1 The principal-agent problems

The principal-agent problems are at the heart of modern economic theory. One of the reasons for this is that it has widespread applicability. We start with some examples.

- Consider a seller trying to sell quantities of a good to a buyer. The value of the buyer for the good is not known to the seller. Indeed, if the value was known then the seller would optimize such that his marginal cost of production equals value. In the absence of this perfect information, the seller is constrained. When it offers certain quantity to a particular buyer type, it needs to ensure that it is optimal for such buyer to accept that offer. This introduces new constraints and distorts the "first-best" optimal.

- A firm has hired a manager to complete a project for him. The firm cannot observe the effort of the manager but observes his output. To incentivize the manager to work, the firm can give wages as a function of output. What is an optimal wage contract? What is the welfare loss due to unobservable effort?

The situation is similar when an insurance company gives insurance contracts to agents where it either cannot observe the characteristics of the agent or efforts put by the agents; a bank giving loans to agents where it cannot observe the characteristics of the lender or efforts put by the lender. The common thread in all these problems is that there are two parties: a principal and an agent. The principal does not have information about the agent. There are two kinds of information asymmetry: (a) the **characteristics** of the agent is not observed and (b) the **actions** of the agent is not observed. Both these information asymmetry lead to different kinds of problems. The first kind of problem is called the **adverse selection** problem (**hidden characteristics**) and the latter one is called the **moral hazard** problem (**hidden action**).

The main takeaways from these models is that the information asymmetry leads to welfare loss and the first-best is no longer possible. The focus of study is the nature of distortion from the first-best.

2 Adverse selection problem

Adverse selection problems involve a principal and an agent. In this model, an agent has a "characteristics", which is often referred to as the **type** of the agent. The principal does not observe the type of the agent. The basic idea is the following. If an insurance company offers a price tailored for the average population, then high risk agents will accept it and company will lose money. As a result, the optimal contract may deny high risk agents insurance.

The other term for adverse selection problem is **screening**. The basic idea can be described as follows. Suppose the principal is a wine seller and the agent is a buyer. There are two types of agents: *low* types (not a big connoisseur of wine) and *high* types (wine fans). High types are willing to pay high price for vintage wines. The principal cannot observe the types. But the principal can offer a menu of different wines with different prices. In particular, since high types are willing to pay more for high quality wines, the principal may offer a high quality wine at high price and a low quality wine at low price.

The hope is that types will then "separate" each other: high types will take the high quality wine and low types will take the low quality one. But the adverse selection story is that there will be some distortions. Some leading examples of this model are as follows.

- In life insurance, the insurer's state of health is not known to the insurance company. Offering a variety of insurance products to target specific risk classes is better for the insurance company. However, this may induce some distortions from efficiency.
- In banking, the borrowers' default risk can be imperfectly known by the bank. In that sense, having different interest rates to target different borrowers is a natural way to discriminate. This may induce credit rationing where high risk borrowers may take up more than their share of credit.
- In labor markets, workers know their abilities better than firms. Hence, firms must screen workers to select correct candidates and reject the bad ones.

2.1 A simple example

We try to understand some basic ideas of adverse selection. There are two types of wine buyers: $\theta_1 < \theta_2$. The buyer can provide a quality of wine to each type of the buyer and charge a price p. If a buyer of type θ is given quality q wine and charged p, then his payoff is

$$\theta q - p.$$

So, utility is quasilinear. There is a commonly known probability π with which a buyer is of type θ_1 and with probability $(1 - \pi)$, he is of type θ_2.

The seller has a cost function: $C : \mathbb{R}_+ \to \mathbb{R}_+$, which is strictly convex and twice differentiable with $C'(0) = 0$. The utility of the seller if he receives a transfer of t is $t - C(q)$, where q is the quality of wine sold.

The perfect information benchmark. If types were known then, the seller would offer each type θ_i a quality q_i and transfer p_i such that

$$\theta_i q_i = p_i.$$

So, it will maximize

$$p_1 - C(q_1) + p_2 - C(q_2) = \theta_1 q_1 + \theta_2 q_2 - C(q_1) - C(q_2).$$

Since C is strictly convex, this is a strictly concave objective function. Hence, first order condition gives us optimal (q_1^*, q_2^*) as $C'(q_1^*) = \theta_1$ and $C'(q_2^*) = \theta_2$. Since $\theta_1 < \theta_2$, $q_1^* < q_2^*$.

Imperfect information. Now, consider the scenario where the seller cannot observe the type of the buyer. Potentially, the seller can set up a complicated contract. However, as we will see later (due to a fact called the *revelation principle*), it is enough to consider a particular kind of *contracts* called the *direct mechanism* (which we simply refer to as a contract). In a direct mechanism, the seller asks buyers his type and commits to awarding a quality and price given the type. Formally, it announces two maps: $q : \{\theta_1, \theta_2\} \to \mathbb{R}_{++}$ and $p : \{\theta_1, \theta_2\} \to \mathbb{R}$. So, the timing of the "game" is as follows:

- The seller announces a contract (and he commits to it).
- The buyer realizes its type.
- The buyer announces a type.
- The buyer gets an outcome (quality, payment) pair according to the announced type and contract.
- The buyer and the seller realize their payoffs.

We will refer to the pair of maps (q, p) as a **contract**. A contract (q, p) needs to satisfy two constraints:

$$\begin{aligned}
q(\theta_1)\theta_1 - p(\theta_1) &\geq q(\theta_2)\theta_1 - p(\theta_2) \\
q(\theta_2)\theta_2 - p(\theta_2) &\geq q(\theta_1)\theta_2 - p(\theta_1) \\
q(\theta_1)\theta_1 - p(\theta_1) &\geq 0 \\
q(\theta_2)\theta_2 - p(\theta_2) &\geq 0.
\end{aligned}$$

The first two constraints are incentive compatibility (IC) constraints and the last two are individual rationality (IR) or participation constraint (outside option gives zero payoff). The seller maximizes his expected payoff:

$$\pi\big[p(\theta_1) - C(q(\theta_1))\big] + (1-\pi)\big[p(\theta_2) - C(q(\theta_2))\big].$$

We make several observations about the IC and IR constraints.

1. Adding the IC constraints, we get $(\theta_2 - \theta_1)(q(\theta_2) - q(\theta_1)) \geq 0$. Since $\theta_2 > \theta_1$, we get $q(\theta_2) \geq q(\theta_1)$.

2. If none of the IR constraints hold, then we can construct another contract (q, p'), where $p'(\theta_i) = p(\theta_i) + \epsilon$, where $\epsilon > 0$ but sufficiently small. It is clear that (q, p') also satisfies IC constraints. It also satisfies IR constraints since none of the IR constraints were binding. Also, the new contract improves seller's expected payoff. Hence, in the **optimal** contract (q, p), one of the IR constraints must bind.

