Moral Hazard

Moral Hazard – Intro

We now consider models with 'hidden action'

- Principal commits to payment schedule
- Agent takes an action
- Principal observes (imperfect) signal about action and pays according to schedule

Examples:

- Insurance contract
- Employment contract
- Rental contract

Moral Hazard – Intro

General setup:

- Agent chooses effort level $e \in E \subset \mathbb{R}_+$
- The profit is $\pi \in [\underline{\pi}, \bar{\pi}] = \Pi \subset \mathbb{R}$.
- Distribution of profit depends on effort: $\pi \sim F(\cdot|e)$ with density $f(\pi|e) > 0$

  $F(\cdot|e)$ is ordered by first-order stochastic dominance:

  $$\text{If } e'' > e', \qquad \text{then } F(\pi|e'') \leq F(\pi|e') \qquad \forall \pi \in \Pi.$$

- Principal observes only π and commits to pay the agent a wage $w(\pi)$
- **Payoffs:** agent: $v(w) - c(e)$ principal: $\pi - w$

  $v(\cdot)$ is increasing and concave, effort cost $c(\cdot)$ is increasing and convex

Moral Hazard

The principal's problem is to choose an effort level e and wage scheme $w(\cdot)$ to solve

$$\max_{e,w(\cdot)} \int_{\underline{\pi}}^{\bar{\pi}} (\pi - w(\pi))\, f(\pi|e)\, \mathrm{d}\pi \quad \text{such that}$$

$$e = \operatorname*{argmax}_{e' \in E} \int_{\underline{\pi}}^{\bar{\pi}} v\,(w(\pi))\, f(\pi|e')\, \mathrm{d}\pi - c(e'), \tag{IC}$$

$$\int_{\underline{\pi}}^{\bar{\pi}} v\,(w(\pi))\, f(\pi|e)\, \mathrm{d}\pi - c(e) \geq 0 \tag{IR}$$

Moral Hazard – Observable effort

Suppose effort is observable to the principal (and contractible) so that wage is $w(e,\pi)$

- The principal can enforce effort e by setting $w(e',\pi) = \begin{cases} 0 & \text{for all } e' \neq e \\ w_e(\pi) & \text{for } e' = e \end{cases}$

Principal's problem is to choose effort level e and wage function $w_e(\cdot)$ to

$$\max_{e, w_e(\cdot)} \int_{\underline{\pi}}^{\bar{\pi}} (\pi - w_e(\pi))\, f(\pi|e)\, \mathrm{d}\pi$$

$$\text{such that} \int_{\underline{\pi}}^{\bar{\pi}} v\left(w_e(\pi)\right) f(\pi|e)\, \mathrm{d}\pi - c(e) \geq 0 \qquad \text{(IR)}$$

Moral Hazard – Observable effort

Moral Hazard – Risk-neutral agent

Suppose effort is unobservable but the agent is risk neutral: $v(w) = w$

- The principal could simply 'sell the firm' to the agent:

$$w(\pi) = \pi - \max_e \left\{ \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e)\, \mathrm{d}\pi - c(e) \right\} = \pi - \left(\int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e^{FB})\, \mathrm{d}\pi - c(e^{FB}) \right)$$

- The agent's expected payoff when choosing some e' is

$$\int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e')\, \mathrm{d}\pi - c(e') - \left(\int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e^{FB})\, \mathrm{d}\pi - c(e^{FB}) \right)$$

Moral Hazard – Risk-neutral agent

We saw: if the agent is risk-neutral and wages are unrestricted, we get efficient outcome

- Agent chooses first-best effort level e^{FB}
- Principal extracts all surplus by 'selling' the firm at expected value

Next, we consider frictions that may induce inefficiencies:

1. Limited Liability:

   if agent has limited funds, requiring $w(\theta) \geq \underline{w}$, selling the firm infeasible

2. Risk-averse agent:

   if $v(w)$ is concave, tradeoff between incentive-provision and risk-sharing arises

Moral Hazard – Limited Liability (and risk-neutral agent)

Suppose

- continuum of effort choices $e \in E = [0, 1]$
- two possible outputs $\pi \in \{0, \bar{\pi}\}$ with $\mathbb{P}[\pi = \bar{\pi} | e] = e$ and $\mathbb{P}[\pi = 0 | e] = 1 - e$
- risk-neutral agent: $v(w) = w$
- **limited liability:** $w(\pi) \geq 0$

