Section 2 – The Principal-Agent Model

Juan Dodyk

Gov 2006 – Harvard

February 4, 2020

Hidden action

The agent has two actions available: $a \in \{0, 1\}$. We can think of $a = 1$ as making effort and $a = 0$ as shirking.

The actions produce a stochastic benefit to the principal $b \in \{b_L, b_H\}$, where $b_L < b_H$.

We have $\Pr(b = b_H | a = 1) = \overline{p}$ and $\Pr(b = b_H | a = 0) = \underline{p}$ with $0 \leq \underline{p} < \overline{p} < 1$.

In words: making effort (choosing $a = 1$) doesn't guarantee a good outcome ($b = b_H$), but it makes it more likely.

The principal chooses ex ante a transfer T contingent on b.

The agent's payoff is $u(T) - ac$ where $u' > 0$, $u'' \leq 0$ and $c > 0$.

The principal's payoff is $b - T$.

First best

Assume that the action a is *contractible*, meaning that the contract can bind the agent to choose it (if he signs the contract).

The agent accepts a contract that requires him to choose $a = 1$ if his expected payoff is non-negative:

$$\mathbb{E}[u(T) - ac] \geq 0.$$

This is

$$\overline{p}u(T_H) + (1 - \overline{p})u(T_L) - c \geq 0. \quad \text{(PC)}$$

We assume that the principal writes the contract, so she chooses T_H, T_L to maximize her expected payoff

$$\mathbb{E}[b - T] = \overline{p}(b_H - T_H) + (1 - \overline{p})(b_L - T_L)$$

subject to the participation constraint (PC).

In lecture we saw that in the optimal contract (PC) must bind, and if u is strictly concave ($u'' < 0$) then we must have $T_H = T_L$.

Let's review the proof. We wanto to choose $T_L, T_H \in \mathbb{R}$ to

$$\begin{aligned} \text{maximize} \quad & \overline{p}(b_H - T_H) + (1 - \overline{p})(b_L - T_L) \\ \text{subject to} \quad & \overline{p}u(T_H) + (1 - \overline{p})u(T_L) - c = 0. \end{aligned}$$

From the (PC) we can get T_H as a function of T_L:

In lecture we saw that in the optimal contract (PC) must bind, and if u is strictly concave ($u'' < 0$) then we must have $T_H = T_L$.

Let's review the proof. We wanto to choose $T_L, T_H \in \mathbb{R}$ to

$$\begin{aligned} \text{maximize} \quad & \overline{p}(b_H - T_H) + (1-\overline{p})(b_L - T_L) \\ \text{subject to} \quad & \overline{p}u(T_H) + (1-\overline{p})u(T_L) - c = 0. \end{aligned}$$

From the (PC) we can get T_H as a function of T_L:

$$\tilde{T}_H(T_L) := T_H = u^{-1}\left[\frac{c - (1-\overline{p})T_L}{\overline{p}}\right].$$

So, replacing T_H, the problem is to choose T_L to maximize

$$V(T_L) := \overline{p}(b_H - \tilde{T}_H(T_L)) + (1-\overline{p})(b_L - T_L).$$

We have $V'(T_L) = -\overline{p}\tilde{T}'_H(T_L) - (1-\overline{p})$.

Two ways of calculating $\tilde{T}'_H(T_L)$.

First, using the inverse function theorem.

Second, differentiating the participation constraint implicitly.

In any case,

$$\tilde{T}_H'(T_L) = -\frac{1-\overline{p}}{\overline{p}} \frac{u'(T_L)}{u'(\tilde{T}_H(T_L))}.$$

So,

In any case,

$$\tilde{T}'_H(T_L) = -\frac{1-\overline{p}}{\overline{p}} \frac{u'(T_L)}{u'(\tilde{T}_H(T_L))}.$$

So,

$$V'(T_L) = (1-\overline{p}) \left[\frac{u'(T_L)}{u'(\tilde{T}_H(T_L))} - 1 \right].$$

Now, if T_L increases, $\tilde{T}_H(T_L)$ decreases, so $u'(T_L)$ decreases and $u'(\tilde{T}_H(T_L))$ increases, and $\frac{u'(T_L)}{u'(\tilde{T}_H(T_L))}$ decreases.