3. The second IR constraint cannot be binding. Suppose it is - then, $q(\theta_2)\theta_2 - p(\theta_2) = 0$. Then, the second IC constraint becomes, $q(\theta_1)\theta_2 - p(\theta_1) \leq 0$. But $\theta_1 < \theta_2$ and $q(\theta_1) > 0$ implies $q(\theta_1)\theta_1 - p(\theta_1) < 0$, which violates the other IR constraint. Hence, the first IR constraint must be binding:
$$q(\theta_1)\theta_1 - p(\theta_1) = 0.$$

4. Once we know that the optimal contract must have $p(\theta_1) = q(\theta_1)\theta_1$, our IC and IR constraints simplify to

$$\begin{aligned}
p(\theta_2) &\geq q(\theta_2)\theta_1 \\
p(\theta_2) &\leq q(\theta_2)\theta_2 - q(\theta_1)\big(\theta_2 - \theta_1\big) \\
p(\theta_2) &\leq q(\theta_2)\theta_2.
\end{aligned}$$

Since $\theta_2 > \theta_1$, the second constraint implies the third constraint. Hence, relevant IC and IR constraints are

$$q(\theta_2)\theta_2 - q(\theta_1)\big(\theta_2 - \theta_1\big) \geq p(\theta_2) \geq q(\theta_2)\theta_1.$$

Clearly, in the optimal contract, we must have

$$p(\theta_2) = q(\theta_2)\theta_2 - q(\theta_1)\big(\theta_2 - \theta_1\big) = q(\theta_1)\theta_1 + \theta_2\big(q(\theta_2) - q(\theta_1)\big).$$

5. For sake of notation, denote $q_1 \equiv q(\theta_1)$ and $q_2 \equiv q(\theta_2)$. Then, our unconstrained objective function is (only a function of (q_1, q_2)):

$$\pi\big(\theta_1 q_1 - C(q_1)\big) + (1-\pi)\big(\theta_1 q_1 + \theta_2(q_2 - q_1) - C(q_2)\big).$$

First order condition with respect to q_2 gives us optimal quantity for type 2 is q_2^* which satisfies

$$\theta_2 = C'(q_2^*),$$

i.e., the perfect information benchmark quality. However, the first order condition with respect to q_1 gives us

$$C'(q_1^*) = \theta_1 - \frac{1-\pi}{\pi}\big(\theta_2 - \theta_1\big) < \theta_1.$$

Hence, quantity assigned to the lower type is less than his perfect information benchmark.

These five insights are common to all screening problems with discrete types. For completeness, we summarize them again below.

1. The highest type gets the perfect information benchmark quality.

2. All types except the highest type get lower quality than their perfect information benchmark quality.

3. The lowest type gets zero payoff (IR of lowest type binds).

4. All types except the lowest type get positive payoff: this is called *information rent* of higher types (IR of higher types do not bind).

5. Each type (except the lowest type) is indifferent between his consumption bundle and that of the immediately lower type (IC constraints of higher types bind).

Exercise. Work out the problem with k types: $\Theta = \{\theta_1, \ldots, \theta_k\}$ with $\theta_1 < \ldots < \theta_k$. Solve for the optimal contract.

Earlier, we had given a sequential timing interpretation of this contracting framework where the buyer was announcing his type. There is also a menu interpretation. Because of incentive compatibility and individual rationality, every buyer type θ finds it optimal to choose $(q(\theta), p(\theta))$ from the **range** of the outcomes of the contract (q, p). In other words, if $R^{f,p}$ is the range of outcomes of the contracts, then the seller can be thought of as announcing a menu of outcomes such that it is optimal for the buyer types to choose the correct outcome.

2.2 General type space

In this section, we flush out some details of the general model where type space is $\Theta = [0, 1]$ (or some closed interval). For simplicity, we assume that agent's utility is *linear*: for consuming quality q at price p, he gets utility equal to $q\theta - p$. This can be generalized by a function $u(q, \theta) - p$, where u is increasing in each argument and satisfies increasing differences property (or, single crossing).

As before, a **contract** is a pair of maps $q : \Theta \to \mathbb{R}_{++}$ and $p : \Theta \to \mathbb{R}$. Denote the net utility of agent of type θ by reporting θ' to the contract as:

$$U^{q,p}(\theta'|\theta) := q(\theta')\theta - p(\theta').$$

Definition 1 *A contract* (q, p) *is* **incentive compatible** *if for every* θ*,*

$$U^{q,p}(\theta|\theta) \geq U^{q,p}(\theta'|\theta).$$

Notice that

$$U^{q,p}(\theta'|\theta) = U(\theta'|\theta') + q(\theta')(\theta - \theta').$$

Incentive constraints say that for all $\theta, \theta' \in \Theta$,

$$U^{q,p}(\theta|\theta) \geq U^{q,p}(\theta'|\theta) = U^{q,p}(\theta'|\theta') + q(\theta')(\theta - \theta').$$

For simplicity of notation, we denote $U^{q,p}(\theta|\theta)$ as $U^{q,p}(\theta)$. Hence, we can write the IC constraints as

$$U^{q,p}(\theta) \geq U^{q,p}(\theta') + q(\theta')(\theta - \theta'). \tag{1}$$

Notice that if two types θ, θ' are such that $q(\theta) = q(\theta')$, then the pair of incentive constraints give us:

$$\theta q(\theta) - p(\theta) \geq \theta q(\theta') - p(\theta') = \theta q(\theta) - p(\theta')$$
$$\theta' q(\theta') - p(\theta') \geq \theta' q(\theta) - p(\theta) = \theta' q(\theta') - p(\theta)$$

Hence, we get $p(\theta) = p(\theta')$. This is called the **taxation principle**. Payment can be reduced to a map from quality to $\mathbb{R}$.

