G5212: Game Theory

Mark Dean

Spring 2017

Moral Hazard

- We will now move on to our final case of asymmetric information
    - Moral Hazard
- Key change: unlike previous models it is the **action** of the informed party that cannot be observed
    - Not type
- Three key components:
    - Agent takes an action which affects the utility of the principal
    - The principal only observes the outcome, imperfectly related to action
    - The action that the agent would choose on their own is not Pareto optimal
- Key point: Principal can condition payment on outcome, **not** on action

Moral Hazard

- Classic example
  - Say you are the boss of a door to door encyclopedia salesman
  - You know that the number of encyclopedias sold depends on
    - How hard the salesman works
    - Luck
  - How do you pay your salesman?
  - If the salesman is risk averse and you are not the first best would be to pay a flat wage
  - But if effort is costly and not observable then salesman will put in zero effort

Moral Hazard

- Moral hazard occurs whenever the objective of the informed and uninformed parties differ
  - Managers and shareholders
  - Property insurance
  - Car owner/mechanic
  - Sharecropping

A Simple Example

- Let's start with a simple example
  - Two actions
  - Two possible output levels
- Action:
  - Work: $a = 1$
  - Shirk $a = 0$
- Utility if receive income w is

$$u(w) - a$$

- u strictly concave
- Let $\bar{u}$ be the outside option: the utility the worker gets from sitting at home in bed eating crisps

A Simple Example

- The project can either succeed or fail
  - If the agent works, probability of success is p_1
  - If not, it is $p_0 < p_1$
- Principal gets payoff x_s if the project is a success and $x_f < x_s$ otherwise

Some Boring Cases

- First best
  - Principal can condition payment on a
  - Will select a^* that maximizes total surplus
  - Pay w^* such that $u(w^*) - a^* = \bar{u}$ if that level of effort is exerted
  - Otherwise pay $w = -\infty$
- Risk neutral agent
  - $u(w) = w$
  - Principal will select a^* that maximizes total surplus (say this is $a = 1$)
  - Total surplus is then $p_1 x_s + (1 - p_1)x_f - 1 = s^*$
  - Will 'sell the firm' to the agent
  - Pay wage $x_s - (s^* - \bar{u})$ if the project is a success
  - Pay wage $x_f - (s^* - \bar{u})$ in the bad state

A Simple Example

- Going back to the case of risk averse agent and asymmetric information
- Let's assume that it is in the best interests of the principal to get the agent to work
- This means that they have to incentivize them to work
- The only way they can do this is by conditioning payment on the outcome
  - Let w_s be the pay if the project is a success and w_f if it is a failure
- It must be the case that

$$\begin{aligned} p_1 u(w_s) + (1-p_1)u(w_f) - 1 &\geq p_0 u(w_s) + (1-p_0)u(w_f) \\ &\Rightarrow (p_1 - p_0)(u(w_s) - u(w_f)) \geq 1 \end{aligned}$$

A Simple Example

$$(p_1 - p_0)(u(w_s) - u(w_f)) \geq 1$$

- This is the IC constraint
- Note that as p_1 gets closer to p_0 the gap in wages must increase
- Incentives must become more ‘high powered’

A Simple Example

- As usual, we need an individual rationality constraint
- Then we need

$$p_1 u(w_s) + (1 - p_1) u(w_f) - 1 \geq \bar{u}$$

- Also as usual, for profit maximizing, this will hold at equality
  - otherwise we can reduce wages for both outcomes in such a way that does not affect IC but lowers wage bill

A Simple Example

- We can also show that the IC constraint must hold at equality
  - Assume not

$$(p_1 - p_0)(u(w_s) - u(w_f)) > 1$$

  - Proposal: reduce wage w_s by

$$\frac{(1 - p_1)\varepsilon}{u'(w_s)}$$

  - Increase wage w_f by

$$\frac{p_1 \varepsilon}{u'(w_f)}$$

A Simple Example

- For ε small enough
  - IC constraint still holds (as we were off the constraint before)
  - IR constrain holds
    - Fall in the utility after a success is $(1 - p_1)\varepsilon$
    - Increase in utility after a failure is $p_1\varepsilon$
    - Total change in utility is

$$-p_1(1 - p_1)\varepsilon + (1 - p_1)p_1\varepsilon = 0$$

A Simple Example

- Change in the wage bill given by

$$\begin{aligned} & -p_1 \frac{(1-p_1)\varepsilon}{u'(w_s)} + (1-p_1)\frac{p_1\varepsilon}{u'(w_f)} \\ = \quad & p_1(1-p_1)\varepsilon\left(\frac{1}{u'(w_f)} - \frac{1}{u'(w_s)}\right) < 0 \end{aligned}$$

- As

$$\begin{aligned} w_s \quad & > \quad w_f \\ & \Rightarrow \quad u'(w_s) < u'(w_f) \\ & \Rightarrow \quad \frac{1}{u'(w_s)} > \frac{1}{u'(w_f)} \end{aligned}$$

A Simple Example

- We therefore have two equations and two unknowns

$$(p_1 - p_0)(u(w_s) - u(w_f)) = 1$$

$$p_1 u(w_s) + (1 - p_1)u(w_f) - 1 = \bar{u}$$

- IR constraint pins down the overall utility level
- IC constraint pins down the difference between w_s and w_f
- Solution is

$$\begin{aligned} u(w_f) &= \bar{u} - \frac{p_0}{p_1 - p_0} \\ u(w_s) &= \bar{u} + \frac{1 - p_0}{p_1 - p_0} \end{aligned}$$

- Wages higher following a success than a failure

A Simple Example

- Is this the solution to the principal's problem?
- Not necessarily
- In this case the principal gets

$$p_1(x_s - w_s) + (1 - p_1)(x_f - w_f)$$

- Alternatively, could let the worker shirk, and only worry about the IR constraint
  - Pay constant wage w such that $u(w) = \bar{u}$
  - Get profit
$$p_0(x_s - w) + (1 - p_0)(x_f - w)$$
- Better to get the worker to work if

$$(p_1 - p_0)(x_s - x_f) \geq p_1 w_s + (1 - p_1)w_f - w$$