ECN 532

Microeconomics II

Hector Chade

Second Quarter 2025

MORAL HAZARD AND INCENTIVE CONTRACTS

Lecture's Objectives

- First topic in information economics: moral hazard
- Fundamental problem in many economic applications
- We will define what moral hazard is and analyze its effect
- Methodology: principal-agent contracting problem
- We will derive the following insights:
  - Crucial difference whether agent's action is observable or not
  - Observable case: pure risk sharing problem, easy contract
  - Unobservable case: moral hazard, incentive contract
  - Moral hazard is costly for principal
  - Moral hazard leads to distortions in action agent takes

The Principal-Agent Problem

Motivation

- Fundamental problem in Economics: moral hazard or hidden action
  - One party to a transaction takes an action that affects the value of the transaction to the other party, who cannot observe the action taken
- Methodology: principal-agent problem with moral hazard
  - The principal offers a contract to the agent, who has a known outside option if she rejects the contract
  - If she accepts the contract, then she takes an action that is unobservable to the principal and costly for the agent, and which affects the principal's payoff
  - The principal can observe a random variable associated with the agent's action, and conditions the contract on that observation
  - Common interpretation of the random variable is output
  - Another interpretation: in insurance, signal is whether agent had an accident
  - Action often interpreted as effort, but could be investment levels, projects, etc.

Historical Remarks

- Historical accounts of insurance numerous, but not of moral hazard
- Insurance literature recognized adverse selection but not moral hazard; first mention is in 1853 in fire insurance
- In the economics literature:
  - Adam Smith (1776): mentions incentive problem that separation of ownership and management entails
  - Marshall (1890): mentions that having insurance may lead to carelessness
  - Haynes (1895): first to use term moral hazard in insurance
  - Knight (1921): describes how moral hazard can limit size of organizations
  - First formal analysis: Arrow (1963) on health care, and Pauly (1968)
  - Current paradigm: Ross (1973), Mirrlees (1975), Shavell (1979), and mainly Holmstrom (1979) and Grossman and Hart (1983)

The Model

- A risk-neutral principal hires a strictly risk-averse agent to perform a task
- Agent's action/effort a can be high, a_h, or low, a_ℓ, unobservable to principal
- Principal offers contract based on a stochastic output x that depends on a
- Principal maximizes expected profits, that is, expected output minus expected compensation to agent
- Agent's utility if wage is w and effort is a is $u(w) - c(a)$, where
  - u twice continuously differentiable with $u' > 0$, $u'' < 0$, unbounded below
  - $0 = c(a_\ell) < c(a_h)$
- If the agent does not work for the principal, she can obtain a level of utility $\bar{u}$ in her next best alternative. We call $\bar{u}$ the agent's reservation utility
- Agent maximizes expected utility using $u(w) - c(a)$

The Model

- Output x can take value $x_1, x_2, ..., x_N$ with probability (given a) $\mathbb{P}[x = x_i|a] = \pi_i(a) > 0$, $i = 1, 2, ..., N$
  - Agent controls with her action a production function that is subject to shocks
- Principal offers a contract: compensation scheme (wages) $w_1, w_2, ..., w_N$, and a recommended action or effort level a
- After observing contract the agent accepts or rejects it
  - If she rejects the game ends and the agent obtains $\bar{u}$
  - If she accepts then the agent privately chooses a, a realization x obtains, the corresponding wage is paid, and the game ends

The Model

- The principal's contracting problem is

$$\max_{w_1,w_2,\dots,w_N,a} \sum_{i=1}^{N}(x_i - w_i)\pi_i(a)$$

$$s.t. \;\; \sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) \geq \bar{u} \quad (P)$$

$$a \text{ solves } \max_{a' \in \{a_\ell . a_h\}} \sum_{i=1}^{N} u(w_i)\pi_i(a') - c(a') \quad (IC)$$

- (P) is the participation constraint, (IC) the incentive constraint
- Remarks:
  - Optimization problem of the principal depends on the optimization of the agent at another level
  - Similar to Stackelberg problem (leader-follower)
  - In fact, solution here is the outcome of the SPE of principal-agent game