A routine exercise to show that if (q, p) satisfies IC constraints (1), then U is convex. A convex function is differentiable almost everywhere. Hence, if we pick any θ, θ' and use the pair of incentive constraints, we get

$$q(\theta)(\theta - \theta') \geq U^{q,p}(\theta) - U^{q,p}(\theta') \geq q(\theta')(\theta - \theta'). \tag{2}$$

Hence, as $\theta' \to \theta$, we see that if U is differentiable at θ, $U'(\theta) = q(\theta)$. So, the **derivative (whenever exists) of U is the quality** q. Hence, by fundamental theorem of calculus, for every $\theta \in [0, 1]$, we must have

$$U^{q,p}(\theta) = U^{q,p}(0) + \int_0^\theta q(\theta')d\theta'. \tag{3}$$

This is sometimes called the **payoff equivalence** formula - if there are two contracts using the same quality assignment rule: (q, p) and (q, p'), then they should differ from each other by the utility assigned to the lowest type. The payoff equivalence formula in Equation (3) also gives us a **revenue equivalence formula** by expanding the U terms: for all $\theta \in [0, 1]$,

$$p(\theta) = p(0) + q(\theta)\theta - \int_0^\theta q(\theta')d\theta'. \tag{4}$$

Now, we turn our attention to the IR constraints. It requires that $U(\theta) \geq 0$ for all θ. But payoff equivalence formula in Equation (3) requires that $U(0) + \int_0^\theta q(\theta')d\theta' \geq 0$. Since $\int_0^\theta q(\theta')d\theta' \geq 0$, this inequality holds if $U(0) \geq 0$ - also, $U(0) \geq 0$ is necessary. Hence, IR holds for all types if it holds for the lowest type: $U(0) \geq 0$ or $p(0) \leq 0$. This gets us to a characterization of IC and IR constraints.

PROPOSITION **1** *A contract (q, p) is incentive compatible and individually rational if and only if*

1. *q is non-decreasing.*
2. *revenue equivalence formula in (4) holds.*
3. *$p(0) \leq 0$.*

Proof: If (q, p) is IC, we have already shown that revenue equivalence formula in (4) holds. The non-decreasing of q is true since for any $\theta > \theta'$, adding the incentive constraints for θ and θ' gives us $\big(q(\theta) - q(\theta')\big)(\theta - \theta') \geq 0$. This gives $q(\theta) \geq q(\theta')$. Also, if IC holds, $p(0) \leq 0$ has been shown to be necessary and sufficient for IR.

To show sufficiency of these conditions for IC, pick θ, θ'. Using revenue equivalence formula

$$U^{q,p}(\theta) - U^{q,p}(\theta') = \int_{\theta'}^{\theta} q(\hat{\theta})d\hat{\theta}.$$

Since q is non-decreasing, the right hand side is greater than or equal to $q(\theta')(\theta - \theta')$, which is the desired incentive constraint. ■

Now, we return to the objective function of the seller. Suppose F is the cdf of types. We assume that F is strictly increasing, differentiable with density f. The seller seeks to maximize the following expression over all contracts:

$$\int_0^1 \Big[p(\theta) - C(q(\theta))\Big] f(\theta)d\theta.$$

Using, revenue equivalence formula (4), we simplify this to

$$\int_0^1 \Big[p(0) + q(\theta)\theta - \int_0^{\theta} q(\theta')d\theta' - C(q(\theta))\Big] f(\theta)d\theta.$$

Since IR implies $p(0) \leq 0$, in any optimal contract, we must therefore have $p(0) = 0$. Hence, the objective function becomes

$$\int_0^1 \Big[q(\theta)\theta - \int_0^{\theta} q(\theta')d\theta' - C(q(\theta))\Big] f(\theta)d\theta.$$

Since this is only a function of q, we only need the constraint that q is non-decreasing. We make a some simplification to this term.

$$\begin{aligned}
&\int_0^1 \Big[q(\theta)\theta - \int_0^{\theta} q(\theta')d\theta' - C(q(\theta))\Big] f(\theta)d\theta \\
&= \int_0^1 \big[q(\theta)\theta - C(q(\theta))\big] f(\theta)d\theta - \int_0^1 \Big(\int_0^{\theta} q(\theta')d\theta'\Big) f(\theta)d\theta \\
&= \int_0^1 \big[q(\theta)\theta - C(q(\theta))\big] f(\theta)d\theta - \int_0^1 \Big(\int_{\theta}^{1} f(\theta')d\theta'\Big) q(\theta)d\theta \\
&= \int_0^1 \big[q(\theta)\theta - C(q(\theta))\big] f(\theta)d\theta - \int_0^1 \big(1 - F(\theta)\big) q(\theta)d\theta \\
&= \int_0^1 \Big(\theta q(\theta) - C(q(\theta)) - \frac{1 - F(\theta)}{f(\theta)} q(\theta)\Big) f(\theta)d\theta.
\end{aligned}$$

Forgetting the fact that q needs to be non-decreasing, we solve this unconstrained objective function. We find the point-wise maximum and that should maximize the overall expression.

Point-wise maximum gives a first order condition for each θ as:

$$\theta - C'(q) - \frac{1 - F(\theta)}{f(\theta)} = 0.$$

Denoting the **virtual value** at θ as $v(\theta) := \theta - \frac{1-F(\theta)}{f(\theta)}$, we se that the optimal quality at type θ must satisfy

$$C'(q(\theta)) = v(\theta).$$

Since C is convex, the objective function at each point θ is concave in q. Hence, this is also a global optimal. However, the optimal solution may not satisfy $q(\theta) \geq 0$. To ensure this, strict concavity implies that if the optimum lies to the left of 0, then under non-negativity constraint, we must have $q(\theta) = 0$ as optimal. So, optimal solution can be described as follows. Let $\hat{q}(\theta)$ be the solution to $C'(\hat{q}(\theta)) = v(\theta)$. Then, the optimal quality contract is: for all θ,

$$q^*(\theta) = \max(0, \hat{q}(\theta))$$

with price

$$p^*(\theta) = \theta q^*(\theta) - \int_0^\theta q^*(\theta')d\theta'.$$

Now, this point-wise optimal solution need not satisfy the fact q is non-decreasing. However, if virtual value is increasing, then it ensures that q is non-decreasing. To see this, assume for contradiction for some $\theta > \theta'$, we have $q(\theta) < q(\theta')$. Then, $q(\theta') > 0$. Further, $\hat{q}(\theta) \leq q(\theta)$ implies $\hat{q}(\theta) < \hat{q}(\theta')$. Then, convexity of C implies $C'(\hat{q}(\theta)) \leq C'(\hat{q}(\theta'))$. But then, $v(\theta) \leq v(\theta')$, which contradicts the fact that v is increasing. Notice that virtual value is increasing can be satisfied if inverse hazard rate $\frac{f(\theta)}{1-F(\theta)}$ is non-decreasing - an assumption satisfied by many distribution including the uniform distribution.

As an exercise, suppose $C(q) = \frac{1}{2}q^2$ with $q \in [0, 1]$ and F is the uniform distribution in $[0, 1]$. Then, we see that for each θ, $v(\theta) = 2\theta - 1$. Hence, $C'(q(\theta)) = q$ must be equal to $2\theta - 1$. Hence, we get $q^*(\theta) = \max(0, 2\theta - 1)$. Notice that in the perfect information case, the seller should ensure $C'(q(\theta)) = \theta$, which gives $q(\theta) = \theta$. So, there is under-provision to lower types due to incentive constraint. This is shown in Figure 1.

