1 Mechanism Design

- This section explores general properties of economic design problems.

- Game theory takes the game form given, and asks how do the players behave in the given game.

- But in many scenarios the interesting questions are quite the opposite: How should the planner design the institution, or game form, so that his goals are achieved.

- Thus the game form is not given but subject to formulation by the planner. The players (agents) in turn take the game form as given and play it rationally (according to the chosen equilibrium). In this scenario, agents' strategic behavior is a constraint that restricts what can be achieved within an institution.

- When information is complete, and the planner all-mighty, the constraint imposed by players' strategic behavior becomes vacuous: The planner forces the ideal outcome (or at least the best outcome that the agents accept).

- Situation become more problematic when information is incomplete. Then it may no longer be known to the planner what the optimal decision is. One has to design a game form that *elicits* the information from the agents.

- Mechanism design studies the means of implementing an allocation when the information is dispersed in the economy.

- In general two questions arise:
  - Which outcomes functions can be implemented?
  - What is the optimal outcome function in the class of implementable functions?
- Applications: public economics, theory of the firm, regulation, contract theory, auction design, computer architecture,... .

Setting

- The set of agents, $I = \{1, ..., n\}$.

- The principal, mechanism designer, has subscript 0.

- A collective choice among a set of possible allocations A.

- Agent i observes privately her type $\theta_i \in \Theta_i$.

- This type, that determines preferences over A described by a utility function $u_i(a, \theta)$ for all.

- Prior distribution over types $p(\theta)$ is common knowledge.

- In general the type could contain any information agent i possesses about his preferences, but it could also contain information about his neighbor, competitors, and alike.

The Model

Definition A *social choice function* is a mapping:

$$f : \Theta_1 \times .... \times \Theta_n \to A.$$

- The problem is that $\theta = (\theta_1, .., \theta_n)$ is not publicly observable when the allocation a is to be decided.

- However each agent can send a message to the principal $s_i \in S_i$.

- The message space can be arbitrary, with

$$S = \times_{i=1}^{n} S_i.$$

- After the agents have transmitted a message s, the center chooses a social allocation $a \in A$ as a function of the message received.

Definition An *outcome function* is a mapping

$$g : S_1 \times ... \times S_n \rightarrow A.$$

- Each agent transmits the message independently and simultaneously with all other agents.

- Thus a mechanism *defines* a game with incomplete information for which must choose an equilibrium concept, denoted by c.

- A *strategy* for agent i is a function

$$s_i : \Theta_i \to S_i.$$

Definition A mechanism $\Gamma = (S, g)$ is a collection of strategy sets $S = S_1 \times ... \times S_n$ and an outcome function

$$g : S \to A.$$

- Thus a mechanism is a strategic game without specification of preferences or information structure.

- With a slight abuse of notation, we use the same notation, S_i, for the set of strategies and their range for a particular agent i.

- Graphically, we have the following commuting diagram:

$$\begin{array}{ccccc} \Theta_1 \times ... \times \Theta_n & & \overset{f(\cdot)}{\longrightarrow} & & A \\ & \searrow & & \nearrow & \\ & s^*(\cdot) & & g(\cdot) & \\ & & S_1 \times ... \times S_n & & \end{array}$$

where $s^*(\theta)$ is the equilibrium message profile under θ, whatever the used equilibrium notion is. The strategy space of each individual agent is often called the message space.

Mechanism as a Game

- The central question we would like to ask is whether a particular objective function, or social choice function $f(\cdot)$ can be realized as an appropriate equilibrium of the game induced by the mechanism.

Definition A mechanism $\Gamma = (S, g)$ *implements* the social choice function f if for all $\theta \in \Theta$ there is an *equilibrium* profile

$$s^*(\theta) = \left(s_1^*\left(\theta_1\right), ..., s_n^*\left(\theta_n\right)\right)$$

of the game induced by Γ such that

$$g\left(s^*\left(\theta\right)\right) = f\left(\theta\right).$$

- Observe that the concept of implementation given in the definition above is not as strong as it might be. Why not require that the condition be satisfied for all equilibrium profiles?

- This stronger sense of implementation is referred to in the literature as strong or unique or full implementation.

