Lecture 7
Asymmetric Information: The Principal-Agent Problem

Information asymmetries exist in a game when a player does not know some relevant fact about another player. Games of this nature are called incomplete information games. In a principal-agent game, one player (the principal) wants to induce some action by another player (the agent). For example, a manager may want to induce a certain level of effort from an employee. It does so by choosing an appropriate compensation scheme for the worker. The choice of a compensation scheme is complicated when the manager can’t observe the worker’s level of effort or the worker’s skill-level. Information is asymmetric in these situations because the worker knows his level of effort and his ability, but the manager does not.

Regulatory problems are principal-agent problems in which a regulatory authority is the principal and regulated firms are the agents. The regulator may wish to choose an efficient regulation, but is hindered in this if it cannot observe the cost structure of a regulated firm. Moreover, the regulator may not be able to perfectly observe the actions of a firm to check if it is complying with a regulation.

There are two basic types of information problems in these games. If the principal cannot observe the behavior of the agent, the information problem is one of hidden actions. If the principal cannot observe some relevant characteristic of the agent (e.g., his skill-level, or a firm’s costs), the problem is one of hidden information. We will examine both problems in the manager-worker example.
**Optimal Manager/Worker Contracts**

This entire lecture is focused on a contractual relationship between a manager and a worker. In this section we begin by investigating the design of optimal contracts when information is complete; that is, the manager can observe the worker’s effort level and knows his utility function. Then we will examine how optimal contracts will be different if the relationship between the worker’s level of effort and the output he produces is stochastic. Throughout this section information is symmetric.

Suppose that a worker’s effort produces an output \( x \geq 0 \) for his manager. For his efforts the worker receives compensation that may depend on the output produced, \( s(x) \). Let the manager’s payoff (utility) function be

\[
 u_m = x - s(x). \tag{1}
\]

Note that since the manager’s utility is linear in a monetary payoff, the manager is risk-neutral. Suppose that the manager seeks to choose a compensation scheme for the worker to maximize her utility.

The worker’s effort is \( e \geq 0 \), and his disutility from effort is given by the cost function \( c(e) \). For now, let the worker’s utility be

\[
 u_w = s(x) - c(e). \tag{2}
\]

**Optimal Contract Under Complete Information**

Suppose that the manager can observe the worker’s effort and that output is completely determined by effort so that \( x = x(e) \). Since effort is observable, its level can be made part of the contract. Therefore, the manager chooses \( e \) and \( s(x(e)) \) to maximize

\[
 x(e) - s(x(e)). \tag{3}
\]

To make sure that the worker accepts a contract and expends the specified amount of effort, the contract \([e, s(x(e))]) must satisfy two constraints.

**Participation** Assume that the worker has other opportunities. To get the worker to accept the contract, it has to be chosen so that the worker achieves the level of utility he could get in his next best alternative. Let \( \bar{u} \) be the worker’s utility in his next best alternative. This is often referred to as the worker’s reservation level of utility. To make sure the worker accepts the contract offered by the manager, the contract \([e, s(x(e))] must satisfy

\[
 u_w = s(x(e)) - c(e) \geq \bar{u}. \tag{4}
\]

**Incentive compatibility** The manager must also make sure that the compensation \( s(x) \) induces the worker to choose the contracted level of \( e \) instead of some other (presumably lower) level of effort. In other words, the worker must find it in his best interest to work the specified level of effort. Thus, the contracted level of effort must satisfy
Asymmetric Information

\[ e \text{ maximizes } u_w = s(x(e)) - c(e). \]

The following is an optimal contract in this setting:

| Compensation: \[ s(x(e)) = \begin{cases} \bar{u} + c(e) & \text{if } e = e^* \\ 0 & \text{if } e \neq e^* \end{cases} \] |
| Effort target: \[ e^* = e \mid x'(e) = c'(e). \] |

To see why this works note first that \( s(x(e)) \) satisfies the participation constraint (4). In fact it is chosen to leave the worker with only his reservation level of utility. This is necessary because maximization of the manager’s utility (3) given some level of effort requires \( s(x(e)) \) to be as low as possible. The incentive-compatibility constraint is also satisfied because if the worker chooses \( e = e^* \), his utility is

\[ s(x(e^*)) - c(e^*) = \bar{u} + c(e^*) - c(e^*) = \bar{u} > 0. \]

If the worker chooses \( e \neq e^* \) instead, his utility is \(-c(e) \leq 0\). Therefore, \( e^* \) is the optimal choice of effort by the worker given the structure of the contract. Lastly, the worker’s compensation, \( s(x(e)) = \bar{u} + c(e) \), leaves the manager with utility

\[ x(e) - s(x(e)) = x(e) - c(e) + \bar{u}. \]

Note that \( \bar{u} \) is a constant. Then, given standard assumptions about \( x(e) \) and \( c(e) \) \([x(e) \text{ is increasing and concave, and } c(e) \text{ is increasing and convex}]\), the optimal level of effort is \( e^* \), the solution to \( x'(e) = c'(e) \).

Another possible contract that doesn’t require that an effort level be specified would make the worker a residual claimant of all output produced. Consider the compensation scheme

\[ s(x(e)) = x(e) - F, \]

Note that the worker receives all the output, but must pay the manager a fixed fee \( F \). Given this compensation, the worker chooses effort to maximize his utility

\[ u_w = s(x(e)) - c(e) = x(e) - c(e) - F; \]

hence, the worker chooses \( e^* = e \mid x'(e) = c'(e) \). Furthermore, the manager’s utility is

\[ u_m = x(e) - s(x(e)) = x(e) - x(e) + F = F. \]

To maximize her utility, she should choose \( F \) to be as high as possible. She does so by making sure that the worker ends up with no more than his reservation utility. Given that the worker chooses \( e^* \), the optimal \( F \) is determined by making sure that the participation constraint (4) is just binding; that is, \( F \) satisfies
which implies \( F = x(e^*) - c(e^*) - \bar{u} \).

Both contracts (and there are other possibilities) achieve the same outcome. The worker accepts the contract and chooses \( e^* \) effort. He receives compensation for his effort that leaves him no better off (or worse off) than his next best alternative. The manager captures the entire surplus generated by their relationship (because she is a monopsony buyer of labor). Furthermore, the contract is efficient because there is no way to make one of them better off without harming the other. Alternatively, the level of effort \( e^* \) maximizes their joint welfare:

\[
\begin{align*}
   u_m + u_w &= [x(e) - s(x(e))] + [s(x(e)) - c(e)] \\
   &= x(e) - c(e).
\end{align*}
\]

\( e^* \) is often referred to as the first-best efficient level of effort. “First-best” refers to situations involving no uncertainty. When there are information problems, the optimal contract must take these into account, and hence, will be very different from what we’ve described thus far.

**Optimal Contract Under Imperfect but Symmetric Information**

In this section we derive the optimal contract when there is uncertainty about how the worker’s effort affects output. Both the manager and the worker know the nature of this uncertainty. In addition, the worker’s effort is observable as are his relevant characteristics.

Suppose that output can take on a finite number of values \((x_1, x_2, \ldots, x_n)\), but suppose that output is a random variable that depends, in part, on the worker’s effort \( e \). Let \( \pi(x_i | e) \) be the probability density function conditional on the worker’s effort. That is, a particular value of \( \pi(x_i | e) \) gives us the probability that output \( x_i \) occurs when the worker applies effort \( e \). Note that we must have:

\[
\begin{align*}
(i) & \quad \pi(x_i | e) \in [0,1], \ \forall x_i \in (x_1, \ldots, x_n) \text{ and } \forall e; \\
(ii) & \quad \sum_{i=1}^n \pi(x_i | e) = 1, \ \forall e.
\end{align*}
\]

Suppose that a contract specifies a level of effort and a compensation that depends on final output, \( s(x_i) \). Expected utility for the manager is

\[
E(u_m) = \sum_{i=1}^n \pi(x_i | e) \times (x_i - s(x_i)).
\]

The manager wishes to choose a contract \((e, s(x_i))\) to maximize (6).