2.3 Constant marginal cost

If marginal cost is constant, then the optimal contract exhibits extreme *pooling*. To see this, suppose that q can take any value in $[0, 1]$ and $C(q) = cq$ for some $c > 0$. Then the

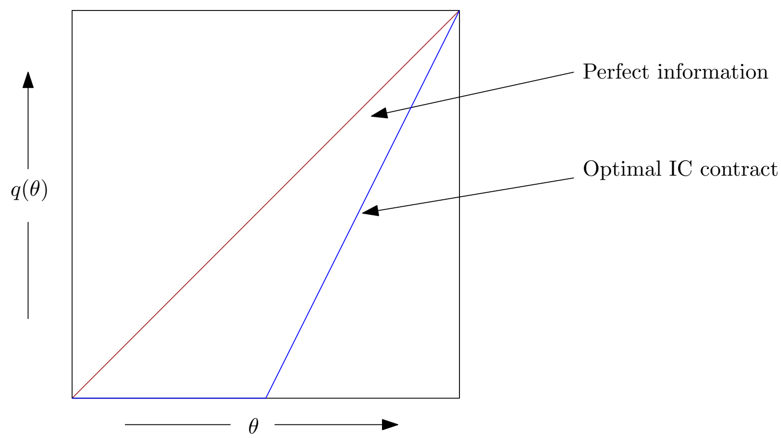

Figure 1: Adverse selection

optimization program is

$$\begin{aligned}
&\max_{q \text{ non-decreasing}} \int_0^1 \Big(\theta q(\theta) - cq(\theta) - \frac{1-F(\theta)}{f(\theta)} q(\theta)\Big) f(\theta) d\theta \\
&= \max_{q \text{ non-decreasing}} \int_0^1 \Big(\theta - c - \frac{1-F(\theta)}{f(\theta)}\Big) q(\theta) f(\theta) d\theta \\
&= \max_{q \text{ non-decreasing}} \int_0^1 \Big(v(\theta) - c\Big) q(\theta) f(\theta) d\theta
\end{aligned}$$

This has a simple optimal solution: whenever $v(\theta) < c$, set $q(\theta) = 0$ and whenever $v(\theta) > c$, set $q(\theta) = 1$. Monotonicity of v ensures monotonicity of q. Notice that if $q(\theta) = 0$, we have $p(\theta) = 0$. By the revenue equivalence formula, if $q(\theta) = 1$ (which implies that $\theta \geq v^{-1}(c)$)

$$p(\theta) = \theta - \int_{v^{-1}(c)}^{\theta} q(\theta')d\theta' = \theta - (\theta - v^{-1}(c)) = v^{-1}(c).$$

Hence, every buyer who gets the maximum possible quality pays the "posted-price" $v^{-1}(c)$. Thus, the optimal contract is equivalent to saying that the seller announces a **posted-price** equal to $v^{-1}(c)$ and the buyer with type greater than the posted price gets maximum quality and those below the posted price get zero quality.

2.4 Non-linear values

In the previous analysis, we assumed that if the agent gets q and pays p, then his utility (with type θ) is $q\theta - p$. However, in many settings, there may be a more general value function u which specifies the value of q given type θ: $u(q, \theta)$. The function u can be assumed to be concave in θ and differentiable sufficient number of times. It is also standard to assume that it is strictly increasing in both arguments and satisfies the **single crossing** condition: for all $q > q'$ and all $\theta > \theta'$, we have

$$u(q, \theta) - u(q, \theta') > u(q', \theta) - u(q', \theta').$$

In other words, $\frac{\partial^2 u(q,\theta)}{\partial q \partial \theta} > 0$.

An analysis similar to above is possible with these assumptions. Now, the point-wise maximization (of the unconstrained problem) will be done of the following function:

$$u(q, \theta) - C(q) - \frac{\partial u(q, \theta)}{\partial \theta} \frac{1 - F(\theta)}{f(\theta)}.$$

The first order condition gives us

$$\frac{\partial u(q, \theta)}{\partial q} - C'(q) - \frac{\partial^2 u(q, \theta)}{\partial \theta \partial q} \frac{1 - F(\theta)}{f(\theta)} = 0.$$

With the single crossing condition, the only missing constraint is monotonicity of q. Again, inverse hazard rate being non-decreasing ensures this.

3 Moral hazard

We now investigate the other principal-agent problem where the hidden feature of the agent is the action he takes. There is no hidden characteristics of the agent in this model. The agent takes some actions which the principal cannot observe. However, the principal observes some signal (say, output) from those actions. Contracts can be written on those signals. The principal would like the agent to take particular actions since it will induce payoffs for him. The objective here is to study what kind of actions can be induced by the principal.

In general, the study of moral hazard is more complicated than that of adverse selection. The optimization program is way more difficult than the adverse selection problem. We will only be studying the tools used to simplify the optimization program and get some insights into properties of the optimal contract in this setting.

3.1 A simple model

A firm (principal) hires a worker (agent) to complete a project. The firm cannot observe the effort level of the agent, which can take values $\{e_L, e_H\}$ with $e_L < e_H$. However, the firm observes the the profit made from the project, which is a signal of the effort put by the agent. Let π denote the level of profit of the project, and let $[\underline{\pi}, \bar{\pi}]$ be the support of this profit level. An important aspect of this profit observation is that, the firm is not able to deduce the effort level from it. Formally, there is a conditional density function $f(\pi|e)$ for each e and for all π. Let $F(\pi|e)$ be the cdf of the conditional distributions.

We will also assume that efforts are "ordered" in a stochastic dominance sense: $F(\pi|e_H) \leq F(\pi|e_L)$ for all π, with strict inequality holding at positive measure of profit levels. A direct implication of this is that the agent derives higher expected profit by exerting high effort than low effort.

The firm can offer a wage contract to the agent. If a wage w is offered and he put effort e, then the profit of the agent is $v(w) - c(e)$, where v is concave, strictly increasing, twice-differentiable, and $c(e_H) > c(e_L)$. The firm receives the profit of the project minus the wage offered to the agent. The agent also has an outside option which gives him a profit of $\underline{u}$.

Formally, a wage contract is a map $w : [\underline{\pi}, \bar{\pi}] \to \mathbb{R}$. Since effort is not observable, this form of the wage contract is appropriate. With the details of the model flushed out, let us start our analysis by looking the first-best.