- We focus on the weaker concept. Our question: Is there a conflict between incentive compatibility and the social choice function?

- The identification of implementable social choice function is at first glance a complex problem because we have to consider all possible mechanism g on all possible domains of strategies S.

- However a useful result (valid for all of the implementation versions above), the *revelation principle*, simplifies the task.

Definition A mechanism is *direct* if $S_i = \Theta_i$ and $g(\cdot) = f(\cdot)$ for all i.

Definition The social choice function $f(\cdot)$ is *truthfully implementable* (or *incentive compatible)* if the direct revelation mechanism

$$\Gamma = (\Theta, f)$$

has an equilibrium $\left(s_1^*(\theta_1), ..., s_n^*(\theta_n)\right)$ in which $s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$, for all i.

Dominant Strategy Equilibrium

- Next we consider implementation in dominant strategies. We show that in the absence of prior restrictions on the characteristics, implementation in dominant strategy equilibria is essentially impossible.

General Environments

- First, we prove the revelation principle for the dominant equilibrium concept. Equivalent results can be proven for Nash and Bayesian Nash equilibria.

Definition The strategy profile $s^*(\theta) = \left(s_1^*(\theta_1), ..., s_n^*(\theta_n)\right)$ is a *dominant strategy equilibrium* of mechanism $\Gamma = (S, g)$ if for all i and all $\theta_i \in \Theta_i$,

$$u_i\left(g\left(s_i^*(\theta_i), s_{-i}\right), \theta_i\right) \geq u_i\left(g\left(s_i, s_{-i}\right), \theta_i\right)$$

for all $s_i \in S_i$, for all $s_{-i} \in S_{-i}$.

Theorem (Revelation Principle) Let $\Gamma = (S, g)$ be a mechanism that implements the social choice function $f(\cdot)$ in dominant strategy equilibrium. Then the direct mechanism (Θ, f) implements f.

Proof (Due to Gibbard). Let $s^* (\theta) = \left(s_1^* (\theta_1), ..., s_n^* (\theta_n) \right)$ be an $n-$tuple of dominant messages for (S, g) under θ, for any $\theta \in \Theta$. By the definition of implementation,

$$g (s^* (\theta)) = f (\theta), \text{ for all } \theta. \tag{1}$$

In fact

$$\Gamma^* = (\Theta, f)$$

is a direct mechanism. Next we want to show that the mechanism implements f as an equilibrium in dominant strategies. Since s^* constitutes a dominant strategy in Γ, we have

$$u_i (g (s_i^* (\theta_i), s_{-i}), \theta_i) \geq u_i (g (s_i, s_{-i}), \theta_i), \text{ for all } s_{-i}, \text{ for all } s_i.$$

In particular,

$$u_i (g (s^* (\theta)), \theta_i) \geq u_i \left(g \left(s^* \left(\theta_i', \theta_{-i} \right) \right), \theta_i \right), \text{ for all } \theta_{-i}, \text{ for all } \theta_i, \theta_i'.$$

By (1),

$$u_i\left(f\left(\theta\right),\theta_i\right)\geq u_i\left(f\left(\theta_i',\theta_{-i}\right),\theta_i\right),\text{ for all }\theta_{-i},\text{ for all }\theta_i,\theta_i'.$$

Thus truthful announcement forms a dominant strategy with (Θ, f).

- Dominant strategy implementation is a very robust mode of implementation in terms of strategies and in informational requirements as the designer doesn't need to know $p(\cdot)$ for the successful implementation. How far do we et with dominant strategy implementation?

Definition The social choice function $f(\cdot)$ is *dictatorial* if there is an agent i such that for all $\theta \in \Theta$,

$$f\left(\theta\right)\in \arg\max_{a\in A} u_i\left(a,\theta_i\right).$$

- Next, we can state the celebrated result by Gibbard and Satterthwaite

Theorem (Gibbard-Satterhwaite) Suppose that A contains at least three elements, that Θ contains all preference combinations of the players, and that f is onto A. Then the social choice function is implementable in dominant strategies if and only if it is dictatorial.

- f is onto A if it is neutral, i.e. the *names* of the candidates do not matter.