Let the worker’s utility be

\[
s(x(e^*)) - c(e^*) = x(e^*) - F - c(e^*) = \bar{u},
\]
Asymmetric Information

\[ u_w = u(s(x_i)) - c(e), \]

where \( u' (\cdot) > 0 \) and \( u'' (\cdot) < 0 \). Since the worker’s utility is strictly concave in a random monetary payoff, he is strictly risk-averse. The worker’s expected utility is

\[ E(u_w) = \sum_{i=1}^{n} \pi(x_i \mid e) \times u(s(x_i)) - c(e). \]  \( \text{(7)} \)

If the worker is going to accept the contract, his expected payoff must be greater than or equal to his reservation utility.

The contract that the manager should offer the worker is the solution to

\[
\begin{align*}
\max_{(e,s(x_i))} \sum_{i=1}^{n} \pi(x_i \mid e) \times (x_i - s(x_i)) \\
\text{s.t.} \quad \sum_{i=1}^{n} \pi(x_i \mid e) \times u(s(x_i)) - c(e) \geq \bar{u}.
\end{align*}
\]  \( \text{(8)} \)

In the problem we looked at in the previous section an incentive-compatibility constraint was not essential. The reason is that, since effort is observable, inducing the optimal level of effort is easily accomplished by making sure that the compensation scheme includes a harsh penalty (perhaps zero compensation) for deviations from the optimal effort. Since effort is observable in this problem as well, we can solve the problem without an incentive-compatibility constraint.

It will be convenient for us to solve \( \text{(8)} \) in two steps. First we will derive optimal compensation for any fixed level of effort. Then, given optimal compensation, we will solve for the optimal level of effort.

To begin the first step, first note that when optimizing \( \text{(8)} \) over \( s(x_i) \), given \( e \), the problem simplifies somewhat to

\[
\begin{align*}
\min_{s(x_i)} \sum_{i=1}^{n} \pi(x_i \mid e) \times s(x_i) \\
\text{s.t.} \quad \sum_{i=1}^{n} \pi(x_i \mid e) \times u(s(x_i)) - c(e) \geq \bar{u}.
\end{align*}
\]  \( \text{(9)} \)

Thus, given that the contract specifies a certain level of effort, the manager chooses \( s(x_i) \) to minimize her expected payment subject to the participation constraint.

Next note that the participation constraint must bind – the worker’s expected utility must be exactly equal to his reservation utility. If the constraint did not bind, the manager could induce the same level of effort with lower compensation.

The Lagrange equation for \( \text{(9)} \) is then

\[
L = \sum_{i=1}^{n} \pi(x_i \mid e) \times s(x_i) + \lambda \left( \bar{u} - \sum_{i=1}^{n} \pi(x_i \mid e) \times u(s(x_i)) + c(e) \right).
\]
The first-order condition for the optimal choice of $s(x_i)$ is $\pi(x_i | e) - \lambda \pi(x_i | e) u'(s(x_i)) = 0$, or equivalently,

$$\frac{1}{u'(s(x_i))} = \lambda. \tag{10}$$

Recall that risk aversion implies that $u(s(x_i))$ is strictly concave, which in turn implies that $u'(s(x_i))$ is strictly decreasing in $s(x_i)$. Therefore, since $\lambda$ is a constant, $s(x_i)$ must be constant, which implies that the optimal compensation $s(x_i)$ should be independent of $x_i$. That is, the manager offers the worker a fixed payment that is independent of random fluctuations in output. With this compensation scheme the worker is fully insured against risk: all of the risk is borne by the manager.

Now given that the contract specifies an effort level, the optimal compensation is a fixed payment $s^*$ that leaves the worker with his reservation utility; that is, $s^*$ is the solution to

$$u_w = u(s^*) - c(e) = \bar{u}. \tag{11}$$

Let $f$ be the inverse of $u$. Then,

$$s^* = f(\bar{u} + c(e)). \tag{12}$$

All that is left now is to determine the contract’s specification of effort. From 6) the manager’s expected utility is

$$\sum_{i=1}^{n} \pi(x_i | e) x_i (x_i - s^*) = \sum_{i=1}^{n} \pi(x_i | e) x_i - \sum_{i=1}^{n} \pi(x_i | e) s^*. \tag{13}$$

Since, $\sum_{i=1}^{n} \pi(x_i | e) = 1$ and $s^*$ is a constant, the manager’s expected utility is

$$\sum_{i=1}^{n} \pi(x_i | e) x_i - s^*. \tag{13}$$

Substitute from 12) to obtain

$$\sum_{i=1}^{n} \pi(x_i | e) x_i - f(\bar{u} + c(e)). \tag{13}$$

The optimal contract specifies the level of effort $e$ that maximizes 13).
Asymmetric Information: Hidden Action (Effort)

In the last section the contract between the manager and the worker specified the effort that the worker needed to expend to fulfill his part of the contract. When effort is not observable, the contract can specify only the compensation. The trick for the manager is to choose a compensation to induce the worker to expend the level of effort she wants him to.

Let us simplify the problem some by assuming that the worker’s effort can take on only two values, high and low. Denote these effort levels by \( e_h \) and \( e_l \), respectively. As for the worker’s disutility of effort, let \( c_h = c(e_h) \) and \( c_l = c(e_l) \). Let us make the natural assumption that, given a compensation payment, the worker prefers to expend low effort. Thus, \( c_h > c_l \).

Unlike the previous problems, there is a serious incentive compatibility difficulty when the worker’s effort is not observable. If the manager wants the worker to expend a high level of effort and does not design the compensation scheme appropriately, the worker may simply choose a low level of effort and the manager would never know. So in this problem we must include an incentive compatibility constraint.

The manager’s problem is to choose an effort level \( e \in (e_l, e_h) \) that it wants to induce and a compensation scheme \( s(x_i) \) to maximize

\[
\sum_{i=1}^{n} \pi(x_i | e) \times (x_i - s(x_i)), \quad 14a)
\]

subject to

\[
\sum_{i=1}^{n} \pi(x_i | e) \times u(s(x_i)) - c(e) \geq \bar{u}. \quad 14b)
\]

\( e \) maximizes

\[
\sum_{i=1}^{n} \pi(x_i | e) \times u(s(x_i)) - c(e). \quad 14c)
\]

14a) is the manager’s expected utility, 14b) is the worker’s participation constraint, and 14c) is the incentive compatibility constraint. Instead of solving this problem all at once, we will determine the optimal compensation for inducing each level of effort separately.

**Inducing \( e_l \):** This is easy. The optimal compensation is exactly the one that would be chosen if effort was observable; namely, from 12), \( s^* = f(\bar{u} + c_l) \). To see why this is the optimal payment for low effort note first that since the worker’s payment is independent of the outcome \( x \), and hence independent of his effort, he will minimize his disutility from working and choose \( e_l \). Thus, the incentive compatibility constraint 14c) is satisfied. The participation constraint 14b) is also satisfied since \( s^* \) leaves the worker with his reservation utility. Lastly, since \( s^* \) is the lowest payment required to induce the worker’s participation and a low level of effort, the manager’s expected utility, given \( e_l \), is maximized subject to the constraints.

**Inducing \( e_h \):** This is more interesting. In this case the incentive compatibility constraint 14c) can be rewritten as

\[
\sum_{i=1}^{n} \pi(x_i | e_h) \times u(s(x_i)) - c_h \geq \sum_{i=1}^{n} \pi(x_i | e_l) \times u(s(x_i)) - c_l,
\]
or equivalently,

\[ \sum_{i=1}^{n} u(s(x_i)) \left[ \pi(x_i \mid e_h) - \pi(x_i \mid e_l) \right] - c_h + c_l \geq 0. \]

14c') forces the manager to choose \( s(x_i) \) so that the worker finds it optimal to expend \( e_h \) rather than \( e_l \).