The first-best solution. If the effort is observable, then the firm can extract any effort it likes from the agent by ensuring his outside option (participation). No incentives are needed. In particular, the firm just solves the following optimization program.

$$\max_{e \in \{e_L, e_H\}, w} \int_{\underline{\pi}}^{\bar{\pi}} \big(\pi - w(\pi)\big) f(\pi|e) d\pi$$

subject to

$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e) d\pi - c(e) \geq \underline{u}.$$

The usual approach to solve this problem is two-stage: fix an effort level and find the optimal wage contract that implements this effort; then compare across effort levels. So, we fix an effort level e and ask what wage contract can ensure participation for e and maximize expected payoff. The expected payoff from a wage contract w at effort level e is

$$\int_{\underline{\pi}}^{\bar{\pi}} \big(\pi - w(\pi)\big) f(\pi|e) d\pi$$

This is equivalent to minimizing the wage bill: $\int_{\underline{\pi}}^{\bar{\pi}} w(\pi)f(\pi|e)d\pi$ subject to participation constraint.

The first step of solving the problem is to observe that the participation constraint must bind. To see this, suppose not:

$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi))f(\pi|e)d\pi > c(e) + \underline{u}.$$

Then, we can define another wage contract $w'(\pi) = w(\pi) - \delta$, where $\delta > 0$ but sufficiently close to zero. Clearly, the expected wage decreases from w to w'. Since v is strictly increasing but continuous, $v(w'(\pi))$ is arbitrarily close to $v(w(\pi))$, and hence, the participation still holds. Hence, any optimal wage contract must satisfy

$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi))f(\pi|e)d\pi = c(e) + \underline{u}.$$

Notice that if v is concave, then Jensen's inequality implies

$$v\Big(\int_{\underline{\pi}}^{\bar{\pi}} w(\pi)f(\pi|e)d\pi\Big) \geq \int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi))f(\pi|e)d\pi = c(e) + \underline{u}.$$

Since v is strictly increasing, we get that

$$\Big(\int_{\underline{\pi}}^{\bar{\pi}} w(\pi)f(\pi|e)d\pi\Big) \geq v^{-1}(c(e) + \underline{u}.$$

This holds for all contracts w. Hence, for every e, the fixed wage contract $w^*(\pi) = v^{-1}(c(e) + \underline{u})$ is optimal contract in the perfect information case.

Note that $v(w^*_{e_H}) > v(w^*_{e_L})$, and hence, $w^*_{e_H} > w^*_{e_L}$. This means that the optimal wage is monotone. So, the firm must make agent choose effort which maximizes

$$\max_{e \in \{e_L, e_H\}} \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e)d\pi - v^{-1}(c(e) + \underline{u}).$$

The second-best solution. Now, the firm does not observe effort of the agent. We follow the same two-step approach for solving the optimal contract. Suppose the firm wants to implement an effort level e. Then, the optimization program it solves is the following:

$$\max_{w} \int_{\underline{\pi}}^{\bar{\pi}} \big(\pi - w(\pi)\big) f(\pi|e) d\pi$$

subject to

$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e) d\pi - c(e) \geq \underline{u}$$
$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e) d\pi - c(e) \geq \int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e') d\pi - c(e') \qquad \forall\, e' \neq e.$$

The new constraint is the **incentive constraint (IC)** to ensure that it is optimal for agent to choose e. We consider the case of implementing each of the effort levels separately.

Implementing e_L. To implement e_L, consider any constant wage contract w. Note that IC holds since $c(e_L) < c(e_H)$. Hence, the only relevant constraint is IR constraint. But the first-best solution is a constant wage contracts $w^*_{e_H}$, which is clearly optimal with the IC constraints. Hence, to implement e_L, the first-best fixed wage contract works.

Lemma **1** *The firm can implement e_L by the first-best wage contract when effort is unobservable.*

Thus, unobservable effort adds no extra wage to the firm if it wishes to implement low effort.

Implementing e_H. For implementing e_H, the incentive constraints matter more. Note that the IC becomes:

$$\int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e_H) d\pi - c(e_H) \geq \int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e_L) d\pi - c(e_L).$$

Now, let λ and μ be the Lagrange multipliers of this optimization. Then, first order condition gives us

$$-f(\pi|e_H) + \lambda v'(w(\pi)) f(\pi|e_H) + \mu v'(w(\pi)) f(\pi|e_H) - \mu v'(w(\pi)) f(\pi|e_L) = 0.$$

This gives us the following necessary condition

$$\frac{1}{v'(w(\pi))} = \lambda + \mu\Big[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}\Big] \;\forall\, \pi.$$

Lemma **2** *Both the IC and IR constraints bind in the optimal solution, i.e., $\lambda > 0, \mu > 0$.*

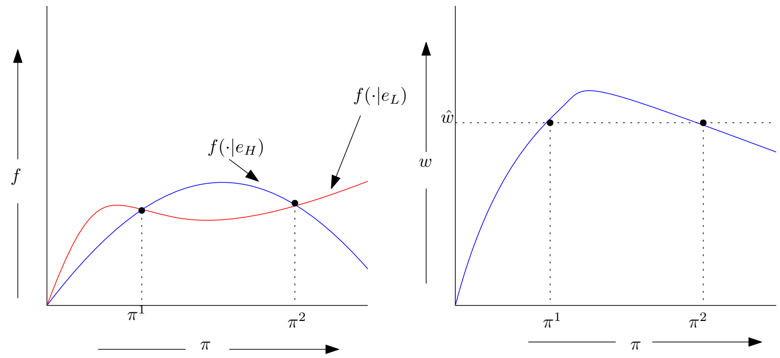

Figure 2: Optimal contract need not be monotone

Proof: If $\mu = 0$, the condition implies a fixed wage contract (as in the first-best case). But under a fixed wage contract, the agent prefers e_L over e_H violating IC.

Now, suppose the IR does not bind. Construct another wage contract $\bar{w}$ such that $v(w(\pi)) - v(\bar{w}(\pi)) = \epsilon > 0$ for all π, i.e. constant change in value. As a result, IC continues to hold. If ϵ is small, $\bar{w}$ differs from w by a small amount at each π. Hence, participation will also hold since it is not binding. But $v(w(\pi)) > v(\bar{w}(\pi))$ implies $w(\pi) > \bar{w}(\pi)$ for all π. Hence, the wage decreases, giving a contradiction to optimality. ■

Lemma 2 throws a surprising conclusion. Let $\lambda = \frac{1}{v'(\hat{w})}$ for some $\hat{w}$. This is possible to define since $\lambda > 0$. Define $\ell(\pi) = \frac{f(\pi|e_L)}{f(\pi|e_H)}$ for all π, called the *likelihood ratio*. Now if $\ell(\pi) > 1$, we see that $\hat{w} > w(\pi)$ and if $\ell(\pi) < 1$, we see that $w(\pi) > \hat{w}$. Unfortunately, $\ell(\pi)$ need not be monotone and the optimal wage can go above and below $\hat{w}$.