Quasi-Linear Private Values Environments

Example: Second price auction (Vickrey) There is a seller with a single good and n buyers with valuations (Θ_i) for the good. Buyers submit bids without knowing the bid of the other people in the auction. The highest bidder wins, but the price paid is the second highest bid. We claim that truthful bidding is a dominant strategy for buyer i with valuation θ_i. Given the other buyers bids $(b_j)_{j \neq i}$, his payoff is

$$\begin{cases} \theta_i - \max_{j \neq i}\{b_j\} & \text{if } \theta_i > \max_{j \neq i}\{b_j\}, \\ 0 & \text{if } \theta_i \leq \max_{j \neq i}\{b_j\}. \end{cases}$$

Overbidding does not affect the payoff when he wins but makes it less likely to win. Since payoff is positive whenever θ_i wins, the deviation is not no matter what the other buyers bid.

Underbidding does not affect the payoff when he wins but makes it more likely to win under those circumstances when winning would not be profitable. Thus the deviation is not no matter what the other buyers bid.

- The second price auction is strategically equivalent the English auction, an ascending auction where bids are raised contiuously until only one bidder is left. This bidder wins and pays his final bid.

- The underlying idea - that the incentive compatibility can be guaranteed if the transfer is not contingent on individual's own announcement - is powerful, and generalized by Groves and Clarke (and Arrow).

- Let there be set A of social actions. A social planner chooses an action a in A and a transfer t_i from each agent $i \in I$.

- Under quasilinearity assumption, agent i's payoff function is

$$u_i(a, \theta_i) - t_i.$$

- Since payoffs are linear in t_i, the first best (efficient) allocations (assume unique) for groups I and $I - \{i\}$ under $\theta = (\theta_i, \theta_{-i})$ are denoted by

$$\begin{aligned} a^*(\theta) &= \arg\max_a \sum_{i=1}^{n} u_i(a, \theta_i), \\ a^*(\theta_{-i}) &= \arg\max_a \sum_{j \neq i} u_j\left(a, \theta_j\right). \end{aligned}$$

- Then a generalization of the Vickrey auctions can be formulated:

Theorem (Vickrey-Groves-Clarke) For each i, construct a transfer scheme $t_i(\cdot)$ such that for all θ,

$$t_i(\theta) = \sum_{j \neq i} u_j\left(a^*(\theta_{-i}), \theta_j\right) - \sum_{j \neq i} u_j\left(a^*(\theta), \theta_j\right). \qquad (2)$$

Then the social choice function

$$f(\theta) = (a^*(\theta), t_1(\theta), ..., t_n(\theta)), \text{ for all } \theta \in \Theta,$$

is truthfully implementable in dominant strategies

Proof Truth is a dominant strategy for agent i if, for all θ_i, θ_i', and θ_{-i},

$$u_i(a^*(\theta_i, \theta_{-i}), \theta_i) - t_i(\theta_i, \theta_{-i}) \geq u_i\left(a^*\left(\theta_i', \theta_{-i}\right), \theta_i\right) - t_i\left(\theta_i', \theta_{-i}\right)$$

Substituting (2) for $t_i(\theta_i, \theta_{-i})$ and $t_i(\theta_i', \theta_{-i})$ gives

$$\sum_{j=1}^{n} u_j\left(a^*(\theta_i, \theta_{-i}), \theta_j\right) \geq \sum_{j=1}^{n} u_j\left(a^*\left(\theta_i', \theta_{-i}\right), \theta_j\right),$$

which holds true since $a^*(\cdot)$ is the efficient policy. Thus, $f(\cdot)$ is truthfully implementable in dominant strategies.

- This is often referred to as the *pivotal mechanism*. It is *ex-post efficient* but it typically fails *budget balance*:

  $t_i(\theta) \geq 0$ for all i and for all θ with strict inequality if $a^*(\theta) \neq a^*(\theta_{-i})$.

- There is no dominant strategy implementable mechanism that satisfies budget balance always.

- An implication: efficiency considerations imply that public services should *not* be free.

2 Bayesian Mechanisms

- In many cases, implementation in dominant strategy equilibria is too demanding.

- Alternative: not dominant strategies but equilibrium strategies.

- Note that with a single agent, these two concepts are equivalent.