The principal solves

$$\max_{e, w(0), w(\bar{\pi})} \{e\,(\bar{\pi} - w(\bar{\pi})) + (1-e)\,(0 - w(0))\} \quad \text{such that}$$

$$e \in \operatorname*{argmax}_{e'} \{e' w(\bar{\pi}) + (1-e') w(0) - c(e')\} \tag{IC}$$

$$e\ w(\bar{\pi}) + (1-e) w(0) - c(e) \geq 0, \tag{IR}$$

$$w(\bar{\pi}) \geq 0, \quad w(0) \geq 0 \tag{LL}$$

Moral Hazard – Limited Liability (and risk-neutral agent)

Moral Hazard – Risk sharing

We saw: principal can implement efficient effort level and extract all expected surplus if

- agent is risk neutral, $v(w) = w$ **and** wages are unrestricted.

In optimal contract, principal 'sells the firm' to agent:

- principal gets $\mathbb{E}_{\pi \sim F(\cdot|e)}\left[\pi | \, e = e^{FB}\right] - c(e^{FB})$ independent of realised π.
- agent gets lottery $w = \pi - \left(\mathbb{E}_{\pi \sim F(\cdot|e)}\left[\pi | \, e = e^{FB}\right] - c(e^{FB})\right)$ with $\mathbb{E}\left[w\right] = c(e^{FB})$

What if the agent is risk-averse, $v(w)$ strictly concave ?

- agent's expected utility $\mathbb{E}\left[v(w)\right] - c(e^{FB}) < 0$

e.g. if v satisfies $\lim_{w \searrow 0} v(w) = -\infty$, then we must have $w(\pi) > 0$ for all π to satisfy IR

Moral Hazard – Risk sharing

Suppose

- binary effort choice $e \in E = \{e_L, e_H\}$ with $e_L < e_H$
- output $\pi \in [\underline{\pi}, \bar{\pi}]$ with distribution $F(\cdot|e)$ satisfying $F(\pi|e_H) \leq F(\pi|e_L)$ for all π
- risk-averse agent: $v(w)$ increasing and strictly concave
- effort cost $c(e_L) = 0$ and $c(e_H) = c_H > 0$

Principal solves

$$\max_{e \in \{e_L, e_H\}, w(\cdot)} \int_{\underline{\pi}}^{\bar{\pi}} (\pi - w(\pi)) \, f(\pi|e) \, \mathrm{d}\pi \quad \text{such that}$$

$$e = \underset{e' \in \{e_L, e_H\}}{\operatorname{argmax}} \int_{\underline{\pi}}^{\bar{\pi}} v\left(w(\pi)\right) f(\pi|e') \, \mathrm{d}\pi - c(e'), \tag{IC}$$

$$\int_{\underline{\pi}}^{\bar{\pi}} v\left(w(\pi)\right) f(\pi|e) \, \mathrm{d}\pi - c(e) \geq 0 \tag{IR}$$

Moral Hazard – Risk sharing – example with binary effort

Example:

- $[\underline{\pi}, \bar{\pi}] = [0, 1]$ with distribution $f(\pi|e_L) = 2 - 2\pi$ and $f(\pi|e_H) = 1$ for all π
- $v(w) = \log(w)$

What is the optimal wage rule $w(\pi)$ to implement e_L?

What is the optimal wage rule $w(\pi)$ to implement e_H?

Moral Hazard – Risk sharing – multiple effort levels

Suppose now there are n effort levels $E = \{e_1, e_2, \ldots, e_n\}$

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 satisfies

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

Moral Hazard – Risk sharing – continuum of effort levels – first-order approach

Suppose now there is a continuum of effort levels $E = [\underline{e}, \bar{e}]$

Optimal wage rule $w(\pi)$ to implement some $e \in [\underline{e}, \bar{e}]$?

- now we have a continuum of incentive constraints: $(IC_{e,e'})$ for all $e' \in E$ with $e' \neq e$

Necessary condition for e to be optimal for the agent:

$$\frac{\partial}{\partial e}\left(\int_{\underline{\pi}}^{\bar{\pi}} v\left(w(\pi)\right) f(\pi|e)\,\mathrm{d}\pi - c(e)\right) = 0.$$

If we ignore all other (non-local) IC constraints, the optimal wage satisfies

$$\frac{1}{v'\left(w(\pi)\right)} = \gamma + \mu\frac{\frac{\partial}{\partial e'} f(\pi|e')}{f(\pi|e)}.$$

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 $c > 0$
- 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 project success (not tasks)
- agent i gets bonus b_i if project is successful; 0 otherwise

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?