The Model

- Questions we will address:
  - What determines the wages $w_1, w_2, ..., w_N$?
    - The information contained in the output about the action taken by the agent
  - Is $w_1 \leq w_2 \leq \cdots \leq w_N$? (Higher output, higher wage)
    - We will provide sufficient conditions that ensure this property
  - Is the action a implemented distorted away from efficiency?
    - The answer is yes, so moral hazard has real effects on the allocation of resources
  - What if in addition to x another signal y could be observed at no cost?
    - We will discuss why it is optimal to introduce y into the contract if informative

Methodology to Solve the Problem

Two-Step Methodology

- The following methodology is similar to what we do in cost and production:
  - Step 1: Cost Minimization. For each $a \in \{a_\ell, a_h\}$, solve

$$\min_{w_1, w_2, \ldots, w_N} \sum_{i=1}^{N} w_i \pi_i(a)$$

    subject to (P) and (IC). Denote the value $C(a)$ (principal's cost function)

  - Step 2: Profit Maximization. Let $B(a) = \sum_{i=1}^{N} x_i \pi_i(a)$ be the expected value of output given action a, and solve

$$\max\{B(a_\ell) - C(a_\ell), B(a_h) - C(a_h)\}$$

- At the end of the second step we have the optimal contract, that is, the compensation scheme and the recommended effort
- We will see that most of the insights come from Step 1

Observable Action Case

Observable Action Case

- As a benchmark, assume a is observable, so there is no moral hazard
- A contract is now a compensation scheme $w_1, w_2, ..., w_N$ and an enforceable action a (since a is observable, its level can be contractually enforced)
- The principal's problem in this case is

$$\max_{w_1, w_2, ..., w_N, a} \sum_{i=1}^{N} (x_i - w_i)\pi_i(a)$$

$$s.t. \quad \sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) \geq \bar{u} \quad (P)$$

- Unlike full problem with moral hazard, constraint (IC) is absent (why?)

Observable Action Case

- First step: principal solves, for each $a \in \{a_\ell, a_h\}$

$$\min_{w_1, w_2, \dots, w_N} \sum_{i=1}^{N} w_i \pi_i(a)$$

$$s.t. \; \sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) \geq \bar{u} \quad (P)$$

- We will see that the solution is very simple
- We will first derive it without calculus and then with calculus
- Let us show that for each a, constraint (P) holds with equality
  - Assume at the optimum $\sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) > \bar{u}$
  - Let $\delta > 0$ be the amount on the left side of the inequality
  - Let $\hat{w}_i$ be $\hat{w}_i = w_i - \eta_i$ for all x_i, where $\eta_i > 0$ solves $u(w_i - \eta_i) = u(w_i) - \delta$
  - $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_N$ is strictly cheaper for principal, contradicting $w_1, \dots, w_N$ optimal

Observable Action Case

- Second, let us show that w_i solves $u(w_i) - c(a) = \bar{u}$ for each i, and so, $w_i = u^{-1}(\bar{u} + c(a))$ for every i, a constant wage
  - If w_i is optimal but is not constant for every i, then the principal can instead offer the constant wage $\hat{w} = \sum_{i=1}^N w_i \pi_i(a)$ and obtain same cost
  - But with $\hat{w} = \sum_{i=1}^N w_i \pi_i(a)$, constraint (P) holds with strict inequality, since

$$u\left(\sum_{i=1}^N w_i \pi_i(a)\right) > \sum_{i=1}^N u(w_i)\pi_i(a) = \bar{u} + c(a)$$

    where the strict inequality follows since for any strictly concave function g, $g(\mathbb{E}[X]) > \mathbb{E}[g(X)]$ (Jensen's inequality), and the equality follows because we are assuming that $(w_1, w_2, ..., w_N)$ is optimal, and thus (P) is binding
  - It follows now that the principal can do strictly better (reduce cost) with another constant wage that is smaller, namely, $\tilde{w} = \sum_{i=1}^N w_i \pi_i(a) - \varepsilon$, where $\varepsilon > 0$ uniquely solve $u(\sum_{i=1}^N w_i \pi_i(a) - \varepsilon) = \bar{u} + c(a)$
  - Thus, $\tilde{w} = u^{-1}(\bar{u} + c(a))$ strictly reduces expected cost, and thus the original contract cannot be optimal