Lemma **3** *The optimal wage contract need not be monotone.*

However, it is immediate that if $\ell(\pi)$ is monotone than the wage contract is monotone. This monotone likelihood ratio property is stronger than first order stochastic dominance. As is clear from the analysis, the optimal wage depends where $\ell(\pi)$ crosses 1 (or $f(\pi|e_L)$ and $f(\pi|e_H)$ cross each other). The non-monotonic nature of the optimal contract is shown in Figure 2. Further, it is unlikely that the optimal wage contract is a fixed wage contract or a linear contract (or any simple contract), which was the case when we were implementing e_L.

Exercise. Show that the monotone likelihood ratio property implies first order stochastic dominance.

Finally, we can also infer that the wage bill of the firm is higher in case of non-observable effort.

Lemma 4 *For implementing e_H, the optimal wage contract has higher expected wage in the non-observable effort case than the observable effort case.*

Proof: In the observable effort case the optimal wage is $v^{-1}(c(e_H) + \underline{u})$. Let w be the optimal wage contract for implementing e_H in the non-observable effort case. Assume for contradiction

$$v^{-1}(c(e_H) + \underline{u}) > \int_{\underline{\pi}}^{\bar{\pi}} w(\pi) f(\pi|e_H) d\pi.$$

Since v is strictly increasing

$$\begin{aligned} c(e_H) + \underline{u} &> v\Big(\int_{\underline{\pi}}^{\bar{\pi}} w(\pi) f(\pi|e_H) d\pi\Big) \\ &\geq \int_{\underline{\pi}}^{\bar{\pi}} v(w(\pi)) f(\pi|e_H) d\pi, \end{aligned}$$

where the second inequality is Jensen's inequality for concave functions. But this implies that the IR constraint is violated, a contradiction. ■

Finally, it is not clear that which effort level should the firm decide to implement. We consider two cases.

Case 1. Suppose it is optimal for the firm to implement e_L when effort is observable. We know that first-best wage contract (Lemma 1) continues to be optimal for implementing e_L in the second-best case. However, by Lemma 4, the expected wage for implementing e_H rises in the second-best case. Hence, e_L continues to be optimal for the firm when effort is non-observable.

Case 2. Suppose it is optimal for the firm to implement e_H when effort is observable. By Lemma 4, now the wage increases with unobservable effort. Hence, to induce e_H, the expected wage may be higher or lower than implementing e_L. In general, the wage may increase so much that the firm may decide to offer a fixed wage contract and implement e_L (which is not first-best). But even if the firm finds it optimal to implement e_H, it pays a higher wage. Hence, there is welfare loss in both the cases.

Exercise. Solve the two effort model when the agent is risk neutral.

Exercise. Solve a k-level of effort model when the agent is risk-averse. How much of the 2-effort results carry over to k case?

3.2 A general model

In this section, we develop a general model of moral hazard and introduce the *first order approach*. The objective is to develop a methodology (rather than a solution itself, which is very difficult to describe) to solve moral hazard problems. The notation and terminology is slightly different from the previous section with two effort levels. We will now call the two participating entities principal (firm earlier) and agent. Instead of effort, we will say that the agent takes actions which is unobservable. Instead of profit, we will say that the principal observes a signal. This is consistent with the wide applicability of the model, where actions need not be effort but, for example, a healthy diet, and signal need not be profit but a good level of blood sugar.

A risk averse agent takes an action which is not verifiable by the principal. Let a denote the action, and assume that a belongs to $\mathcal{A} \equiv [\underline{a}, \bar{a}]$. We will denote by $int(\mathcal{A}) := (\underline{a}, \bar{a})$. Agent's action generates a signal (effort) which is observed by the principal. Let x denote the signal and assume that it belongs to $\mathcal{X} \equiv [\underline{x}, \bar{x}]$. If the action taken is a, then it generates a distribution of signals, denoted by the cdf $F(x|a)$. We will assume F to be well-behaved - it has a density function $f(x|a)$, it is continuously differentiable up to requisite degree.

The principal offers a wage contract to the agent. Formally, a **wage contract** is a map $w : \mathcal{X} \to \mathbb{R}$. The agent incurs a value from the wage contract, which is given by the map $v : \mathbb{R} \to \mathbb{R}$, which is assumed to be strictly increasing and thrice differentiable. Since the agent is risk averse, we assume that v is strictly concave, i.e., $v''(\cdot) < 0$. Agent incurs a cost by taking actions, and this is described by a map $c : \mathcal{A} \to \mathbb{R}$. The cost function c is assumed to be strictly increasing, differentiable, and convex. Hence, agent's net utility from a wage contract w when he takes action a and principal observes signal x is given by an additively separable function

$$v(w(x)) - c(a).$$

Since the agent does not know what signal the principal is going to observe when he is deciding on his action, he computes his expected utility from his action. So, agent's expected utility from a wage contract w when takes action a is given by

$$EU_w(a) := \int_{\underline{x}}^{\bar{x}} v(w(x))f(x|a)dx - c(a).$$

The following is the main definition of this section.

Definition 2 *An action* $a^* \in int(\mathcal{A})$ *is* **implementable** *if there exists a wage contract* w *such that*

$$EU_w(a^*) \geq EU_w(a) \ \ \forall \ a \in \mathcal{A}.$$

In this case, we say that w *is a* **globally incentive compatible (GIC)** *for* a^**.*

A necessary condition for w to be GIC for a^* is that the first order condition must hold. This inspires our local incentive compatibility.

Definition 3 *A wage contract* w *is* **locally incentive compatible (LIC)** *for* a^* *if*

$$EU_w'(a^*) = \int_{\underline{x}}^{\bar{x}} v(w(x)) f_a(x|a^*)dx - c'(a^*) = 0.$$

3.3 When does LIC imply GIC?

A critical question in analyzing the moral hazard problem lies in analyzing when LIC implies GIC. Here, we employ a somewhat indirect (but easy) approach. The observation that inspires this approach is the following. LIC requires that EU_w has derivative equal to zero at a^* (stationary point) - in other words, the tangent to EU_w at a^* is flat. A necessary and sufficient condition for GIC is that this tangent is above EU_w at all the points. Hence, if we can create an *auxiliary problem* where the tangent to EU_w is the actual expected utility of the agent and maintain incentive compatibility in that problem, then we are done. Since the tangent is a linearization of EU_w around a^*, checking incentive constraints in the auxiliary problem is probably simpler. This inspires us to look at the following new auxiliary problem.

