Lecture 1: Introduction to Mechanism Design

August 31, 2001

Announcement

- Introduction
- Almost paperless course
- Course URL: www.haas.berkeley.edu/~hermalin/econ206.html

Lecture

1. Course Overview
   (a) What is mechanism design & agency theory?
      i. application of game theory
      ii. method of studying how *strategic* parties attempt to govern selves
      iii. means of understanding *certain* aspects of organization
   (b) Topics
      i. basics of agency and mechanism design
      ii. incomplete contracts and renegotiation
      iii. organization
      iv. agency and market interactions
2. Contractual Situations
   (a) Consider Figure 1.
   (b) Examples
      i. parties are employer & employee; endowments are ownership of productive assets (employer) and skills (employee); information structure is employee's action, but not consequence of action, hidden from employer; etc.
      ii. parties are social planner and citizens; endowments are transferable asset; information structure is citizens have private knowledge of benefit of public good; etc.
   (c) The basic division

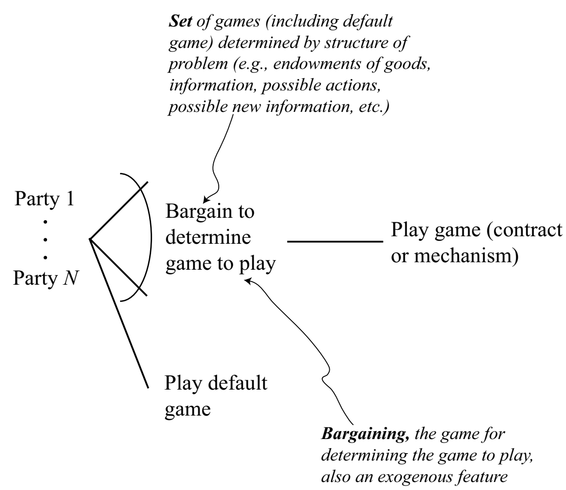

Figure 1: Schematic of contract design.

   i. **Hidden action:** Symmetric distribution of information until one player takes *actions* not observable by other.
   ii. **Hidden information:** Asymmetric distribution of information *regardless* of any player's actions.

3. Basics of Hidden-Information Agency

   (a) Structure

      i. two players, P and A.
      ii. P has bargaining power.
      iii. but A has information.
      iv. *example:* P is buyer who wants A to provide a custom item. A's expertise allows A to know cost of production. P's ignorant of cost of production.
      v. *example:* P is a monopolist who wants to sell to A. A is buyer who knows his valuation for product, which P doesn't know.
      vi. **timing.** A learns information. P proposes contract. A accepts or rejects. Rejection "ends" game. If A accepts, game ensues as dictated by contract.

   (b) Example: Saturday-night restrictions

      i. $P =$ airline; $A =$ passenger.
      ii. A is either tourist or business traveller (T or B).
      iii. two kinds of tickets, restricted or unrestricted (R and U).

iv. A's value for a ticket a function of type of traveller, $\theta \in \{T, B\}$, and kind of ticket, $k \in \{R, U\}$:

$$V(\theta, k) = \begin{cases} \$1000, & \text{if } \theta = B \text{ and } k = U \\ \$600, & \text{if } \theta = B \text{ and } k = R \\ \$400, & \text{if } \theta = T \text{ and } k = U \\ \$300, & \text{if } \theta = T \text{ and } k = R \end{cases}.$$

v. suppose P has one seat, $MC = 0$, and half the population is B.

vi. **benchmark.** Ideal is for P to sell only U tickets at \$1000 to business travellers and \$400 to tourists. *Problem: business travellers won't admit they're business travellers.*

vii. **next best solution.** Charge \$300 for R ticket and \$700 for U ticket for expected profit of \$500.

viii. why work? Clearly no one pays more for a ticket than $V(\theta, k)$. Both types of travellers maximize their surplus, $V(\theta, k) - p_k$:

$$V(\theta, k) - p_k = \begin{cases} \$1000 - \$700 = \$300 & \text{if } \theta = B \text{ and } k = U \bigstar \\ \$600 - \$300 = \$300 & \text{if } \theta = B \text{ and } k = R \\ \$400 - \$700 = -\$300 & \text{if } \theta = T \text{ and } k = U \\ \$300 - \$300 = \$0 & \text{if } \theta = T \text{ and } k = R \bigstar \end{cases}.$$

ix. **exercise.** Consider a general situation in which $V(\theta, U) > V(\theta, R)\ \forall \theta$, $V(B, k) > V(T, k)\ \forall k$, $MC = 0$, and $\text{Prob}\{\theta = B\} \in (0, 1)$. Work out what the expected profit-maximizing values of p_U and p_R are.