Observable Action Case

- That the wage is constant is not surprising, since without moral hazard this is just a risk sharing problem between a risk neutral party and a strictly risk averse one, and full insurance obtains
- The value of the problem for each action a is $C^o(a) = u^{-1}(\bar{u} + c(a))$
- Note that C^o is strictly increasing in $\bar{u}$ and a
- In short, in the first step we have shown the following (recall $c(a_\ell) = 0$) :
  - If principal wants the agent to take $a = a_\ell$, then offer a contract that says "I will pay you $\tilde{w} = u^{-1}(\bar{u})$ and you will have to exert effort level a_ℓ"
  - If principal wants the agent to take $a = a_h$, then offer a contract that says "I will pay you $\tilde{w} = u^{-1}(\bar{u} + c(a_h))$ and you will have to exert effort level a_h"
  - In each case the agent is willing to accept the contract (why?)
- In the second step, principal decides whether to induce a_ℓ or a_h depending on $\max\{B(a_\ell) - C^o(a_\ell), B(a_h) - C^o(a_h)\}$

Observable Action Case

- As a consistency check, let us re-derive the optimal contract using calculus
- Form the Lagrangian with a multiplier λ

$$\mathcal{L} = \sum_{i=1}^{N} w_i \pi_i(a) + \lambda \left(\bar{u} - \sum_{i=1}^{N} u(w_i)\pi_i(a) + c(a) \right)$$

- The FOC with respect to w_i, $i = 1, 2, ..., N$ is

$$\pi_i(a) - \lambda u'(w_i)\pi_i(a) = 0 \ \Rightarrow \ \frac{1}{u'(w_i)} = \lambda, \ i = 1, 2, ..., N$$

- We immediately obtain the following results:
  - $u' > 0$ implies $\lambda > 0$, which implies that (P) must hold with equality
  - Since for every i, $\frac{1}{u'(w_i)}$ is equal to the same constant, w_i is constant for all i
  - But since (P) holds with equality and w_i is constant in i (pays the same amount for every x_i observed, it follows that the only candidate for an optimal compensation scheme given a is $u(\tilde{w}) = \bar{u} + c(a)$ or $\tilde{w} = u^{-1}(\bar{u} + c(a))$
- The rest of the analysis is now the same

Unobservable Action (Moral Hazard) Case

Unobservable Action Case

- Henceforth action is unobservable, that is, there is moral hazard
- Recall that $a \in \{a_\ell, a_h\}$, $a_\ell < a_h$, with $c(a_\ell) = 0 < c(a_h)$, and that the principal's problem is given by

$$\max_{w_1, w_2, \ldots, w_N, a} \sum_{i=1}^{N} (x_i - w_i)\pi_i(a)$$

$$s.t. \quad \sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) \geq \bar{u} \quad (P)$$

$$a \text{ solves } \max_{a' \in \{a_\ell . a_h\}} \sum_{i=1}^{N} u(w_i)\pi_i(a') - c(a') \quad (IC)$$

- We are going to apply our two-step methodology to solve it

Unobservable Action Case

- Let us solve the cost-minimization problem for each a
- Cost-minimization problem trivial for $a = a_\ell$:
  - If principal pays constant wage, then (IC) constraint holds with strict inequality (why?)
  - Thus principal can offer same wage from observable case
  - It follows that $C(a_\ell) = C^o(a_\ell) = u^{-1}(\bar{u})$

Unobservable Action Case

- Things are much more difficult but more interesting for $a = a_h$
- Principal solves

$$\min_{w_1, w_2, \ldots, w_N} \sum_{i=1}^{N} w_i \pi_i(a)$$

$$s.t. \quad \sum_{i=1}^{N} u(w_i)\pi_i(a) - c(a) \geq \bar{u} \quad (P)$$

$$\sum_{i=1}^{N} u(w_i)\pi_i(a_h) - c(a_h) \geq \sum_{i=1}^{N} u(w_i)\pi_i(a_\ell) \quad (IC)$$

- Fundamental trade-off in moral hazard: risk sharing versus incentives
  - For risk sharing, principal prefers constant wage, but this violates (IC)
  - For incentives, variability in wages is needed, which is costly to the principal

Unobservable Action Case

- Let $\lambda \geq 0$ be the multiplier of (P) and $\mu \geq 0$ of (IC)
  - Fact: If a multiplier is strictly positive the constraint holds with equality
- The Lagrangian of the problem is

$$\mathcal{L} = \sum_{i=1}^{N} w_i \pi_i(a) + \lambda \left(\bar{u} - \sum_{i=1}^{N} u(w_i) \pi_i(a) + c(a) \right)$$
$$+ \mu \left(\sum_{i=1}^{N} u(w_i) \pi_i(a_\ell) - \left(\sum_{i=1}^{N} u(w_i) \pi_i(a_h) - c(a_h) \right) \right)$$