Let $a^* \in int(\mathcal{A})$ be the action that we seek to implement. To construct the new problem, we first *linearize* the primitives of the problem: the conditional distribution of signals and the cost function. Let $F^L(x|a, a^*)$ be the new "cdf" with "density" $f^L(x|a, a^*)$, where $F^L(x|a, a^*)$ is defined as

$$F^L(x|a, a^*) := F(x|a^*) + (a - a^*)F_a(x|a^*) \ \ \forall \ x \in \mathcal{X}, \ \forall \ a \in \mathcal{A},$$

where F_a denotes the derivative with respect to a. We argue that F^L has many nice properties of a cdf. Notice that for all $a \in \mathcal{A}$ and for all $x \in \mathcal{X}$, we see that

$$f^L(x|a, a^*) = f(x|a^*) + (a - a^*)f_a(x|a^*).$$

Since $F(\underline{x}|a) = 0$ for all a and $F(\bar{x}|a) = 1$ for all a, we get that $F_a(\underline{x}|a) = F_a(\bar{x}|a) = 0$ for all a. Hence, $F^L(\underline{x}|a, a^*) = F(\underline{x}|a^*) = 0$. Similarly, $F^L(\bar{x}|a, a^*) = 1$. However, F^L may not be increasing in x. This is not a problem as we are just constructing an imaginary problem.

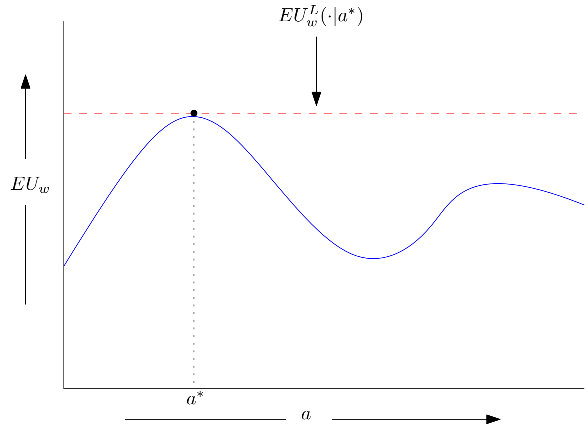

Figure 3: The first order approach

Similarly, let $c^L(a, a^*)$ denotes the new cost function:

$$c^L(a, a^*) := c(a^*) + (a - a^*)c'(a^*) \ \forall \ a \in \mathcal{A}.$$

Now, in this imaginary problem, the expected utility of the agent from contract w by choosing action a is given by

$$\begin{aligned} EU^L_w(a|a^*) &:= \int_{\underline{x}}^{\bar{x}} v(w(x)) f^L(x|a, a^*)dx - c^L(a, a^*). \\ &= \int_{\underline{x}}^{\bar{x}} v(w(x))\big[f(x|a^*) + (a - a^*)f_a(x|a^*)\big]dx - \big[c(a^*) + (a - a^*)c'(a^*)\big]. \\ &= EU_w(a^*) + (a - a^*)\Big[\int_{\underline{x}}^{\bar{x}} v(w(x))f_a(x|a^*)dx - c'(a^*)\Big] \end{aligned}$$

If w is LIC for a^*, the second term in the above expression vanishes. Hence, if w is LIC for a^*, it is GIC if and only if

$$EU^L_w(a|a^*) \geq EU_w(a) \ \forall \ a \in \mathcal{A}.$$

This is shown in Figure 3 - the tangent to EU_w curve at a^* must dominate the curve for it to be a globally optimal solution.

Expanding terms, we get

$$\int_{\underline{x}}^{\bar{x}} v(w(x))f^L(x|a, a^*)dx - c^L(a, a^*) \geq \int_{\underline{x}}^{\bar{x}} v(w(x))f(x|a)dx - c(a).$$

Since c is convex, its tangent at any point always lies below c, i.e., $c^L(a, a^*) \leq c(a)$ for all a. Hence, the above condition holds if

$$\int_{\underline{x}}^{\bar{x}} v(w(x))f^L(x|a, a^*)dx \geq \int_{\underline{x}}^{\bar{x}} v(w(x))f(x|a)dx.$$

So, we have proved a fundamental result on moral hazard which gives a sufficient condition under which a LIC contract becomes GIC.

Proposition **2** *Suppose $a^* \in int(\mathcal{A})$ and w is LIC for a^*. Then, w is GIC for a^* if*

$$\int_{\underline{x}}^{\bar{x}} v(w(x)) f^L(x|a,a^*)dx \geq \int_{\underline{x}}^{\bar{x}} v(w(x)) f(x|a)dx \qquad \forall\ a \in \mathcal{A}. \tag{5}$$

Even though f^L is not a proper density function, Inequality (5) is like comparing a continuum of risky prospects. This is because the stochastic dominance criteria for comparing risky prospects apply even if the “probabilities” are not “distributions”. These intuition are formalized in the following theorem.

Theorem **1** *Suppose $a^* \in int(\mathcal{A})$ and w is LIC for a^*. Then, w is GIC for a^* if w is increasing and for all $a \in \mathcal{A}$ and for all $x \in \mathcal{X}$,*

$$F_{aa}(x|a) \geq 0. \tag{6}$$

Proof: Pick $a^* \in int(\mathcal{A})$ and suppose w is LIC for a^*.

Proof of (1). Suppose w is increasing. Then, v is increasing in x. Since $F_{aa}(x|a) \geq 0$ for all a and for all x, $F(x|a)$ is convex in a for all x. As a result, the tangent line of F at a^* must lie below F for all a, i.e.,

$$F^L(x|a,a^*) \leq F(x|a) \qquad \forall\ a \in \mathcal{A},\ \forall\ x \in \mathcal{X}. \tag{7}$$

Now, we establish Inequality (5) using standard first-order-stochastic-dominance arguments: for all $a \in \mathcal{A}$, note that

$$\begin{aligned}
\int_{\underline{x}}^{\bar{x}} v(w(x)) f^L(x|a,a^*)dx &= \Big[v(w(x))F^L(x|a,a^*)\Big]_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} v'(w(x))F^L(x|a,a^*)dx \\
&\geq v(w(\bar{x})) - \int_{\underline{x}}^{\bar{x}} v'(w(x))F(x|a)dx \qquad \text{(by } v'(w(x)) \geq 0 \text{ and Inequality (7))} \\
&= \Big[v(w(x))F(x|a)\Big]_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} v'(w(x))F(x|a)dx \\
&= \int_{\underline{x}}^{\bar{x}} v(w(x))f(x|a)dx.
\end{aligned}$$