x. *second-degree price discrimination via quality distortions*

A. different classes of service

B. "pro" versions versus other versions of software

C. luxury vs. no luxury features on cars, stereos, etc.

(c) The problem more generally

- Two players are involved in a strategic relationship; that is, each player's well being depends on the play of the other player.
- One player is better informed (or will become better informed) than the other; that is, he has *private information* about some state of nature relevant to the relationship. As is typical in information economics, we refer to the player with the private information as the *informed player* and the player without the private information as the *uninformed player*.
- Critical to our analysis of these situations is the bargaining game that determines the contract. We will refer to the contract proposer as the principal and the player who receives the proposal as the agent. Moreover, we assume contracts are proposed on a take-it-or-leave-it basis: The agent's only choices are to accept or reject the contract proposed by the principal. Rejection

ends the relationship between the players. A key assumption is that the principal is the *un*informed player. Models like this, in which the uninformed player proposes the contract, are referred to as *screening* models. In contrast, were the informed player the contract proposer, we would have a type of signaling model.

- A contract can be seen as setting the rules of a secondary game to be played by the principal and the agent.

4. The Two-Type Model — Set Up

   (a) Timing

   - A learns his *type*, $\theta \in \Theta$. *Critically,* this is his *private* information.
   - P proposes a contract on a take-it-or-leave-it (tioli) basis.
   - In equilibrium, A accepts the contract and plays contract.
   - Payoffs realized.

   (b) Preferences

   i. two goods, x the allocated good and y money (all other goods)
   ii. A's preferences, $y - C_\theta(x)$, $C'_\theta > 0$ and $C''_\theta \geq 0$.
   iii. P's preferences, $r(x) + y$, $r' > 0$.

   (c) Interpretation: A is producing x units of some good desired by P. A's cost depends on its type, t.

   (d) If $C_\theta(\cdot)$ is a cost, then $C_\theta(0) = 0$. ***Why?***

   (e) A *contract* between P and A is an agreement on x and a payment *from* P to A of s; yielding utilities $r(x) - s$ and $s - C_\theta(x)$, respectively.

   (f) Efficiency dictates that x maximize $r(x) - C_\theta(x)$.

   (g) Define $x^F_\theta = \arg\max r(x) - C_\theta(x)$.

   (h) Assume $C_\theta(0) = 0$ and $C'_1(x) > C'_2(x)$ for all x. Observe $\Theta = \{1, 2\}$. Observe, consequence, $C_1(x) > C_2(x)$ for all $x > 0$.

   (i) **Exercise.** Prove $x^F_1 < x^F_2$ (unless $x^F_2 = 0$); henceforth assume $x^F_2 > 0$.

   (j) **The first best.** The optimal contract from P's perspective is an efficient contract — i.e., one that achieves x^F_θ — at minimum cost. *A participates* only if $s - C_\theta(x) \geq 0$, so minimum cost means $s^F_\theta = C_\theta(x^F_\theta)$.

   (k) Observe the first best is not achievable — $s^F_1 - C_2(x^F_1) = C_1(x^F_1) - C_2(x^F_1) > 0 = s^F_2 - C_2(x^F_2)$.

   (l) The difference $C_1(x^F_1) - C_2(x^F_1)$ is a potential *information rent.*

   (m) Define $R(x) = C_1(x) - C_2(x)$ as the information-rent function. Observe $R(x) > 0$ for all $x > 0$. Observe, too, that $R'(\cdot) = C'_1(\cdot) - C'_2(\cdot) > 0$; *i.e.,* $R(\cdot)$ is an increasing function.

5. The Two-Type Model — Solution

   (a) P's problem:
   $$\max \mathbb{E}_\theta\{r(x_\theta) - s_\theta\}$$

   (b) Let $f = \text{Prob}\{\theta = 1\}$.

   (c) Constraints: sorting, incentive-compatibility, or truth-telling —
   $$u_2 \equiv s_2 - C_2(x_2) \geq s_1 - C_2(x_2) = u_1 + R(x_1) \qquad \text{(IC2)}$$
   $$u_1 \equiv s_1 - C_1(x_1) \geq s_2 - C_1(x_2) = u_2 - R(x_2). \qquad \text{(IC1)}$$

   (d) Observe these imply
   $$R(x_1) \leq u_2 - u_1 \leq R(x_2)\,,$$
   which means $x_1 \leq x_2$. ***Why?***

   (e) Constraints: participation — $u_\theta \geq 0$. Individual rationality constraints.

   (f) So P's problem is
   $$\max_{\{s_1,s_2,x_1,x_2\}} f[r(x_1) - s_1] + (1-f)[r(x_2) - s_2]$$
   subject to (IC1), (IC2), and (IR) constraints.