From 14a-b), the optimal compensation to induce \( e_h \) solves

\[ \max_{s(x_i)} \sum_{i=1}^{n} \pi(x_i \mid e_h) \times (x_i - s(x_i)) \]

subject to

\[ \sum_{i=1}^{n} \pi(x_i \mid e_h) \times u(s(x_i)) - c_h \geq \bar{u} \]

and 14c'). The Lagrange equation is

\[ L = \sum_{i=1}^{n} \pi(x_i \mid e_h) \times (x_i - s(x_i)) + \lambda \left( \sum_{i=1}^{n} \pi(x_i \mid e_h) \times u(s(x_i)) - c_h - \bar{u} \right) \]

\[ + \mu \left( \sum_{i=1}^{n} u(s(x_i)) \left[ \pi(x_i \mid e_h) - \pi(x_i \mid e_l) \right] - c_h + c_l \right), \]

with first-order condition

\[ -\pi(x_i \mid e_h) + \lambda \pi(x_i \mid e_h) \times u'(s(x_i)) + \mu u'(s(x_i)) \left[ \pi(x_i \mid e_h) - \pi(x_i \mid e_l) \right] = 0. \]

Rearrange terms to obtain

\[ \frac{1}{u'(s(x_i))} = \lambda + \mu \left[ 1 - \frac{\pi(x_i \mid e_l)}{\pi(x_i \mid e_h)} \right]. \]

15)

We shall compare 15) to the marginal condition that determines the optimal compensation when effort is observable,

\[ \frac{1}{u'(s(x_i))} = \lambda. \]

10) Under reasonable circumstances, it can be shown that the participation and incentive-compatibility constraints are binding. Therefore, let us assume that the multipliers in 15) are strictly positive. In comparing 10) and 15), the obvious difference is that the compensation to induce \( e_h \) when it cannot be observed is not independent of output as it would be in 10).

Furthermore, the optimal compensation with unobservable effort is a random variable (because \( x_i \) is a random variable). This implies that when effort is not observable, the compensation scheme needed to induce \( e_h \) depends on the realization of output, and hence, forces the worker to bear
some risk. Recall that when effort is observable, the manager offers a fixed compensation and bears all of the risk.

Consider the ratio $\pi(x_i | e_l)/\pi(x_i | e_h)$ in 15). Recall from basic statistical theory that this is a likelihood ratio, which gives us the likelihood of observing $x_i$ when effort $e_l$ is applied relative to the likelihood of observing $x_i$ when $e_h$ is applied.

Now as before, let $s^*$ be the solution to 10), and let $\hat{s}(x_i)$ be the solution to 15). Then:

$$
\hat{s}(x_i) > s^* \text{ if } \pi(x_i | e_l) < \pi(x_i | e_h) ;
$$

$$
\hat{s}(x_i) < s^* \text{ if } \pi(x_i | e_l) > \pi(x_i | e_h).
$$

This implies that relative to the fixed payment when effort is observable, the payment to induce $e_h$ when it is not observable is higher (lower) when the likelihood of observing a particular $x_i$ when low effort is applied is less than (greater than) the likelihood of observing $x_i$ when $e_h$ is applied.

You might expect that the optimal compensation scheme given by 15) is increasing in the observed output, but this need not be the case. To see why, suppose that $s(x_i)$ is increasing in $x_i$.

Then since $u'' < 0$, $u'(s(x_i))$ is decreasing in $x_i$ and the left-hand-side of 15) is increasing in $x_i$.

Therefore, if $s(x_i)$ is to be increasing in $x_i$, the right-hand-side of 15) must also be increasing in $x_i$.

This requires the ratio $\pi(x_i | e_l)/\pi(x_i | e_h)$ to be decreasing in $x_i$. If it is not, then the optimal payment is not increasing in output.

Often it is simply assumed, or distribution functions are chosen so that is $\pi(x_i | e_l)/\pi(x_i | e_h)$ is decreasing in $x_i$. This property is often called the \textit{monotone likelihood ratio property}.

Although for some realizations of output the compensation to induce $e_h$ when it is unobservable may be less than the required payment when $e_h$ is unobservable, it turns out that the worker’s expected payment when $e_h$ is not observable is greater than his payment when his effort is observable; that is

$$
\sum_{i=1}^{n} \pi(x_i | e_h) \times \hat{s}(x_i) > s^* ,
$$

where $s^*$ induces $e_h$ when $e_h$ is observable. To establish this result we need the following:

\textbf{Jensen’s Inequality}

For a discreet random variable $z$ that is distributed according to $f(z)$, consider a strictly concave (convex) function $h(z)$. Then, the value of $h$ evaluated at the expectation of $z$ is greater than (less than) the expectation of $h(z)$. That is,

$$
h\left(\sum_i z_i f(z_i)\right) > (<) \sum_i h(z_i)f(z_i).
$$
To establish (16) using Jensen’s inequality, recall from 11) that $s^*$ satisfies $u(s^*) - c_h = \bar{u}$ or rather

$$u(s^*) = \bar{u} + c_h.$$ \hfill (17)

The worker also just achieves his reservation utility in the unobservable effort case. From the participation constraint 14b),

$$\sum_{i=1}^{n} \pi(x_i \mid e) \times u(\hat{s}(x_i)) = \bar{u} + c_h.$$ \hfill (18)

17) and 18) imply

$$\sum_{i=1}^{n} \pi(x_i \mid e) \times u(\hat{s}(x_i)) = u(s^*).$$ \hfill (19)

Since $u(\cdot)$ is strictly concave, Jensen’s Inequality implies

$$u\left(\sum_{i=1}^{n} \pi(x_i \mid e_h) \times \hat{s}(x_i)\right) > \sum_{i=1}^{n} \pi(x_i \mid e_h) \times u(\hat{s}(x_i)).$$ \hfill (20)

19) and 20) yield

$$u\left(\sum_{i=1}^{n} \pi(x_i \mid e_h) \times \hat{s}(x_i)\right) > u(s^*).$$ \hfill (21)

Since $u(\cdot)$ is monotonically increasing, (21) holds if and only if

$$\sum_{i=1}^{n} \pi(x_i \mid e_h) \times \hat{s}(x_i) > s^*,$$

which is 16), the result we set out to prove.

The fact that the worker’s expected compensation for high effort when it is not observable is greater than when effort is observable is very intuitive. When high effort is not observable, the worker bears some risk. To get the worker to bear risk the manager must compensate him for doing so. This extra compensation is not necessary when effort is observable.

There is a welfare consequence of the fact that the expected compensation for high effort when it unobservable is greater than the compensation when it is observable. Since the manager’s expected cost of inducing $e_h$ is greater when it is unobservable, she may choose to implement $e_l$ instead even though $e_h$ would be preferred if it were observable.

Now that we have determined the optimal payments to induce $e_l$ and $e_h$ when they are not observable, all that is left is for the manager to choose which effort level to induce. Using 14a), her expected payoff from inducing $e_l$ is

$$\sum_{i=1}^{n} \pi(x_i \mid e_l) \times x_i - s^*,$$ \hfill (22)
where \( s^* = f(\bar{\mu} + c(e)) \) from 12). Her expected payoff from inducing \( e_h \) is

\[
\sum_{i=1}^n \pi(x_i | e_h) \times (x_i - \hat{s}(x_i)).
\]

The manager should offer \( s^* \) to induce \( e_l \) if 22) is greater than 23). She should offer \( \hat{s}(x_i) \) to induce \( e_h \) if 23) is greater than 22).

**The Hidden-Action Model with Mean-Variance Utility**

The model of the previous section does not lend itself readily to comparative static analysis. In this section we construct a model of contracting when the agent’s effort is not observable that allows us to examine how uncertainty about the production (effort-output) relationship and the agent’s attitudes toward risk affect the principal’s choice of contract.

Suppose that the manager offers the agent compensation that is a linear function of realized output \( \hat{x} \):

\[
s(\hat{x}) = \alpha + \beta \hat{x}.
\]

Note that \( \alpha \) is a fixed payment to the worker while \( \beta \in [0,1] \) is his share of output. Continue to assume that there is symmetric uncertainty about how the worker’s effort affects output; specifically,

\[
\hat{x} = x(e) + \varepsilon,
\]

where \( \varepsilon \) is an error term with zero mean and variance \( \sigma^2 \).

Suppose that the agent has a mean-variance utility function

\[
U = E(u_w) - r \text{ var}(u_w),
\]

where \( u_w = s - c(e) \) is his compensation net of his personal costs of effort, and \( r \) is his degree of risk aversion. (See Silberberg and Suen, pp 406-408, for a discussion of mean-variance utility functions). Using 24) and 25),

\[
u_w = s(\hat{x}) - c(e)
= \alpha + \beta(x(e) + \varepsilon) - c(e)
= \alpha + \beta x(e) - c(e) - \beta \varepsilon,
\]

with

\[
E(u_w) = \alpha + \beta x(e) - c(e),
\]

and

\[
\text{var}(u_w) = \beta^2 \sigma^2.
\]
26) - 28) imply that the agent’s utility is

\[ U = \alpha + \beta x(e) - c(e) - \epsilon \beta^2 \sigma^2. \quad 29) \]

To design an optimal compensation scheme, the manager has to know how the worker will respond to any contract she might offer. Assuming that 29) is strictly concave \( [\beta x''(e) - c''(e) < 0] \), the first order condition for maximizing 29) is

\[ \beta x'(e) - c'(e) = 0, \quad 30) \]

provided that the worker accepts the contract (the manager will use \( \alpha \) to make sure that he does).