Using Proposition 2, w is GIC for a^*. ■

3.4 The principal's problem: first order approach

Here, we return to principal's problem. The principal wants to maximize his expected payoff given incentive and participation constraint. We will assume that there is an outside option of $\underline{v}$ for the agent, and participation must ensure that expected payoff is at least $\underline{v}$. For incentive constraints, we will only impose LIC - this is called the first order approach (FOA). Of course, FOA is **valid** if LIC implies GIC. In the previous section, we have derived sufficient conditions for that. By defining $B : \mathcal{A} \to \mathbb{R}$ to be the payoff of the principal if action a is chosen, we define the principal's problem is as follows.

$$\max_{w,a} B(a) - \int_{\underline{x}}^{\bar{x}} w(x) f(x|a) dx$$

$$\text{subject to}$$

$$\int_{\underline{x}}^{\bar{x}} v(w(x)) f(x|a) dx - c(a) \geq \underline{v}$$

$$\int_{\underline{x}}^{\bar{x}} v(w(x)) f_a(x|a) dx - c'(a) = 0$$

Definition 4 *The first order approach (FOA) is valid if the solution to the above program is the principal's optimal solution when LIC is replaced by GIC in the above program.*

Here, the participation constraints must bind in the optimal solution - else, the wage can be slightly decreased at all signals, and continuity of v will guarantee that participation holds (GIC holds trivially because wage is constantly decreased for all signals, and hence, LIC holds).

Fix a particular action a. Denote the Lagrangian multipliers of IC and IR constraints be μ and λ respectively. So, the Lagrange of this problem is

$$L(w) = \Big[B(a) - \int_{\underline{x}}^{\bar{x}} w(x) f(x|a) dx\Big] + \lambda\Big[\int_{\underline{x}}^{\bar{x}} v(w(x)) f(x|a) dx - c(a) - \underline{v}\Big] + \mu\Big[\int_{\underline{x}}^{\bar{x}} v(w(x)) f_a(x|a) dx - c'(a)\Big].$$

So, the first order condition (with respect to w) yields

$$f(x|a) - \lambda v'(w(x)) f(x|a) - \mu v'(w(x)) f_a(x|a) = 0.$$

Denoting the ratio of $f_a(x|a)$ and $f(x|a)$ as the **likelihood ratio**:

$$\ell_a(x|a) := \frac{f_a(x|a)}{f(x|a)} \ \forall\ x,$$

we see that the first order condition is for all x,

$$\frac{1}{v'(w(x))} = \lambda + \mu \ell_a(x|a).$$

Now, if $\ell_a(x|a)$ is increasing in x, $v'(w(x))$ is decreasing in x. Since the agent is strictly risk averse, it implies that w is increasing. Combining with Theorem 1, we see that if the optimal action a^* is in the interior, $F_{aa}(x|a) \geq 0$ for all x and for all a, and the monotone likelihood ratio property holds, then the first order approach is valid. We summarize this finding below.

Proposition **3** *Suppose the optimal action for the principal is $a^* \in int(\mathcal{A})$. Further, suppose that $F_{aa}(x|a) \geq 0$ for all x and for all a and $\ell_{ax}(x|a) \geq 0$ for all x and for all a. Then, the first order approach is valid.*

4 Applications of moral hazard

We present various applications of moral hazard. In general, it can be applied to a variety of problems where one side (agent) has risk to share and the other side (principal) cannot observe outcomes and needs to provide incentives.

4.1 Efficiency wage

A firm hires an agent. The agent can exert an effort $e \in \{0, E\}$ - if effort level is e it costs him e. With effort $e = E$, the production level is guaranteed to be Q (with probability 1). With effort $e = 0$, the production level is Q with probability p and 0 with probability $(1-p)$. If production level is is zero, the principal fires the agent and he gets his outside option $\underline{U}$.

The principal does not observe effort but observes production level. It rewards the agent with wage w if production Q is observed and zero otherwise. If the effort was completely observable, then the principal would just offer $\underline{U} + e$ to the agent. Since effort is not observable, he needs to satisfy the incentive constraint of the agent. In particular, if the principal offers a contract of w and wants to implement effort E, then it must be that

$$w - E \geq pw + (1-p)\underline{U} \quad \Leftrightarrow \quad (1-p)(w - \underline{U}) \geq E.$$

So, optimal $w = \underline{U} + \frac{E}{1-p} > \underline{U} + E$.

Now, suppose we change the participation constraint. Instead of $\underline{U}$, we assume that if the production level is zero, then the agent is hired back (by another firm) with probability $(1-\delta_u)$ at wage w and forced to put effort E, but gets zero with probability δ_u. So, effectively, the agent's outside option is now changed to $(1-\delta_u)(w-E)$. So, incentive constraint now becomes

$$w - E \geq pw + (1-p)(1-\delta_u)(w-E).$$

This is equivalent to

$$w \geq E + E\frac{p}{\delta_u(1-p)}.$$

So, a wage minimizing firm must have this constraint binding. The optimal contract should then offer a wage equal to $E\Big(1+\frac{p}{\delta_u(1-p)}\Big)$. Notice that as δ_u (rate of unemployment) increases, wages decrease.

4.2 Moral hazard in teams

Consider a team of two agents. The agents are in a team and put in effort to jointly produce a quantity. The outputs are shared by the agents. In a simple model, we will assume that if e_1 and e_2 are the effort levels, then agents produce $(e_1 + e_2)$, which is shared. Suppose the cost of effort is e^2. Then, in a first-best world, agents should choose e_1, e_2 so as to maximize $e_1 + e_2 - e_1^2 - e_2^2$. This gives us $e_1 = e_2 = \frac{1}{2}$.

Now, suppose the agents cannot observe each other's effort. So, the contract specifies the share of each agent: $s_i : \mathbb{R}_{++} \to \mathbb{R}_+$. Hence, incentive constraint requires that each agent chooses an effort level that maximizes her payoff:

$$\max_{e_i} s_i(Q) - e_i^2.$$

Notice that $s_1(Q) + s_2(Q) = Q = e_1 + e_2$. The first order condition implies that the IC contracts must have $s_i'(Q) = 2e_i$ for each i. So, if the principal wanted to implement first-best effort levels (which were $\frac{1}{2}$), then $s_1'(1) = s_2'(1) = 1$. But $s_1(Q) + s_2(Q) = e_1 + e_2$ implies that $s_1'(Q) + s_2'(Q) = 1$ for all Q. So, first-best cannot be implemented.