- The FOC with respect to w_i, $i = 1, 2, ..., N$, is (check)

$$\frac{1}{u'(w_i)} = \lambda + \mu \left(1 - \frac{\pi_i(a_\ell)}{\pi_i(a_h)} \right), \quad i = 1, 2, ..., N$$

- Note that $\lambda > 0$ and $\mu > 0$ at the optimum
  - Multiplying by $\pi_i(a_h)$ and adding yield $\lambda = \sum_{i=1}^{N} \frac{1}{u'(w_i)} \pi_i(a_h) > 0$
  - If $\mu = 0$, then $\frac{1}{u'(w_i)} = \lambda$ for all i, and w_i is constant in i, violating (IC)

Unobservable Action Case

- There is an important economic insight emerging from FOC
$$\frac{1}{u'(w_i)} = \lambda + \mu\left(1 - \frac{\pi_i(a_\ell)}{\pi_i(a_h)}\right), \;\; i = 1, 2, ..., N$$
- Interpretation:
  - Wages vary in x_i now, unlike observable case
  - w_i paid when output is x_i is pinned down by the likelihood ratio $\frac{\pi_i(a_h)}{\pi_i(a_\ell)}$
  - Likelihood ratio: information x_i contains about agent taking a_h instead of a_ℓ
  - Optimal compensation pays more when x_i observed has a higher $\frac{\pi_i(a_h)}{\pi_i(a_\ell)}$
  - If $\frac{\pi_i(a_h)}{\pi_i(a_\ell)}$ increases in i (called monotone likelihood ratio property, or MLRP), then agent gets a higher wage when the output observed is higher (something we commonly observe)
- $C(a_h) > C^o(a_h)$ since wages vary: moral hazard strictly increases costs
  - To impose variability on agent's compensation, that is, risk, principal must on average pay the agent more for her to bear the risk

Unobservable Action Case

- The second step is trivial: choose a_ℓ or a_h according to

$$\max\{B(a_\ell) - C(a_\ell), B(a_h) - C(a_h)\}$$

- Since $C(a_\ell) = C^o(a_\ell)$ and $C(a_h) > C^o(a_h)$, the action a_h is implemented "less often" under moral hazard

- Cases with observable action with a_h optimal but under moral hazard a_ℓ is

- Moral hazard distorts the optimal action downward in this case

Unobservable Action Case

- Example
  - Assume $u(w) = \sqrt{w}$ and $N = 2$ (two output levels)
  - To simplify notation, let $\pi_2(a_h) = p$, and $\pi_2(a_\ell) = q$, with $p > q > 0$
  - Note that in this case if $v = u(w)$, then $u^{-1}(v) = v^2$
  - Observable case is easy: $C^o(a_\ell) = (\bar{u})^2$ and $C^o(a_h) = (\bar{u} + c(a_h))^2$
  - Since $B(a_\ell) = qx_2 + (1-q)x_1$ and $B(a_h) = px_2 + (1-p)x_1$, principal chooses to implement a_ℓ or a_h depending on $\max\{B(a_\ell) - C^o(\bar{u}), B(a_h) - C^o(a_h)\}$
  - In the moral hazard case, we know that $C(a_\ell) = C^o(a_h) = (\bar{u})^2$, since it is enough to pay a constant wage $w = (\bar{u})^2$

Unobservable Action Case

- Continuation of the example:
  - The cost-minimizing contract when the principal wants to induce the agent to exert high effort a_h solves

$$\min_{w_1, w_2} (1-p)w_1 + pw_2$$

$$s.t.\ \ (1-p)\sqrt{w_1} + p\sqrt{w_2} - c(a_h) \geq \bar{u}$$

$$(1-p)\sqrt{w_1} + p\sqrt{w_2} - c(a_h) \geq (1-q)\sqrt{w_1} + q\sqrt{w_2}$$

  - When $N = 2$, solving this problem is easy since we know from above that $\lambda > 0$ and $\mu > 0$ and thus both constraints hold with equality by the "fact"
  - Let $v_1 = \sqrt{w_1}$ and $v_2 = \sqrt{w_2}$. Then (P)–(IC) with equality become (check)

$$(1-p)v_1 + pv_2 = \bar{u} + c(a_h)$$

$$-(p-q)v_1 + (p-q)v_2 = c(a_h)$$

  - These are two linear equations in two unknowns, v_1 and v_2
  - We can solve them and then invert to obtain w_1 and w_2