30) implicitly defines the agent’s choice of effort \( e = e^*(\beta) \). Note that the first-best efficient level of effort is the solution to \( x'(e) - c'(e) = 0 \). Therefore, the manager can induce the first-best level of effort only if she chooses \( \beta = 1 \). However, this would make the agent the sole residual claimant of output and would, consequently, make him bear all of the production risk.

Now substitute \( e^*(\beta) \) into 30):

\[ \beta x'(e^*(\beta)) - c'(e^*(\beta)) = 0. \quad 31) \]

Differentiate 31) with respect to \( \beta \) to obtain

\[ x' + \beta x^* \left( \frac{de^*(\beta)}{d \beta} \right) - c^* \left( \frac{de^*(\beta)}{d \beta} \right) = 0. \]

Rearrange this to obtain

\[ \frac{de^*(\beta)}{d \beta} = \frac{-x'}{\beta x'^* - c'^*} > 0. \quad 32) \]

The sign follows because \( \beta x'' - c'' < 0 \) and \( x' > 0 \).

32) suggests that the principal can induce greater effort from the agent if she increases the agent’s share of output. However, increasing the agent’s share of output involves two costs for the principal. The first is obvious – the principal will receive a smaller share of output. In addition, if the agent receives a larger share of output he bears more risk, which the principal has to compensate for to induce him to accept the contract.

Now that the manager knows how the worker will respond to any linear contract she may offer, she can determine the optimal contract. Assume that she is risk-neutral. Using 24) her expected payoff is

\[ x(e) - s(x(e)) = x(e) - \alpha - \beta x(e). \quad 33) \]
The principal chooses the compensation scheme to maximize 33) subject to a participation constraint and an incentive compatibility constraint. Using 29) the participation constraint is

$$\alpha + \beta x(e) - c(e) - r \beta^2 \sigma^2 \geq \bar{u}. \tag{34}$$

The incentive compatibility constraint is

$$e = e^*(\beta). \tag{35}$$

Since 34) will bind, substitute 34) and 35) into 33) to obtain

$$x(e^*(\beta)) - c(e^*(\beta)) - r \beta^2 \sigma^2 - \bar{u}. \tag{36}$$

The optimal compensation scheme involves choosing $\beta$ to maximize 36) and then choosing $\alpha$ to satisfy the participation constraint 34).

The first-order condition for maximizing 36) with respect to $\beta$ is

$$[x'(e^*(\beta)) - c'(e^*(\beta))][de^*(\beta)/d \beta] - 2r \beta \sigma^2 = 0. \tag{37}$$

Since $de^*/d \beta > 0$ and $2r \beta \sigma^2 > 0$, 37) holds only if $x'(e^*) - c'(e^*) > 0$. From this we have several conclusions:

a) The principal’s compensation scheme does not induce the agent to choose the first-best, efficient level effort (which recall would require $x'(e^*) - c'(e^*) = 0$).

b) Under typical curvature conditions ($x' > 0$, $x'' < 0$, $c' > 0$, and $c'' > 0$), $e^*(\beta)$ is less than the efficient level. Graphically:
c) Result b) implies that $\beta < 1$; hence the principal and agent share residual claimancy as well as the production risk.

d) a) thru c) require that the agent be risk-averse ($r > 0$). For a risk-neutral agent ($r = 0$), 37) implies $x'(e^*) - c'(e^*) = 0$ and $e^*$ is efficient. Furthermore, if $r = 0$ then $\beta = 1$. This implies that if the worker is risk-neutral, the optimal contract makes him the sole residual claimant of output.

**Comparative statics**

Now, let us find the comparative statics that reveal how the optimal $\beta$ changes with changes in $r$ and $\sigma^2$. Note that an increase in $r$ implies that the worker is more risk averse – he has a lower tolerance for risk. Also, an increase in $\sigma^2$ implies an increase in output variability – an increase in the uncertainty about how the worker’s effort will determine output.

To derive the comparative statics, first differentiate the first-order condition 37) with respect to $\beta$ to determine the second-order condition:

$$s(\beta) = (x^* - c^*) \left( \frac{d e^*}{d \beta} \right)^2 + (x' - c') \left( \frac{d^2 e^*}{d \beta^2} \right) - 2r\sigma^2 < 0. \quad 38$$

Provided that $s(\beta) < 0$, the first order condition (37) implicitly defines $\beta^* = \beta(r, \sigma^2)$. 37) evaluated at $\beta^*$ holds as an identity:

$$[x'(e^*(\beta^*)) - c'(e^*(\beta^*))] \left[ \frac{de^*(\beta)}{d \beta} \right] - 2r \beta^* \sigma^2 = 0. \quad 39$$

Differentiate 39) with respect to $r$ to obtain

$$\left( x^* \frac{de^*}{d \beta} \frac{\partial \beta^*}{\partial r} - c^* \frac{de^*}{d \beta} \frac{\partial \beta^*}{\partial r} \right) \left( \frac{de^*}{d \beta} \right) + \left( \frac{d^2 e^*}{d \beta^2} \frac{\partial \beta^*}{\partial r} \right) (x' - c') - 2\sigma^2 \left( \beta^* + r \frac{\partial \beta^*}{\partial r} \right) = 0.$$

Collect terms:

$$\frac{\partial \beta^*}{\partial r} \left[ (x'' - c'') \left( \frac{de^*}{d \beta} \right)^2 + (x' - c') \frac{d^2 e^*}{d \beta^2} + 2r\sigma^2 \right] - 2\sigma^2 \beta^* = 0.$$

Observe that the term in hard brackets is $s(\beta)$ defined by 38). Therefore

$$\frac{\partial \beta^*}{\partial r} = \frac{2\sigma^2 \beta^*}{s(\beta^*)} < 0. \quad 40$$

To obtain $\partial \beta^*/\partial \sigma^2$, first differentiate 39) with respect to $\sigma^2$: \[\text{...}\]
Asymmetric Information

\[ \left( x^* \frac{d^2 e^*}{d \beta^* \partial \sigma^2} - c^* \frac{d e^*}{d \beta^* \partial \sigma^2} \right) + \left( d^2 e^* \frac{d^2 \beta^*}{d \beta^2 \partial \sigma^2} \right) (x^* - c^*) - 2r \left( \beta^* + \sigma^2 \frac{\partial \beta^*}{\partial \sigma^2} \right) = 0. \]

As we did above, collect terms, substitute \( s(\beta) \), and rearrange the result to obtain

\[ \frac{\partial \beta^*}{\partial \sigma^2} = \frac{2r \beta^*}{s(\beta^*)} < 0. \]  

41)

Since \( \beta \) is the worker’s share of output it is also an indicator of how much risk he will take on in this relationship. 40) indicates that the optimal contract will have him take on less if he is more risk-averse, and 41) indicates he will take on less risk the more uncertain the production relationship.

**Costly Monitoring of Actions**

We have been assuming that the principal (the manager) can either observe the agent’s behavior (the worker’s level of effort) perfectly and without cost, or that is completely impossible for the principal to observe the agent’s behavior. In more realistic settings, the principal will be able to observe that agent’s behavior, but not perfectly and only if she undertakes costly monitoring.

In this section we will examine our manager/worker relationship when the manager monitors the worker’s effort with some cost and only imperfectly. First, we will examine the trade-off the manager faces in choosing a compensation package and a monitoring scheme to induce a certain level of effort from the worker. Then we will examine the manager’s choice of effort by the worker when monitoring is imperfect and costly. In general, the manager’s optimal level of effort when costly monitoring is necessary is less than the first-best, efficient level of effort.