- In the context of incomplete information, Bayes Nash or perfect Bayesian equilibria are the appropriate equilibrium concepts.

- Consider the following general question: Given the underlying physical structure (I, A, u, Θ, p), what outcome functions can be induced in a perfect Bayesian equilibrium of *any* extensive form game?

- A priori this is a difficult question since the class of all extensive form games is enormous. One could not possibly go through all of them.

- The next result simplifies the problem remarkably.

- Recall that the strategy profile $s^* = \left(s_1^*(\theta_1), ..., s_n^*(\theta_n)\right)$ is a Bayesian Nash equilibrium of mechanism $\Gamma = (S_1, .., S_n, g)$ if for all i and all $\theta \in \Theta$:

$$\sum_{\theta_{-i}} p(\theta_{-i} : \theta_i)\, u_i\left(g\left(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})\right), \theta\right) \geq \sum_{\theta_{-i}} p(\theta_{-i} : \theta_i)\, u_i\left(g\left(s_i, s_{-i}^*(\theta_{-i})\right), \theta\right.$$

for all s_i.

Definition An outcome function $f : \Theta \rightarrow A$ is Bayesian incentive compatible if truthful announcement forms a Bayes Nash equilibrium of a direct mechanism (Θ, f). That is, for all i and all $\theta_i, \theta_i' \in \Theta_i$,

$$\sum_{\theta_{-i}} p\left(\theta_{-i} : \theta_i\right) u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta\right) \geq \sum_{\theta_{-i}} p\left(\theta_{-i} : \theta_i\right) u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta\right).$$

Theorem (Bayesian Revelation Principle) An outcome function $f : \Theta \rightarrow A$ can be implemented in Bayes Nash equilibrium if and only if f is Bayesian incentive compatible.

Proof: "Only if" Let mchanism $(S_1, .., S_n, g)$ implement f in Bayes nash equilibrium. That is, there is the Bayes Nash equilibrium strategy $s^* : \Theta \rightarrow S$

such that

$$f = g \circ s^*,$$

i.e.

$$f(\theta) = g(s^*(\theta)), \text{ for all } \theta. \tag{3}$$

We want to show that there is a Bayes Nash equilibrium of the direct mechanism (Θ, f) that implements f.

By the definition of Bayes Nash equilibrium, for all $\theta \in \Theta$,

$$\sum_{\theta_{-i}} p(\theta) u_i \left(g \left(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}) \right), \theta \right) \geq \sum_{\theta_{-i}} p(\theta) u_i \left(g \left(s_i, s_{-i}^*(\theta_{-i}) \right), \theta \right),$$

for all s_i. In particular,

$$\sum_{\theta_{-i}} p(\theta) u_i (g(s^*(\theta)), \theta) \geq \sum_{\theta_{-i}} p(\theta) u_i \left(g \left(s^* \left(\theta_i', \theta_{-i} \right) \right), \theta \right),$$

for all θ_i'. By (3), this can be written

$$\sum_{\theta_{-i}} p(\theta)u_i\left(f\left(\theta\right), \theta\right) \geq \sum_{\theta_{-i}} p(\theta)u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta\right).$$

But this contradicts our hypothesis that s^* is a Bayes Nash equilibrium of (S, g) and, a fortiori, that outcome function f is induced in PBE of some extensive game.

"If". If f is Bayesian incentive compatible, then f is induced in Bayes Nash equilibrium of the particular strategic game (Θ, f).

- The results is actually more general: In the revious slides we argued that any perfect Bayesian equilirium of any extensive form game forms also a Bayes Nash equilibrium in th reduced strategic form of the extensive game. Thus the Revelation Principle covers also fs that are implementable via extensive form games. It genuinely suffices to focus on the "simple" direct mechanisms (cf. Myerson, 1979) if one is interested in implementing somthing in incomplete information framework.

2.1 Price Discrimination

- Let there be a monopoly (seller) with a single buyer. Consumer's (buyer) preferences are his private information.

- The monopoly faces cost $c(a)$ from production of output $a \in \mathbb{R}_+$, where $c', c'' > 0$ and $c(0) = c'(0) = 0$.