Hence $V'(T_L)$ is decreasing, so V is strictly concave, and $V'(T_L) = 0$ implies that T_L maximizes it.

Now, $V'(T_L) = 0$ is $u'(T_L) = u'(\tilde{T}_H(T_L))$, so the optimal T_L satisfies $T_L = \tilde{T}_H(T_L) = T_H$, as we wanted to prove.

There is a more conceptual proof that I'd like to show you, but first we need to review concavity and risk preferences.

Concave functions and risk

A function $u : S \to \mathbb{R}$ is concave if for every $x, y \in S$ and $\lambda \in [0, 1]$ we have

$$u(\lambda x + (1 - \lambda)y) \geq \lambda u(x) + (1 - \lambda)u(y).$$

If u is differentiable twice and $S \subset \mathbb{R}$ then this is equivalent to $u''(x) \leq 0$ for all $x \in S$.

A function $u : S \to \mathbb{R}$ is strictly concave if for every $x, y \in S$, $x \neq y$, and $\lambda \in (0, 1)$ we have

$$u(\lambda x + (1 - \lambda)y) > \lambda u(x) + (1 - \lambda)u(y).$$

If u is differentiable twice and $S \subset \mathbb{R}$ then this is equivalent to $u''(x) < 0$ for all $x \in S$.

Jensen's inequality. *Let $u : S \to \mathbb{R}$ be a concave function, and X be a random variable with support in S. We have*

$$u(\mathbb{E}[X]) \geqq \mathbb{E}[u(X)].$$

If u is strictly concave and X is not constant then

$$u(\mathbb{E}[X]) > \mathbb{E}[u(X)].$$

In the case $X = x$ with proba λ and $X = y$ with proba $1 - \lambda$, Jensen's theorem is just the definition of concavity. It is a generalization of the definition.

Meaning: an agent with concave utility u always prefers to receive $\mathbb{E}[X]$ for sure than the lottery X.

Hence, concavity means risk-aversion. The agent doesn't like variance, i.e., risk.

Risk-neutrality means that the agent only cares about $\mathbb{E}[X]$, so $\mathbb{E}[u(X)] = u(\mathbb{E}[X])$ for every random variable X. This is satisfied if and only if u is linear, say $u(x) = x$.

There are two popular measures of risk aversion: absolute risk aversion, defined as

$$r(x) = -\frac{u''(x)}{u'(x)},$$

and relative risk aversion, defined as

$$\rho(x) = -\frac{xu''(x)}{u'(x)}.$$

These lead to the popular constant absolute risk aversion (CARA) and constant relative risk aversion (CRRA) utility functions:

$$u_r(x) = \begin{cases} -\frac{1}{r}e^{-rx}, & \text{if } r \neq 0, \\ x, & \text{if } r = 0, \end{cases} \text{ and } u_\rho(x) = \begin{cases} \frac{1}{1-\rho}x^{1-\rho}, & \text{if } \rho \neq 1, \\ \ln(x), & \text{if } \rho = 1. \end{cases}$$

Back to the first best

The principal chooses $T_L, T_H \in \mathbb{R}$ to

$$\begin{aligned} \text{maximize} \quad & \overline{p}(b_H - T_H) + (1-\overline{p})(b_L - T_L) \\ \text{subject to} \quad & \overline{p}u(T_H) + (1-\overline{p})u(T_L) - c \geq 0. \end{aligned}$$

We want to show that at the optimum we have $T_L = T_H$.