Suppose that the manager wants to induce \( \bar{e} \) effort from the worker. The worker may have the incentive to supply less effort \((e < \bar{e})\). To avoid this the manager will monitor the worker’s effort, but does so imperfectly and with some cost. Let \( \pi \) be the probability that shirking by the worker is detected. Achieving this level of monitoring (observability) costs the manager \( K(\pi) \), with \( K' > 0 \) and \( K'' > 0 \). In addition the compensation package includes a penalty if the worker is caught shirking along with a standard payment \( s \). If the worker shirks \((e < \bar{e})\) and is detected, his payment is \( s_0 < s \). The manager could make sure the worker chooses \( e = \bar{e} \) with very little monitoring if she can make \( s_0 \) very low (perhaps \( s_0 = 0 \) implying dismissal) or even negative (the worker pays a fine for shirking). However, there are likely to be practical or legal limits to how low the penalty wage can be, so let’s assume that \( s_0 \geq \bar{s}_0 \).

Assuming that the manager and worker are risk-neutral the manager chooses the level of monitoring \( (\pi) \), the standard wage \( (s) \), and the penalty wage \( (s_0) \) to minimize her costs of inducing \( \bar{e} \) level of effort from the worker. That is, she solves
\[
\min_{\pi, s, s_0} \ s + K(\pi) \quad 42)
\]
subject to
\[
\begin{align*}
\quad & s - c(\bar{e}) \geq u_0; \quad 43) \\
\quad & s - c(\bar{e}) \geq \pi s_0 + (1 - \pi) s - c(e); \quad 44) \\
\quad & s_0 \geq \bar{s}_0; \quad 45) \\
\quad & \pi \in [0, 1]. \quad 46)
\end{align*}
\]

Equation 43) is the participation constraint, while equation 44) is the incentive-compatibility constraint. The left-hand side of 44) is the worker’s payoff from choosing \( e = \bar{e} \), while the right-hand side is his expected payoff from shirking – from choosing \( e < \bar{e} \). Note however that \( \pi s_0 + (1 - \pi) s - c(e) \) is strictly decreasing in \( e \). Therefore, if the worker decides to shirk he will choose \( e = 0 \), and 44) should be \( s - c(\bar{e}) \geq \pi s_0 + (1 - \pi) s \). Rewrite this as

\[
\pi(s - s_0) \geq c(\bar{e}). \quad 44')
\]

Note that the left-hand side of 44') includes all of the instruments available to the manager to induce \( \bar{e} \) level of effort. To make sure the constraint holds the manager can increase \( \pi \) or \( s \), both of which are costly, or she can decrease \( s_0 \), which isn’t costly for her. Therefore, she will set \( s_0 \) as low as possible, which implies that 45) will be binding; that is,

\[
s_0 = \bar{s}_0. \quad 45')
\]

The incentive-compatibility constraint 44') also implies \( \pi > 0 \). Therefore, 46) can be simplified to \( \pi \leq 1 \). To simplify matters further let’s assume that the manager will never be able to monitor the worker perfectly, so that \( \max \pi < 1 \). Taken together, \( 0 < \pi < 1 \), allows us to drop 46) as a constraint on the problem.

With our simplifications the manager chooses \( \pi \) and \( s \) to minimize \( s + K(\pi) \) subject to 43), 44'), and 45'). The Lagrange equation is:

\[
L = s + K(\pi) - \lambda_1(s - c(\bar{e}) - u_0) - \lambda_2[\pi(s - \bar{s}_0) - c(\bar{e})],
\]

with Kuhn-Tucker conditions:

\[
\begin{align*}
\lambda_1 = 1 - \lambda_1 - \lambda_2 \pi = 0; \quad 47) \\
L_\pi = K'(\pi) - \lambda_2 s = 0; \quad 48) \\
L_{\lambda_1} = -(s - c(\bar{e}) - u_0) \leq 0, \quad \lambda_1 \geq 0, \quad \lambda_1(s - c(\bar{e}) - u_0) = 0; \quad 49) \\
L_{\lambda_2} = -[\pi(s - \bar{s}_0) - c(\bar{e})] \leq 0, \quad \lambda_2 \geq 0, \quad \lambda_2[\pi(s - \bar{s}_0) - c(\bar{e})] = 0. \quad 50)
\end{align*}
\]
Since \( K'(\pi) > 0 \) and \( s > 0 \), (48) implies that \( \lambda_2 > 0 \). Therefore, the incentive-compatibility constraint is binding, and (50) implies that the optimal contract involves \( \pi(s - s_o) = c(\varepsilon) \). The optimal amount of monitoring is therefore

\[
\pi^* = \frac{c(\varepsilon)}{(s - s_o)}. \tag{51}
\]

Note that:

\[\frac{\partial \pi^*}{\partial c(\varepsilon)} > 0 \quad \text{– The more costly is } \varepsilon \text{ for the worker the greater his incentive to shirk. To offset this incentive the manager needs to monitor him more closely.}\]

\[\frac{\partial \pi^*}{\partial \varepsilon_o} = c(\varepsilon)/(s - \varepsilon_o)^2 > 0 \quad \text{– By setting the penalty wage as low as possible the manager can minimize the amount of monitoring needed to make sure the worker chooses } \varepsilon = \varepsilon. \text{ If the available penalty wage is lower, the manager is able to monitor the worker less closely.}\]

\[\frac{\partial \pi^*}{\partial s} = -c(\varepsilon)/(s - \varepsilon_o)^2 < 0 \quad \text{– Higher regular compensation for the worker reduces his incentive to shirk; hence, the manager needs to monitor him less closely. This implies that monitoring and regular compensation are substitute instruments for inducing } \varepsilon. \text{ However, both are costly. The manager must therefore choose the combination of } \pi \text{ and } s \text{ that minimizes total costs.}\]

To investigate how to set the optimal wage \( s \), given optimal monitoring \( \pi^* \), substitute (51) into the manager’s objective to obtain

\[
v^*(s, \varepsilon_o, \varepsilon) = s + K \left( \frac{c(\varepsilon)}{(s - \varepsilon_o)} \right). \tag{52}
\]

Differentiate with respect to \( s \):

\[
\frac{\partial v^*}{\partial s} = 1 - K'(\pi^*) \left( \frac{c(\varepsilon)}{(s - \varepsilon_o)^2} \right). \tag{53}
\]

Differentiate again to obtain

\[
\frac{\partial^2 v^*}{\partial s^2} = - \left[ K''(\pi^*) \left( \frac{c(\varepsilon)}{(s - \varepsilon_o)^2} \right) \left( \frac{c(\varepsilon)}{(s - \varepsilon_o)^2} \right) \right] - K'(\pi^*) \left( \frac{-2(s - \varepsilon_o)c(\varepsilon)}{(s - \varepsilon_o)^4} \right)
\]

\[
= K''(\pi^*) \left( \frac{c(\varepsilon)}{(s - \varepsilon_o)^2} \right)^2 + 2K'(\pi^*) \left( \frac{(s - \varepsilon_o)c(\varepsilon)}{(s - \varepsilon_o)^4} \right) > 0.
\]

[You will need to use \( \frac{\partial \pi^*}{\partial s} = -c(\varepsilon)/(s - \varepsilon_o)^2 \) to do this derivation].
Since \( v^* \) is strictly convex in \( s \), the optimal wage exists and is unique. Recall, however that we must satisfy the participation constraint (43). Therefore, the lowest possible wage is \( s = c(\bar{e}) + u_0 \).

There are two possibilities for the optimal wage.

**(A)** \( s^* = c(\bar{e}) + u_0 \). In this case, (53) evaluated at \( s^* = c(\bar{e}) + u_0 \) is strictly positive. Graphically,

![Graph showing the optimal wage](image)

**(B)** \( s^* > c(\bar{e}) + u_0 \). In this case, (53) evaluated at \( c(\bar{e}) + u_0 \) is negative and \( s^* \) is the solution to

\[
\frac{\partial v^*}{\partial s} = 1 - K'(\pi^*) \left( \frac{c(\bar{e})}{(s - \bar{s}_0)^2} \right) = 0,
\]

where \( \pi^* = c(\bar{e})/(s - \bar{s}_0) \) [equation (51)]. Graphically,

![Graph showing the optimal wage (B)](image)

In this case the worker earns a wage that gives him a payoff that is higher than his reservation level of utility. To get some intuition, reconsider

\[
\frac{\partial v^*}{\partial s} = 1 - K'(\pi^*) \left( \frac{c(\bar{e})}{(s - \bar{s}_0)^2} \right).
\]
The second term is marginal monitoring costs. The first term (1) is the marginal cost of offering a higher wage. Since \( s^* > c(\bar{e}) + u_0 \) when \( \partial s^*/\partial s < 0 \) at \( s = c(\bar{e}) + u_0 \), the worker earns at rent \([ s^* > c(\bar{e}) + u_0 \)] when marginal monitoring costs are relatively high. To conserve on monitoring costs the manager offers a higher wage than \( s = c(\bar{e}) + u_0 \).

