Recap

- Last Thursday (February 13, 2003)
  - Results and analysis of the one-card poker game
  - Principal-agent problem
    - First-best contract
    - Moral hazard
    - Contracts with wages conditioned on outcomes
- Today (February 18, 2003)
  - Principal-agent problem: general model

Principal-Agent Problem

- If the agent doesn’t accept the contract, his payoff is his reservation utility, $U$.
- If the agent accepts the contract, he chooses between $n$ possible actions: $a_1, \ldots, a_n$.
- These actions produce $m$ possible outcomes: $x_1, \ldots, x_m$.
- There is a stochastic relationship between actions and outcomes (called “technology”). When the action is $a_i$, the principal observes outcome $x_j$ with probability $P_{ij}$.
- If the principal observes outcome $x_j$, she pays the agent $w_j$.
- The agent’s payoff is $U(w) - e(a_i)$, where $U$ is the utility of wage $w$ to the agent and $e(a_i)$ is the cost of action $a_i$ to the agent.
- $U$ is increasing, differentiable and concave.
- Assuming the principal is risk-neutral, her payoff is $x_j - w_j$.

Principal’s problem: Step 1

- Given an action $a_i$, how to set the wages such that the agent chooses $a_i$ and the principal’s payoff is maximized?

- Participation constraint: $\sum_{j=1}^{m} P_j U(w_j) - e(a_i) \geq U$
- Incentive constraint: $\sum_{j=1}^{m} P_j U(w_j) - e(a_i) \geq \sum_{j=1}^{m} P_j U(w_k) - e(a_k) \quad \forall k \neq i$
- Principal's objective (minimize the expected cost of inducing the agent to choose $a_i$):
  $$\max \sum_{j=1}^{m} P_j (x_j - w_j) = \min \sum_{j=1}^{m} P_j w_j = C(a_i)$$

Principal’s problem: Step 2

- From step 1:
  $$\max \sum_{j=1}^{m} P_j (x_j - w_j) = \sum_{j=1}^{m} P_j x_j - \min \sum_{j=1}^{m} P_j w_j = R(a_i) - C(a_i)$$
- What is the action $a_i$ that maximizes $R(a_i) - C(a_i)$?
What if the agent’s actions can be observed?

- The principal can design a contract where the wages are conditioned on the actions, i.e., \( w(a_i) \)

\[
\max \sum_{j=1}^{m} P_j(x_j - w_j) = \min w_i
\]

Participation constraint: \( U(w_i) - e(a_i) \geq U \)

Incentive constraint:
\[
U(w_i) - e(a_i) \geq U(w_i) - e(a_j) \quad \forall k \neq i
\]

To induce the agent to choose action \( a_i \), set \( w_i \) such that \( U(w_i) = U + e(a_i) \) and set all other \( w_k \) sufficiently low.

Principal’s problem: Step 1

\[
L(w_i, \lambda, \mu) = \sum_{j=1}^{m} P_j(x_j - w_j) + \mu \sum_{j=1}^{m} P_j U(w_j) - e(a_i) - U
\]

\[
\sum_{i=1}^{n} \lambda_i \left( \sum_{j=1}^{m} P_j U(w_j) - e(a_i) - \sum_{j=1}^{m} P_j U(w_j) + e(a_i) \right)
\]

\[
\frac{\partial L(w_i, \lambda, \mu)}{\partial w_i} = -P_0 + \mu P_i \frac{\partial U(w_j)}{\partial w_j} + \sum_{i=1}^{n} \lambda_i \left( P_i \frac{\partial U(w_j)}{\partial w_j} - P_0 \frac{\partial U(w_j)}{\partial w_j} \right) = 0
\]

\[
-\frac{P_0}{P_i} + \frac{\mu}{P_i} + \sum_{i=1}^{n} \lambda_i \left( P_i - P_0 \right) = 0
\]

\[
1 + \frac{\partial U(w_j)}{\partial w_j} \left( \mu + \sum_{i=1}^{n} \lambda_i \left( P_i - P_0 \right) \right) = 0 \quad \Rightarrow \quad \frac{1}{U(w_j)} = \mu + \sum_{i=1}^{n} \lambda_i \left( \frac{P_i}{P_0} \right)
\]

Principal’s problem: Step 1

\[
\frac{1}{U'(w_j)} = \mu + \sum_{i=1}^{n} \lambda_i \left( 1 - \frac{P_i}{P_0} \right)
\]

\[
P_0 \quad : \text{likelihood ratio}
\]

Two actions and two outcomes

- Suppose the agent has two possible actions, \( a \) and \( b \), and there are two possible outcomes, \( x_1 \) and \( x_2 \).