   (g) Can rewrite problem as
   $$\max_{\{u_1,u_2,x_1,x_2\}} f[r(x_1) - u_1 - C_1(x_1)] + (1-f)[r(x_2) - u_2 - C_2(x_2)]$$
   subject to
   $$u_2 \geq u_1 + R(x_1) \text{ and } u_1 \geq u_2 - R(x_2) \qquad \text{(IC);}$$
   and
   $$u_1 \geq 0 \text{ and } u_2 \geq 0\,. \qquad \text{(IR)}$$

   (h) Ideal is set $u_\theta = 0$; can't because violates one of (IC)

   (i) Because $R > 0$, A_2's participation constraint must be slack.

   (j) Because can lower u_t by uniform amount without Δing (IC), A_1's participation is binding.

   (k) Know $R(x_1) \leq u_2 \leq R(x_2)$. Want u_2 as small as possible, so left constraint binding.

   (l) Substituting the binding constraints, the problem is
   $$\max_{\{x_1,x_2|x_1\leq x_2\}} f[r(x_1) - C(x_1)] + (1-f)[r(x_2) - \underbrace{R(x_1)}_{u_2} - C(x_2)]$$

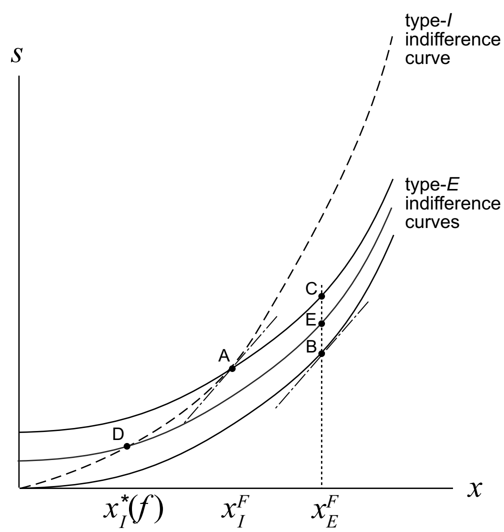

Figure 2: Two-type model.

or, equivalently

$$\max_{\{x_1,x_2|x_1\leq x_2\}} f[r(x_1) - \underbrace{\left(C(x_1) + \frac{1-f}{f}R(x_1)\right)}_{\text{virtual cost}}] + (1-f)[r(x_2) - C(x_2)]\,.$$

(m) *results.*

i. $x_2^*(f) = x_2^F$ — no distortion at the top;

ii. $x_1^*(f) \leq x_1^F$ (and strictly less if $x_1^F > 0$) — distortion at the bottom. ***Why?***

6. Two-Type Model — Graphical Interpretation

(a) Assume $r(x) = x$

(b) Discuss figure 2.

7. Interpreting the Contract

(a) Contract is a menu. An x off the menu is punished.

(b) Alternatively, a payment schedule:

$$S(x) = \begin{cases} 0, \text{ if } x < x_1^*(f) \\ C_1[x_1^*(f)], \text{ if } x_1^*(f) \leq x < x_2^*(f) \\ R[x_1^*(f)] + C_2[x_2^*(f)], \text{ if } x_2^*(f) \leq x \end{cases}.$$

General Problem

8. Structure

   (a) Nature determines state of nature (A's type), $\theta \in \Theta$, according to *commonly* known distribution function, F.

   (b) A learns *type*. This is his *private* information.

   (c) P proposes a contract (*mechanism*) to A on a take-it-or-leave-it basis. [**Digression:** screening v. signaling.]

   (d) A contract defines a game. The game includes actions, $m \in \mathcal{M}$ for A and $n \in \mathcal{N}$ for P and a mapping, possibly stochastic, $\sigma : \mathcal{M} \times \mathcal{N} \to \Delta(\mathcal{X} \times \mathcal{S})$, where $\mathcal{X}$ is the set of possible allocations of a good x, $\mathcal{S}$ are the possible transfers of a transferable good from P to A, and Δ are possible distributions over $\mathcal{X} \times \mathcal{S}$, including degenerate "mass" distributions.

   (e) The players have von Neumann-Morgernstern preferences expressed by utility functions $\mathcal{W} : \mathcal{X} \times \mathcal{S} \times \Theta \to \mathbb{R}$ and $\mathcal{U} : \mathcal{X} \times \mathcal{S} \times \Theta \to \mathbb{R}$ for P and A, respectively.

   (f) As a slight abuse of notation, define $\mathcal{W}(\sigma, \theta) = \mathbb{E}_\sigma\{\mathcal{W}(x, s, \theta)\}$ and $\mathcal{U}(\sigma, \theta) = \mathbb{E}_\sigma\{\mathcal{U}(x, s, \theta)\}$.