To sum up, the solution to the manager’s problem of inducing \( e = \bar{e} \) in the cheapest manner possible involves

(i) \( s_0 = \bar{s}_0 \) -- set the penalty wage as low as possible.

(ii) \( \pi^*(s^*, \bar{e}, \bar{s}_0) = c(\bar{e})/(s^* - \bar{s}_0) \) -- choose the level of monitoring that gives the worker the correct incentive to choose \( e = \bar{e} \) rather than shirk

(iii) \( s^* = \begin{cases} 
    c(\bar{e}) + u_0 & \text{if } \partial v^*(s = c(\bar{e}) + u_0, \bar{e}, \bar{s}_0)/\partial s \geq 0; \\
    s & \text{such that } \partial v^*(s, \bar{e}, \bar{s}_0)/\partial s = 0 & \text{if } \partial v^*(s = c(\bar{e}) + u_0, \bar{e}, \bar{s}_0)/\partial s < 0.
\end{cases} \)

Thus far we have treated \( \bar{e} \) as exogenous. However, monitoring and its costs may have an impact on the optimal choice of \( \bar{e} \). Suppose that output is determined by \( x(\bar{e}) \). When effort is perfectly observable, the optimal level of effort is determined from

\[ x'(\bar{e}) - c'(\bar{e}) = 0. \tag{55} \]

To examine the optimal level of effort with costly monitoring write the cost-minimizing levels of the wage and monitoring as \( s^* = s(\bar{e}, \bar{s}_0) \) and \( \pi^*(s^*, \bar{e}, \bar{s}_0) = c(\bar{e})/(s^* - \bar{s}_0) \). Use (52) to write the indirect cost function for the manager:

\[ v(\bar{e}, \bar{s}_0) = s^* + K\left( \frac{c(\bar{e})}{(s^* - \bar{s}_0)} \right). \tag{56} \]

The optimal level of effort is chosen to maximize

\[ x(\bar{e}) - v(\bar{e}, \bar{s}_0). \tag{57} \]

Assuming strict concavity of \( x(\bar{e}) - v(\bar{e}, \bar{s}_0) \), the optimal \( \bar{e} \) will be the solution to

\[ x'(\bar{e}) - \partial v(\bar{e}, \bar{s}_0)/\partial \bar{e} = 0. \tag{58} \]

Determining how the need for costly monitoring affects the optimal choice of effort involves comparing (58) to (55), which reduces to a comparison of the manager’s marginal costs; that is,
\[ c'(\bar{e}) - \partial v(\bar{e}, \bar{s}_0)/\partial e. \]  

Return to 56). When the participation constraint is not binding \((s^* > c(\bar{e}) + u_0)\), we can apply the envelope theorem to 56) to obtain

\[ \frac{\partial v(\bar{e}, \bar{s}_0)}{\partial e} = K'(\pi^*) \left( \frac{c'(\bar{e})}{(s^* - \bar{s}_0)} \right). \]  

Recall the first order condition for determining \(s^* > c(\bar{e}) + u_0\), equation 54),

\[ 1 - K'(\pi^*) \left( \frac{c'(\bar{e})}{(s - \bar{s}_0)^2} \right) = 0, \]  

which can be rewritten as

\[ \frac{s^* - \bar{s}_0}{c(\bar{e})} = \frac{K'(\pi^*)}{s^* - \bar{s}_0}. \]

Since \(\pi^* = c(\bar{e})/(s^* - \bar{s}_0)\),

\[ \frac{K'(\pi^*)}{s^* - \bar{s}_0} = \frac{1}{\pi^*}. \]  

Substitute 62) into 60) to obtain

\[ \frac{\partial v(\bar{e}, \bar{s}_0)}{\partial e} = \frac{c'(\bar{e})}{\pi^*}. \]  

Since \(\pi^* < 1\), 63) implies that \(\partial v(\bar{e}, \bar{s}_0)/\partial e > c'(\bar{e})\). This reveals that with costly monitoring and \(s^* > c(\bar{e}) + u_0\), the optimal level of effort is lower than when effort is perfectly observable. Graphically,
The same result obtains when the participation constraint is binding in the costly monitoring problem. In this case the wage is \( s^* = c(\bar{e}) + u_0 \) and the monitoring probability is

\[
\pi^* = \frac{c(\bar{e})}{c(\bar{e}) + u_0 - s_0}.
\]

Since \( \pi^* < 1 \), \( c(\bar{e}) < c(\bar{e}) + u_0 - s_0 \), which requires \( u_0 - s_0 > 0 \). We will use this in a moment.

Substitute \( s^* = c(\bar{e}) + u_0 \) into (56) to obtain the manager’s indirect cost function when the participation constraint binds:

\[
v(\bar{e}, \bar{s}_0) = c(\bar{e}) + u_0 + K \left( \frac{c(\bar{e})}{c(\bar{e}) + u_0 - \bar{s}_0} \right).
\]

Differentiate (64) with respect to \( \bar{e} \):

\[
\frac{\partial v(\bar{e}, \bar{s}_0)}{\partial \bar{e}} = c'(\bar{e}) + K'(\pi^*) \left( \frac{c'(\bar{e})(c(\bar{e}) + u_0 - \bar{s}_0) - c'(\bar{e})c(\bar{e})}{(c(\bar{e}) + u_0 - \bar{s}_0)^2} \right)
\]

\[
= c'(\bar{e}) + K'(\pi^*) \left( \frac{c'(\bar{e})(u_0 - \bar{s}_0)}{(c(\bar{e}) + u_0 - \bar{s}_0)^2} \right).
\]

Since \( K' > 0 \), \( c' > 0 \), and \( u_0 - s_0 > 0 \), the second term of (65) is strictly positive. Therefore, \( \frac{\partial v(\bar{e}, \bar{s}_0)}{\partial \bar{e}} > c'(\bar{e}) \), which again implies that the manager should try to induce a level of effort that is lower than the level she would choose if she could observe the worker’s effort perfectly and without cost.

The model of costly monitoring suggests that imperfect observability generates two sorts of costs: (1) the direct costs of monitoring, (2) the loss that is due to the fact that the optimal level of effort when it is imperfectly observable is less than the first-best efficient level of effort.

**Hidden Information**

Now let us assume that the principal is uncertain about some relevant characteristic of the agent, perhaps his skill level. Let us also assume that the worker’s effort is perfectly observable without cost, and that there is no uncertainty about how effort affects output; that is, let \( x = x(e) \). In this setting, a contract can specify either output or the worker’s effort, it doesn’t matter. Let’s assume that the contract specifies output.

Suppose that there are just two types of workers, high cost workers (type \( h \)) and low cost workers (type \( l \)). A type \( t \) worker can produce output \( x \) at cost \( c_t(x) \). Suppose that:

7.21
\[ c_h(x) > c_l(x) \text{ and } c_h'(x) > c_l'(x) \quad \forall x \]

The low cost worker is more efficient. Our assumptions about \( c_t(x), t=h, l \), imply that for any \( x^0 \) and \( x^1 \) such that \( x^0 > x^1 \),

\[ c_h(x^0) - c_h(x^1) > c_l(x^0) - c_l(x^1). \quad (66) \]

To illustrate, consider:

\[ c_h(x^0) - c_h(x^1) = A + B, \text{ while } c_l(x^0) - c_l(x^1) = A, \text{ which confirms } 66). \]

The manager is going to offer to the worker a contract consisting of a compensation \( s_t \) and an output target \( x_t \) that depends on the worker’s type. The problem is that the manager doesn’t know the workers type, but the worker does.

Let the manager’s utility be

\[ u_m = x_t - s_t, \quad t = h, l, \]

and let the worker’s utility be

\[ u_{wt} = s_t - c_t(x_t), \quad t = h, l. \]

For simplicity let us suppose that the worker’s reservation utility is zero.