Unobservable Action Case

- Continuation of the example:
  - Solving the system yields (check)

$$v_1 = \bar{u} + c(a_h) - p\frac{c}{p-q}$$
$$v_2 = \bar{u} + c(a_h) + (1-p)\frac{c}{p-q}$$

  - Note both utility levels cover outside option and disutility of effort
  - But with moral hazard v_1 is lowered by $-p\frac{c}{p-q}$ (a penalty), and v_2 is raised by $(1-p)\frac{c}{p-q}$ (a bonus)
  - Since $v_i = \sqrt{w_i}$, we have $w_i = v_i^2$ and thus wages and $C(a_h)$ are

$$w_1 = \left(\bar{u} + c(a_h) - p\frac{c}{p-q}\right)^2$$
$$w_2 = \left(\bar{u} + c(a_h) + (1-p)\frac{c}{p-q}\right)^2$$
$$C(a_h) = (1-p)\left(\bar{u} + c(a_h) - p\frac{c}{p-q}\right)^2 + p\left(\bar{u} + c(a_h) + (1-p)\frac{c}{p-q}\right)^2$$

Unobservable Action Case

- Continuation of the example:
  - Finally, the optimal contract solves $\max\{B(a_\ell) - C(a_\ell), B(a_h) - C(a_h)\}$
  - If $B(a_\ell) - C(a_\ell) > B(a_h) - C(a_h)$, then moral hazard is severe enough that principal prefers to implement a_ℓ at a constant wage
  - If $B(a_h) - C(a_h) \geq B(a_\ell) - C(a_\ell)$, then the principal implements a_h with the contract derived above

Real-World Examples

- Incentive contracts are ubiquitous, and at all levels of the organization:
  - CEO compensation: many items tied to profits of the firm
  - Middle and lower level managers: bonuses tied to performance
  - Salespeople: commissions depend on level of sales
  - Factory workers: piece-rate contracts provide incentives to produce
- Moral hazard is also pervasive in insurance markets
  - Policies with less than full insurance (e.g., deductibles) can be explained by moral hazard: e.g., a driver has more incentives to drive carefully if they are responsible for a significant fraction of the repair
  - Similarly with the threat of increase in future premium upon an accident

Real-World Examples

CEO PAY | FIGURE C

Stock-related components of CEO compensation constitute a large share of total compensation, 2024

CEO compensation, by components, 2024 (millions)

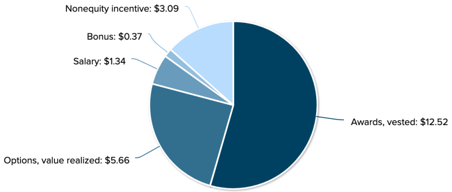

Chart | Data

Economic Policy Institut

Notes: Average annual compensation for CEOs at the top 350 U.S. firms ranked by sales is measured in two ways. Both include salary, bonus, and long-term incentive payouts, but the "granted" measure includes the value of stock options and stock awards when they were granted, whereas the "realized" measure captures the value of stock-related components that accrues after options or stock awards are granted by including "stock options exercised" and "vested stock awards." Projected value for 2024 is based on the percent change in CEO pay in the sample available in June 2023 and in August 2024 applied to the full-year 2023 value.

Source: Authors' analysis of data from Compustat's ExecuComp database.

Additional Signals: Sufficient Statistic Result

- In many settings under moral hazard, in addition to x, the principal can observe, at no cost, another signal y
- Let y take values $y_1 < y_2 < \cdots < y_K$ with some probability distribution
- The main question is the following:
  - Should principal condition compensation on realization of x and y or just x?
- The answer is simple and intuitive (sufficient statistic result):
  - If y provides information about the agent's action, then it should be included
  - So wage paid should be $w(x, y)$, e.g., if $x = x_i$ and $y = y_j$, then wage is w_{ij}
  - By "information" we mean likelihood ratio $\frac{\pi_{ij}(a_h)}{\pi_{ij}(a_\ell)}$ is not independent of j
- This result has many applications
  - CEO and stock options with industry-index strike price; salespeople and information about state of demand; relative performance evaluation
  - In all these cases, there is an additional signal that is used in the contract