optimal contract is inefficient since the agent exerts too low an effort. The case in which  $\bar{u} > \frac{1}{2}$  is not pertinent, since in this case the principal is not able to make a profit even in the verifiable effort case, since the maximum value that the surplus can attain is  $\frac{1}{2}$

---

<sup>10</sup> $w$  cannot be set any lower because of the agent's wealth constraint.