**Complete information benchmark:** If the manager knows that the worker is type \( t \), she will offer a contract \( (s_t, x_t) \) that maximizes \( x_t - s_t \), subject to satisfaction of the worker’s participation constraint \( s_t - c_t(x_t) \geq 0 \). The manager will set \( s_t \) as low as possible so that \( s_t = c_t(x_t) \). Her utility is then \( x_t - c_t(x_t) \). Maximizing this function requires \( c_t'(x_t) = 1 \). Denoting the solution to this as \( \bar{x}_t, \quad t = h, l \), we have:
To summarize, the complete information contract for a type $t$ worker is

$$(s_t, x_t) = (c_t(x_t), x_t). \quad (67)$$

From the graph, $\bar{s}_h = A + B$ and $\bar{s}_l = B + C$.

**Asymmetric Information:** When the manager does not know the worker’s type the contract just described encounters a problem; namely, a low cost worker would like the manager to believe that he is a high-cost worker. To see this note that if a type $l$ worker reveals that he is type $l$, his payoff is $(B + C) - c_l(x_l) = 0$. If he pretends to be a type $h$ worker instead, his compensation is $\bar{s}_h = A + B$ but his costs are only $c_h(\bar{x}_h) = B$: therefore, he would get a surplus of $A$. [It is easy to show that a high cost worker will never pretend that he is a low cost worker].

The problem for the manager is to design a contract that takes account of this incentive problem. The manager’s problem of offering a type-specific contract without knowing what type she has encountered may appear rather daunting. Fortunately, the remarkable result below greatly simplifies her problem.

**The Revelation Principle:** Suppose that the set of possible types is $T$, and suppose further that the manager presents the worker with a menu of contracts $(s_t, x_t)$, one for each $t \in T$, and lets the worker choose for himself which contract he prefers. Then, the manager can do no better than to design $(s_t, x_t) \forall t \in T$, so that the worker has a dominant strategy to choose the specific contract that is designed for his type.

The Revelation Principle allows us to write the manager’s problem as follows: the manager chooses a pair of type-specific contracts $\{(s_t, x_t), (s_h, x_h)\}$ to maximize
\[ \phi (x_l - s_l) + (1 - \phi) (x_h - s_h) \]  

subject to

\[ s_l - c_l(x_l) \geq 0 \]  
\[ s_h - c_h(x_h) \geq 0 \]  
\[ s_l - c_l(x_l) \geq s_h - c_h(x_h) \]  
\[ s_h - c_h(x_h) \geq s_l - c_l(x_l). \]

The manager’s objective function 68) is her expected utility, where \( \phi \) is the probability she has encountered a type-\( l \) worker and \( (1 - \phi) \) is the probability she has encountered a type-\( h \) worker. Constraints 69) and 70) are participation constraints, one for each type of worker. Constraints 71) and 72) are incentive-compatibility constraints that guarantee that the worker chooses the contract intended for his type. For example 71) guarantees that an \( l \) type worker will not chose the contract \( (s_h, x_h) \).

Let the solution to the manager’s problem be \([ (s^*_l, x^*_l), (s^*_h, x^*_h) ] \). Our task now is to characterize this solution and compare the contracts to the complete information contracts \( (s_t, x_t), t = h, l \). It is convenient to simplify the manager’s problem by deriving a number of results before we actually carry out the manager’s optimization problem.

**Result 1:** The optimal contracts must involve \( x^*_h \leq x^*_l \).

**Proof:** Rearrange the incentive-compatibility constraints 71) and 72):

\[ s_h \leq s_l + c_l(x_h) - c_l(x_l); \]
\[ s_h \geq s_l + c_h(x_h) - c_h(x_l). \]

Together these imply \( s_l + c_l(x_h) - c_l(x_l) \geq s_l + c_h(x_h) - c_h(x_l) \), or equivalently,

\[ c_h(x_l) - c_h(x_h) \leq c_l(x_l) - c_l(x_h). \]

If \( x_h > x_l \), 66) would be violated. [Recall that 66) is a consequence of our assumption that \( c'_h > c'_l (x), \forall x \).] Therefore, \( x^*_h \leq x^*_l \). Q.E.D.

That a low-cost worker should be required to produce a level of output that is not less than a less efficient worker is very intuitive.

**Result 2:** The incentive-compatibility constraint 71) is binding, but the participation constraint 69) is not; that is,

\[ s^*_l - c_l(x^*_l) > 0; \]
\[ s^*_l = c_l(x^*_l) + [s^*_h - c_l(x^*_h)]. \]  

\[ 7.24 \]
Proof: Rewrite 69) and 71) as

\[ s_l \geq c_l(x_l); \]
\[ s_l \geq c_l(x_l) + [s_h - c_l(x_h)]. \]

Since the manager wants to set \( s_l \) as low as possible, either one of these constraints is binding and the other is not, or both are binding. Take the participation constraint 71):

\[ s_l - c_l(x_l) \geq s_h - c_l(x_h). \]

Note that the other participation constraint 70), \( s_h - c_h(x_h) \geq 0 \), and \( c_l(x_h) < c_h(x_h) \) imply \( s_h - c_l(x_h) > 0 \). Therefore, for 71) to hold, \( s_l > c_l(x_l) \), which implies that 69) is not binding. Since 69) is not binding, 71) must be. Q.E.D.

Result 2 implies that the optimal contract involves

\[ s_l^* = c_l(x_l^*) + [s_h^* - c_l(x_h^*)] > c_l(x_l^*). \]  \hspace{1cm} 74)

Equation 74) says something quite important, namely a low-cost worker earns a surplus of \( s_h^* - c_l(x_h^*) \) in his dealings with the manager. This surplus is paid to keep an \( l \)-type worker from masquerading as an \( h \)-type worker. The surplus is often call an information rent – it is an extra payment necessary to keep the worker from using his private information about his type to manipulate the contract.

Result 3: If \( x_l^* > x_h^* \), the participation constraint 70) is binding but the incentive-compatibility constraint 72) is not; that is,

\[ s_l^* - c_l(x_l^*) = 0; \] \hspace{1cm} 75)
\[ s_h^* > c_h(x_h^*) + [s_l^* - c_l(x_l^*)]. \]

Proof: Rewrite 70) and 72) as

\[ s_h \geq c_h(x_h); \] \hspace{1cm} 76)
\[ s_h \geq c_h(x_h) + [s_l - c_l(x_l)]. \] \hspace{1cm} 77)

Now, one of these constraints is binding while the other is not, or both are binding. Suppose that 77) is binding. Substitute for \( s_l \) from 73) and rewrite 77) as

\[ s_h = c_h(x_h) + c_l(x_l) + s_h - c_l(x_h) - c_h(x_l). \]

This implies \( c_h(x_h) + c_l(x_l) - c_l(x_h) - c_h(x_l) = 0 \), or rather

\[ c_h(x_l) - c_l(x_h) = c_l(x_l) - c_l(x_h). \]
But with $x_l > x_h$, 66) requires
\[ c_h(x_l) - c_h(x_h) > c_l(x_l) - c_l(x_h). \]

Therefore, 77), and hence, 72) cannot be binding. It follows then that 77), and hence 70), are binding. Q.E.D.

We have only shown $x_l^* \geq x_h^*$ in Result 1, but we have assumed $x_l^* > x_h^*$ in Result 3. From here, we are going to continue to assume $x_l^* > x_h^*$ and show that at the optimal solution this is in fact true.

With Results 2 and 3, we can simplify the manager’s problem by setting $s_l = c_l(x_l) + s_h - c_l(x_h)$ and $s_h = c_h(x_h)$ [constraints 70) and 71) are binding] and ignoring constraints 69) and 72).

Substitute $s_l = c_l(x_l) + s_h - c_l(x_h)$ and $s_h = c_h(x_h)$ into the manager’s objective to obtain her choice problem:

\[
\max_{(x_l, x_h)} \phi(x_l - c_l(x_l) - c_h(x_h) + c_l(x_h)) + (1 - \phi)(x_h - c_h(x_h)).
\]

The first-order conditions for $(x_l^*, x_h^*)$ are
\[
\phi(1 - c_l'(x_l^*)) = 0;
\]
\[
-\phi(c_h'(x_h^*) - c_l'(x_l^*)) + (1 - \phi)(1 - c_h'(x_h^*)) = 0,
\]

which can be rewritten as
\[ c_l'(x_l^*) = 1; \quad 78) \]
\[ c_h'(x_h^*) = 1 + \frac{\phi}{(1 - \phi)}[c_l'(x_l^*) - c_h'(x_h^*)] < 1. \quad 79) \]

The inequality in 79) follows because $c_l'(x_l^*) < c_h'(x_h^*)$.