Now $\overline{p}(b_H - T_H) + (1-\overline{p})(b_L - T_L) = \mathbb{E}[b - T] = \mathbb{E}[b] - \mathbb{E}[T]$.

So, the principal chooses T_L, T_H to minimize what she has to pay, i.e., $\mathbb{E}[T]$. Note that she only cares about the expectation.

The participation constraint is $\mathbb{E}[u(T)] - c \geq 0$. Note that the agent is risk-averse if we assume that $u'' < 0$. He wants as little risk as possible.

Proof that $T_L = T_H$

Suppose that the optimal T is not constant, i.e., $T_L \neq T_H$.

By Jensen, we have $u(\mathbb{E}[T]) > \mathbb{E}[u(T)]$.

By the PC, we have $\mathbb{E}[u(T)] \geq c$, so $u(\mathbb{E}[T]) > c$.

So, we can take $\epsilon > 0$ such that $u(\mathbb{E}[T] - \epsilon) > c$.

Now $T'_L = T'_H = \mathbb{E}[T] - \epsilon$ satisfy the PC, and $\mathbb{E}[T'] = \mathbb{E}[T] - \epsilon$, so the principal prefers T' to T (she has to pay less money in expectation), contradicting the assumption that T was optimal.

Hence the optimal T must be constant, i.e., $T_L = T_H$, as we wanted to prove.

A couple of comments

Note that this proof works for any number of actions and outcomes.

If $u'' = 0$, i.e., the agent is risk neutral, then any T_L, T_H such that $\mathbb{E}[T] = u^{-1}(c)$ are optimal. Hence we need strict concavity for the result.

In fact, if the principal is risk-averse and the agent is risk-neutral, we get the opposite result: in the first best the principal transfers all the risk to the agent. Conceptually, the principal makes the agent the residual claimant, i.e., she effectively "sells" the job to the agent.

Questions?

Review of contracting with hidden type

To make this slightly different from the lecture, let's study this as a political accountability problem.

There is an incumbent (the agent) and a representative voter (the principal), and two periods.

In the first period, The incumbent chooses $a \geq 0$, her effort, at cost ca, where $c \in \{c_L, c_H\}$, $0 < c_L < c_H$, is her type, which is private information. We have $\Pr(c = c_L) = \mu \in (0, 1)$.

The voter's utility is $u(a)$, where $u' > 0$, $u'' < 0$, $u(0) = 0$. For example, $u(x) = \sqrt{x}$.

The voter offers a menu of contracts: $(a_L, p_L), (a_H, p_H)$, which mean that if the incumbent chooses a_T then the voter reelects her with probability p_T, for $T \in \{L, H\}$. If the incumbent chooses another a, the voter doesn't reelect.

In the second period, the incumbent gets a wage δ and chooses a_2 at cost ca_2. Since it's the last period, there are no incentives for her to choose $a_2 > 0$, so we can assume that $a_2 = 0$. Hence every politician behaves the same way in period 2, so the voter is indifferent between reelecting the period-1 incumbent or not. So the promise to reelect with proba p_T is credible.

The problem is then to choose $(a_L, p_L), (a_H, p_H)$ to maximize

$$\mu u(a_L) + (1 - \mu)u(a_H)$$

subject to

$$p_L\delta - c_L a_L \geq p_H\delta - c_L a_H, \tag{ICL}$$
$$p_H\delta - c_H a_H \geq p_L\delta - c_H a_L, \tag{ICH}$$
$$p_L\delta - c_L a_L \geq 0, \tag{IRL}$$
$$p_H\delta - c_H a_H \geq 0. \tag{IRH}$$

Solution

It's given by

$$\begin{aligned} a_H &= \frac{p_H}{c_H}\delta, \\ a_L &= a_H + \frac{1-p_H}{c_L}\delta, \\ p_L &= 1, \\ u'(a_L) &= \frac{\mu}{1-\mu}\frac{c_H - c_L}{c_L}u'(a_H). \end{aligned}$$