   (g) What happens if A rejects P's proposal?

      i. then play default game Γ_0.

      ii. $s = 0$ in Γ_0.

      iii. critical issue: Who/what determines x in default game? In C&H notes, a bit sloppy — assume either (i) P does (*i.e.,* $x_0 = \arg\max \mathbb{E}_\theta\{\mathcal{W}(x, 0, \theta)\}$); (ii) there's some automatic default, x_0; or (iii) A does and $x_0 = \arg\max \mathcal{U}(x, 0, \theta)\ \forall\theta$. Examples: (i) P is some regulator who can impose solution, x_0, if A doesn't cooperate; (ii) P owns productive assets and A is necessary worker, idle assets yield x_0; (iii) trade — absent compensation A optimally wishes not to produce (*e.g.,* $x_0 = 0$). But can conceive of situations in which A chooses and x_0 not constant (*e.g.,* A retains control rights if doesn't come to agreement with regulator, P).

      iv. we'll assume fixed x_0.

      v. define reservation utilities $W_R \equiv \mathbb{E}_\theta\{\mathcal{W}(x_0, 0, \theta)\}$ and $U_R(\theta) = \mathcal{U}(x_0, 0, \theta)$.

      vi. if $(x_0, 0) \in \mathcal{X} \times \mathcal{S}$, then no loss of generality in assuming that an agreement is reached in equilibrium.

9. What Game?

   (a) Seems like too much freedom if can pick any $\mathcal{M}$, $\mathcal{N}$, and σ.

   (b) Fortunately, have the revelation principle.

   (c) A mechanism is *direct* if $\mathcal{M} = \Theta$; that is, A's action in the mechanism is limited to making announcements about his type.

(d) A mechanism is *direct revelation* if it's a direct mechanism *and* an equilibrium strategy is truth-telling for all types; that is, if $m : \Theta \to \Theta$ is a strategy for A, then an equilibrium is $m(\theta) = \theta; \forall\theta$; that is, for any θ and $\theta' \in \Theta$ we have

$$\mathcal{U}(\sigma(\underbrace{\theta}_{m}), \theta) \geq \mathcal{U}(\sigma(\theta'), \theta)\,.$$

(e) **revelation principle.** For any general contract $(\mathcal{M}, \mathcal{N}, \sigma)$ and associated Bayesian equilibrium, there exists a direct-revelation mechanism such that the associated truthful Bayesian equilibrium generates the same equilibrium outcome as the general contract.

Proof: The proof of the revelation principle is standard but informative. A Bayesian equilibrium of the game $(\mathcal{M}, \mathcal{N}, \sigma)$ is a pair of strategies $(m(\cdot), n)$.[1] Let us consider the following direct mechanism: $\hat{\sigma}(\cdot) = \sigma(m(\cdot), n)$. Our claim is that $\hat{\sigma}(\cdot)$ induces truth-telling (is a direct-*revelation* mechanism). To see this, suppose it were not true. Then there must exist a type θ such that the agent does better to lie—announce some $\theta' \neq \theta$—when he is type θ. Formally, there must exist θ and $\theta' \neq \theta$ such that

$$\mathcal{U}(\hat{\sigma}(\theta'), \theta) > \mathcal{U}(\hat{\sigma}(\theta), \theta).$$

Using the definition of $\hat{\sigma}(\cdot)$, this means that

$$\mathcal{U}(\sigma\left[m(\theta'), n\right], \theta) > \mathcal{U}(\sigma\left[m\left(\theta\right), n\right], \theta);$$

but this means the agent prefers to play $m(\theta')$ instead of $m(\theta)$ in the *original* mechanism against the principal's equilibrium strategy n. This, however, can't be since $m(\cdot)$ is an equilibrium best response to n in the original game. Hence, truthful revelation must be an optimal strategy for the agent under the constructed direct mechanism. Finally, when the agent truthfully reports the state of nature in the direct truthful mechanism, the same outcome $\hat{\sigma}(\theta) = \sigma(m(\theta), n)$ is implemented in equilibrium. ■

(f) Figure 3 illustrates.

[1] Observe that the agent's strategy can be conditioned on θ, which he knows, while the principal's cannot be (since she is ignorant of θ).

$$
\begin{tikzcd}
\Theta \arrow[rr, "\hat{\sigma}(\bullet) \equiv \sigma[m(\bullet),n]"] \arrow[dr, "m(\bullet)"'] & & \Delta(\mathcal{X}\times\mathcal{S}) \\
& \mathcal{M} \arrow[ur, "\sigma(\bullet,n)"'] &
\end{tikzcd}
$$

Figure 3: The revelation principle.