Recall from Results 2 and 3, [equations 73) and 75)] that the compensation payments are:
\[ s_l^* = c_l(x_l^*) + [s_h^* - c_l(x_h^*)]; \]
\[ s_h^* = c_h(x_h^*). \]

Furthermore, recall that under complete information, $\bar{x}_l = x_l \mid c_l'(x_l) = 1$ and $\bar{x}_h = c_l'(x_h)$. So relative to the complete information outcome, the contract for a low-cost worker involves the same output but the payment gives the worker a surplus, while the contract for a high-cost worker leaves him no surplus buts distorts his output target below the complete information target. Graphically:
It is easy to see that $s_h^* = c_h(x_h^*) = A + B$. Furthermore, since

$$s_l^* = c_l(x_l^*) + s_h^* - c_l(x_h^*) = c_l(x_l^*) + c_h(x_h^*) - c_l(x_h^*),$$

$$s_l^* = A + B + C + D.$$

Area $A$ is the information rent earned by a low-cost worker. It is therefore clear why the manager chooses $x_h^*$ lower than the first-best choice of $x_h$ (as characterized by the complete information choice): it does so to decrease the information rent it would have to pay to a low-cost worker. This trade-off between information rents and efficiency is a fundamental characteristic of hidden-information, principal agent problems.

There is one last detail: Recall that we proved $x_l^* \geq x_h^*$ but assumed that $x_l^* > x_h^*$ and went ahead with the problem. In fact, it is clear from the graph that $x_l^* > x_h^*$, so our assumption was justified.
Exercises

[1] Consider the “hidden action model with mean-variance utility” from your lecture notes. In addition to what is assumed in the notes, assume that $x(e) = pe$ and $c(e) = ce^2$. With these assumptions the model is fully parameterized, implying that we can now solve for explicit values of the choice variables. To help in your interpretations of your results note that $p$ is the worker’s marginal product of effort, and that $c$ is an indicator of the worker’s cost and marginal cost of effort.

[a] Derive the complete information, efficient level of effort.

[b] Given a compensation package, $s(\hat{x}) = \alpha + \beta \hat{x}$, where $\hat{x}$ is defined in the notes, derive the worker’s choice of effort. How does this choice compare to the full-information, efficient level of effort?

[c] Denote the worker’s choice of effort as $e^*$. Find $\partial e^*/\partial p$ and $\partial e^*/\partial c$. Interpret these comparative static results.

[d] State the principal’s (manager’s) problem of determining the worker’s compensation. Make sure that you identify the participation and incentive-compatibility constraints.

[e] Derive the level of $\beta$ the manager will offer to the worker. Denote your solution as $\beta^*$. How do we interpret $\beta^*$? What is $\beta^*$ when the worker is risk-neutral? What is $\beta^*$ when there is no randomness in the production relationship?

[f] Derive $\partial \beta^*/\partial r$, $\partial \beta^*/\partial \sigma^2$, $\partial \beta^*/\partial p$, and $\partial \beta^*/\partial c$. Interpret each result.

[g] Derive the level of effort the manager is trying to induce from the worker. Hint: this is $e^*(\beta^*)$. Derive $\partial e^*(\beta^*)/\partial r$, $\partial e^*(\beta^*)/\partial \sigma^2$, $\partial e^*(\beta^*)/\partial p$, and $\partial e^*(\beta^*)/\partial c$, and interpret each result.

[h] How does $e^*(\beta^*)$ compare to the complete information, efficient level of effort.

[i] Denote the full-information, efficient level of effort as $\bar{e}$. How does $\bar{e} - e^*(\beta^*)$ change with $p$, $c$, $r$ and $\sigma^2$? Interpret your results.

[j] Given $\beta^*$ from [e], derive $\alpha^*$. Derive $\partial \alpha^*/\partial r$, $\partial \alpha^*/\partial \sigma^2$, $\partial \alpha^*/\partial p$, and $\partial \alpha^*/\partial c$. Interpret each result by comparing them the comparative static results on $\beta^*$ from [f].

[k] Denote the expected compensation package as $s^* = \alpha^* + \beta^* x(e^*(\beta^*))$. Derive $\partial s^*/\partial r$, $\partial s^*/\partial \sigma^2$, $\partial s^*/\partial p$, and $\partial s^*/\partial c$. Interpret your results. You may want to refer to your answers from [f] and [j].
A single risk-neutral individual chooses the level of an activity \( x \geq 0 \). The net benefit the individual receives from this activity is \( b(x) \), which is strictly concave and reaches a maximum at \( \bar{x} \). Perhaps because the activity causes an external cost, an upper limit \( s < \bar{x} \) has been imposed. For example, \( x \) may stand for driving speed and \( s \) is a speed limit, or \( x \) is the amount of emissions released by a plant and \( s \) is an emission standard.

An enforcement authority monitors the individual with probability \( \pi \). If the individual is found to be in violation, \( x > s \), a penalty \( \phi(x - s) \) is imposed, where \( \phi \) is a positive constant. The individual chooses \( x \) to maximize its benefit from \( x \) minus the expected penalty from exceeding \( s \).

[a] Derive the combination of \( \pi \) and \( \phi \) that just induces the individual to choose \( x = s \).

Interpret and graph this outcome. Suppose that monitoring is costly but setting the unit penalty is not. How then should \( \pi \) and \( \phi \) be set?

[b] Suppose that \( \phi \) is restricted to be less than an exogenous value \( \phi^{\text{max}} \). Suppose further that \( \phi^{\text{max}} < b'(s) \) and \( \pi \leq 1 \). Show that the individual will be noncompliant; that is, he or she will choose \( x > s \). Graph your result and interpret.

[3] There are two basic types of principal/agent models, hidden information and hidden action models.

[a] In most principal/agent models the principal must satisfy two constraints. What is the purpose of the participation constraint in hidden action and hidden information problems?

[b] What is the purpose of the incentive compatibility constraint in hidden action models?

[c] What is the purpose of the incentive compatibility constraint in hidden information models?

[d] Risk sharing is an important characteristic of principal/agent relationships. In class we studied the relationship between a risk-neutral manager and a risk-averse worker when there is uncertainty about how the worker’s effort determined final output. Write an essay that describes and contrasts the allocation of risk between the manager and the worker when (1) the manager is able to observe the worker’s effort, and (2) the manager is unable to observe the worker’s effort. No derivations are required, but you must completely describe the nature of the contract between the manager and the worker in both cases, and then describe how each contract determines the allocation of risk between them.
Two firms supply output to a market with inverse demand \( p(Y) = A - y_1 - y_2 \). The firms are identical and have cost functions \( c(y_i) = ky_i \). The firms are Cournot duopolists in the sense that they choose their outputs simultaneously with complete information. However, while firm 2 is operated by its owner who seeks to maximize the firm’s profit, firm 1 is operated by a manager hired by its owner. The owner of firm 1 pays compensation \( w \) to its manager, which is a linear function of the firm’s profits \( (\pi_1) \), its sales \( (y_1) \), minus a lump sum \( F \). That is,

\[
w(\pi_1, y_1) = \alpha \pi_1 + \beta y_1 - F,
\]

where \( \alpha \in (0, 1) \) so that the manager receives of percentage of the firm’s profits, and \( \beta \) is a fixed constant that can be positive, zero, or negative. The parameters \( \alpha, \beta, \) and \( F \) have been fixed and the manager has agreed to operate firm 1. The structure of the contract is public knowledge; in particular, it is known to firm 2.

[a] Compute the Nash equilibrium choices of output. Remember that the operator of firm 2, its owner, seeks to maximize the firm’s profit, while the manager of firm 1 seeks to maximize his compensation.

[b] Show that these output choices are the same as the standard Cournot output choices (when each firm seeks to maximize its profits) if and only if \( \beta = 0 \).

[c] How do the equilibrium output choices compare to the standard Cournot output choices when:

- \( i \) the manager is penalized for increased sales \( (\beta < 0) \);
- \( ii \) the manager is rewarded for increased sales \( (\beta > 0) \).

Provide some intuition for your findings.

[d] Compute the equilibrium profit levels for each firm and compare them to the profit levels generated by standard Cournot competition when \( \beta < 0 \) and when \( \beta > 0 \).