- Suppose that action \( b \) is preferred by the principal

For action \( b \):

\[
\max \sum_{j=1}^{2} P_j(x_j - w_j)
\]

\[
P_{b1} U(w_1) + P_{b2} U(w_2) - e(b) \geq U \quad \mu
\]

\[
P_{b1} U(w_1) + P_{b2} U(w_2) - e(b) \geq P_{a1} U(w_1) + P_{a2} U(w_2) - e(a) \quad \lambda
\]

Incentive compatibility constraint:

\[
P_{a1} U(w_1) + P_{a2} U(w_2) - e(a) \geq P_{b1} U(w_1) + P_{b2} U(w_2) - e(b)
\]

When is the agent indifferent between \( a \) and \( b \)?

\[
P_{b1} U(w_1) + P_{b2} U(w_2) - e(b) = P_{a1} U(w_1) + P_{a2} U(w_2) - e(a)
\]

\[
U(w_1) = \frac{P_{a1} - P_{b1}}{P_{a1} - P_{b2}} U(w_1) + \frac{P_{b1} - P_{a1}}{P_{a2} - P_{b2}} U(w_2) + \frac{P_{b1} - P_{a1}}{P_{a2} - P_{b2}} = U(w_1) + \frac{P_{b1} - P_{a1}}{P_{a2} - P_{b2}} \frac{e(b) - e(a)}{P_{a2} - P_{b2}}
\]

since \( P_{a1} + P_{a2} = P_{b1} + P_{b2} = 1 \)
Two actions and two outcomes

Action b is preferred to action a

\[ U(w_2) = U(w_1) + \frac{e(b) - e(a)}{P_{b2} - P_{a2}} \]

Participation constraint:

Action b: \( P_{a2} U(w_1) + P_{a2} U(w_2) - e(b) \geq U \)

Action a: \( P_{a2} U(w_1) + P_{a2} U(w_2) - e(a) \geq U \)

\[ U(w_2) \geq \frac{P_{a2}}{P_{b2}} U(w_1) + \frac{e(b) + U}{P_{b2}} \]

Incentive compatibility constraint:

Principal's objective function

For action b:

\[ \max P_{a2}(x_i - w_1) + P_{a2}(x_j - w_2) \]

Let \( f \) be the inverse of \( U \):

\[ f(U(w_j)) = w_j \]

\[ \max P_{a2}(x_i - f(U(w_1))) + P_{a2}(x_j - f(U(w_2))) \]

Slope of the objective function:

\[ -P_{a2} \frac{\partial f(U(w_1))}{\partial U(w_1)} - P_{a2} \frac{\partial f(U(w_1))}{\partial U(w_1)} \frac{\partial U(w_2)}{\partial U(w_1)} = 0 \]

\[ \frac{\partial U(w_2)}{\partial U(w_1)} = -\frac{P_{a2} f'(U(w_1))}{P_{b2} f'(U(w_2))} \]

Principal’s isoprofit curves

\[ P_{a2}(x_i - f(U(w_1))) + P_{a2}(x_j - f(U(w_2))) = C \]
Two actions and two outcomes

If $U(w_1) = U(w_2)$ (45° line)

$$\frac{\partial U(w_1)}{\partial U(w_1)} = \frac{P_{a1}}{P_{a2}}$$

Same as the slope of the participation constraint!

The principal's isoprofit curve is tangent to the agent's participation constraint along the 45° line.

Two actions and two outcomes

$$\frac{1}{U'(w_j)} = \mu + \lambda \left( 1 - \frac{P_{a1}}{P_{a2}} \right)$$ likelihood ratio

If the incentive constraint is not binding:

$\lambda = 0$, $U'(w_j) = 1/\mu$ and $w_j = w$.

Substituting into the incentive compatibility constraint (the agent prefers action $a$ over action $b$):

$$\sum_{j \in a} P_j U(w_j) \geq e(a) \rightarrow e(a) \geq e(b)$$

Two actions and two outcomes

$$\frac{1}{U'(w_j)} = \mu + \lambda \left( 1 - \frac{P_{a1}}{P_{a2}} \right)$$ likelihood ratio

If the incentive constraint is binding:

$\lambda > 0$, and $U'(w_j)$ depends on the likelihood ratio (LR).

A low value of LR is in favor of the agent choosing $a$.

The optimal incentive scheme is a linear function of LR.