5-4-22 Lecture 08

yesterday

- Limited Liability (& risk neutral agent)
- Risk-averse agent step ① (optimal wage rule $w_e(\cdot)$ for any given effort e.)
  - for lowest effort: constant wage st. IR (holds)
  - for $e = e_H$, wage is increasing in likelihood ratio $\frac{f(\pi \mid e_H)}{f(\pi \mid e_L)}$

today

- more than two possible efforts in step ①
- Step ②: How to determine optimal effort level?
- Moral hazard with multiple agents

Moral Hazard – Risk sharing – multiple effort levels

Suppose now there are n effort levels $E = \{e_1, e_2, \ldots, e_n\}$ with cost $c_1 < c_2 < \cdots < c_n$

Optimal wage rule $w(\pi)$ to implement e_i?

- same techniques as with $n = 2$ can be applied
- but now we have $n-1$ incentive constraints: (IC_{e_i,e_k}) for all $k \neq i$

The optimal wage to implement e_i satisfies

$$\frac{1}{v'\left(w_{e_i}(\pi)\right)} = \gamma + \sum_{\substack{k=1 \\ k\neq i}}^{n} \mu_k^i \left[1 - \frac{f(\pi|e_k)}{f(\pi|e_i)}\right]$$

- If we interpret μ_n^i as the shadow cost of constraint (IC_{e_i,e_n}), then the agent is rewarded most for outputs π with high $\frac{f(\pi|e_i) - f(\pi|e_n)}{f(\pi|e_i)}$ for those k where incentive to deviate to e_n is strongest.
- Compliction: ex ante we have no idea which constraints bind (i.e. which $\mu_n^i \neq 0$).

Step ②: Optimal effort level (principal's second best)

- From step ① we get optimal wage function $w_e(\cdot)$ for each e.
- Let $W(e) = \int_{\underline{\pi}}^{\overline{\pi}} w_e(\pi) f(\pi | e) \, d\pi$ be the expected wage paid to agent to implement effort level e (= cost of principal)
- compare $\int_{\underline{\pi}}^{\overline{\pi}} \pi f(\pi | e) \, d\pi - W(e)$ across e

choose maximiser (if net profit is ≥ 0).

Moral Hazard – Example

Exercise:

- 3 effort levels $e \in \{e_L, e_M, e_H\}$
- 2 output levels $\pi \in \{\underline{\pi}, \bar{\pi}\}$
- $\mathbb{P}[\pi = \bar{\pi} \mid e = e_i] = f_i$ with $0 \leq f_L < f_M < f_H$ $\Rightarrow$ $\Pr[\pi = \underline{\pi} \mid e_i] = 1 - f_i$
- cost $0 \leq c_L < c_M < c_H$

Q:

Find conditions on the π, f, c such that e_M cannot be implemented in any contract.

↳ hint 1: no need to know utility function $v(\cdot)$. write $\bar{u} = v(w(\bar{\pi}))$

$\underline{u} = v(w(\underline{\pi}))$

hint 2: values $\underline{\pi}$ and $\bar{\pi}$ don't matter.

- To make agent choose e_M, we need to satisfy the IC constraints:

$(IC_{e_M, e_L}:)$ $\quad (1-f_M)\underline{u} + f_M \bar{u} - c_M \geq (1-f_L)\underline{u} + f_L \bar{u} - c_L$

$(IC_{e_M, e_H}:)$ $\quad (1-f_M)\underline{u} + f_M \bar{u} - c_M \geq (1-f_H)\underline{u} + f_H \bar{u} - c_H$

- Rewrite $(IC_{e_M, e_L}:)$ $\quad \cancel{\underline{u}} + \underline{f_M}(\bar{u} - \underline{u}) - c_M \geq \cancel{\underline{u}} + \underline{f_L}(\bar{u} - \underline{u}) - c_L$

$\iff (f_M - f_L)(\bar{u} - \underline{u}) \geq c_M - c_L \iff \boxed{\bar{u} - \underline{u} \geq \frac{c_M - c_L}{f_M - f_L}}$

The difference in expected utility $\geq$ difference in effort cost

- Rewrite $(IC_{e_M, e_H}):$ $\quad \cancel{\underline{u}} + \underline{f_M}(\bar{u} - \underline{u}) - c_M \geq \cancel{\underline{u}} + \underline{f_H}(\bar{u} - \underline{u}) - c_H$

$\iff c_H - c_M \geq (\bar{u} - \underline{u})(f_H - f_M) \iff \boxed{\bar{u} - \underline{u} \leq \frac{c_H - c_M}{f_H - f_M}}$

If $\dfrac{c_M - c_L}{f_M - f_L} > \dfrac{c_H - c_M}{f_H - f_M}$, then effort e_M cannot be implemented by any wage pair $(w(\bar{I}), w(\bar{F}))$

Intuitively, the utility difference that makes agent prefer e_M to e_L also makes agent prefer e_H to e_M when this condition holds.

Moral Hazard

Multiple Agents (a glimpse)

Moral Hazard – Multiple Agents

If principal interacts with multiple agents, organisation design matters

- should principal foster competition or collaboration?
- when tasks are substitutes, agents may want to free-ride on others' effort
- if tasks are complements, multiple equilibria may arise in agents' game
- agents may collude
- ...

We will only consider one specific example model that asks
'(when) should ex-ante symmetric agents be rewarded differently for same outcome?'

Moral Hazard – Multiple Agents

Economic policy questions often contain a tradeoff between equality and efficiency

- rewards for qualification attract more skilled employees
- rewards for good performance foster incentives
- ...

However, most would agree that **favouritism** (treating identical agents unequally) is bad

Winter (2004): Incentives and Discrimination. *Am Econ Review.* presents possible tension: discrimination may be effective to coordinate agents on the right actions

Moral Hazard – Multiple Agents

Model:

- 2 agents $i = 1, 2$ work on joint project
- each agent has a task and privately chooses $e_i \in \{0, 1\}$ with effort cost $c > 0$ $e_i \cdot c$
- task i ends successfully with probability $\begin{cases} 1 & \text{if } e_i = 1 \\ \alpha \in (0,1) & \text{if } e_i = 0 \end{cases}$
- project is successful only if both tasks end successfully
- principal can only observe joint project success (not individual tasks)
- agent i gets reward w_i if project is successful; 0 otherwise
  (w_1, w_2)

1: optimal rewards for **project** success s.t. $(e_1, e_2) = (1, 1)$ is **a** Nash equilibrium?

2: optimal rewards for **project** success s.t. $(e_1, e_2) = (1, 1)$ is **unique** Nash equilibrium?

1. When is $e_1 = 1$ a best response to $e_2 = 1$

$$1 \cdot 1 \cdot w_1 - c \geq 1 \cdot \alpha \cdot w_1 - 0 \iff w_1 \geq \frac{c}{1-\alpha}$$

↳ if $e_1 = 1$

↳ probability that $i=1$ succeeds if $e_1 = 0$

↳ probability that $i=2$ succeeds at task 2 if $e_2 = 1$

The cheapest rewards for the principal to make sure that $(e_1, e_2) = (1,1)$ is a NE are

$$w_1 = w_2 = \frac{c}{1-\alpha}$$ → but then $(e_1, e_2) = (0,0)$ is also a NE ! !!

2. How to make $(e_1, e_2) = (1,1)$ the unique NE ?

↳ we make sure that $e_1 = 1$ is also the best response to $e_2 = 0$

$$\alpha \cdot 1 \cdot w_1 - c \geq \alpha \cdot \alpha \cdot w_1 - 0 \iff \boxed{w_1 \geq \frac{c}{\alpha(1-\alpha)}}$$

Note: it is sufficient to make $e_i = 1$ dominant for one agent. The other agent then knows that $e_i = 1$ and responds with $e_j = 1$ iff $\boxed{w_j \geq \frac{c}{1-\alpha}}$

$(w_1, w_2) = \left(\frac{c}{\alpha(1-\alpha)}, \frac{c}{1-\alpha}\right)$ or $(w_1, w_2) = \left(\frac{c}{1-\alpha}, \frac{c}{\alpha(1-\alpha)}\right)$ are optimal rewards to make $(e_1, e_2) = (1,1)$ the unique NE.

$\Rightarrow$ It is optimal to treat agents differently for same performance to rule out bad equilibria.

Result in the paper: discrimination is optimal if and only if tasks are complements

Since some of you have expressed interest in this model, here are some (recently published) papers that are based on this main idea:

— Halac, Lipnowski, Rappoport (2021, AER): Rank uncertainty in organizations (this one deals with the information about wages of others that I mentioned)
— Halac, Kremer, Winter (2020 AER): Raising capital from heterogeneous investors
— Bernstein, Winter (2012 AEJ:Micro): Contracting with heterogeneous externalities
— Winter (2010 RAND): Transparency and incentives among peers

Don‘t hesitate to reach out if you want to talk about potential research ideas related to this (or any other topic connected to the course)