- Consumer payoff from consumption is θa where the marginal utility θ is known only by the consumer. The seller knows the ex ante the cdf P of θ on the interval $\Theta = [0, 1]$. Denote the corresponding density by p.

- The value of the outside option to the buyer is 0.

- Given the produced amount a and monetary transfer t from the buyer to the monopoly, monopoly profit is $t - c(a)$ and buyer θ's surplus is $\theta a - t$.

- It may be possible to *separate* the consumers based on his type. This opens the possibility for the seller to exercise *price discrimination* against the buyer: prices are no longer independent of the buyer types nor the amount to be bought (which is also known as nonlinear pricing).

- By the revelation principle, it suffices to focus on Bayesian incentive compatible contract $(a(\cdot), t(\cdot)) : [0, 1] \to \mathbb{R} \times \mathbb{R}$ that specifies an allocation of resources for each valuation type of the buyer. We seek to identify the contract maximizes seller's profits.

- An alternative interpretation: The monopolist faces a mass S of buyers on the continuum whose valuations are distributed according to SP.

Complete Information - First Best Solution

- Type θ is known by the monopoly, and he makes the optimal take-it-or-leave-it offer $(a(\theta), t(\theta))$ to the buyer. Thus his problem is then to maximize profits subject to the buyer's *individual rationality* (or participation) constraint:

$$\max_{a(\theta),t(\theta)} \{t(\theta) - c(a(\theta))\}$$
$$\text{s.t. } \theta a(\theta) - t(\theta) \geq 0.$$

- In the optimum $\theta a(\theta) = t(\theta)$. Thus the problem reduces to

$$\max_{a(\theta)} \{\theta a(\theta) - c(a(\theta))\} \quad (4)$$

whose solution $a^*(\theta)$ satisfies

$$c'(a^*(\theta)) = \theta, \text{ for any } \theta \in [0, 1]. \quad (5)$$

- Since c is concave, a^* is an increasing function of θ. The monopoly wants to sell more to a buyer with higher marginal utility. Since $c'(0) = 0$, we have $a^*(\theta) > 0$ for any θ

- As a consequence, the *first order price discriminatory* pricing policy is implemented: the seller extracts all the buyers' surplus. Scheme a^* is also efficient (first best).

Incomplete Information - Second Best Solution

- When information is asymmetric, the seller is no longer able to extract all the surplus from the buyer - she does not know the level at which the individual rationality is binding. First degree price discrimination is no longer feasible.

- Since the monopoly cannot make a type-conditional offer to the buyer, the best it can do is to make a *menu* of offers from which the buyer can choose from. As a consequence, the monopoly leaves the buyer some discretion over the choice of the price-output combinations.

- Denote the payoff of type $\theta \in [0,1]$ buyer who *acts* as type θ' buyer by

$$U\left(\theta' : \theta\right) = \theta a(\theta') - t(\theta').$$

- By incentive compatibility:

$$U(\theta : \theta) \geq U\left(\theta' : \theta\right), \text{ for all } \theta, \theta' \in [0,1],$$

- We identify a scheme $(a(\theta), t(\theta))_{\theta \in \Theta}$ that maximizes monopoly profits *in the class of incentive compatible* and *individually rational* schemes. Such pricing policy is called the *second degree price discrimination*, or nonlinear pricing.

- The monopoly's problem is to choose a function $(a, t) : [0, 1] \rightarrow \mathbb{R}^2_+$ (assume twice differentiable) that maximizes her expected payoff subject to incentive compatibility and individual rationality constraints.

$$\max_{(a,t):[0,1]\rightarrow\mathbb{R}^2_+} \int_0^1 \{t(\theta) - c(a(\theta))\}p(\theta)d\theta \quad (6)$$

$$\text{s.t. } U(\theta : \theta) \geq U\left(\theta' : \theta\right), \text{ for all } \theta, \theta' \in [0, 1], \quad (7)$$

$$U(\theta : \theta) \geq 0, \text{ for all } \theta \in [0, 1]. \quad (8)$$

- First we focus on incentive compatibility.

Lemma (Incentive compatibility) Allocation function (a, t) satisfies incentive compatibility if and only if, for all $\theta \in [0, 1]$,

$$\begin{aligned} a'(\theta) &\geq 0, \\ U(\theta : \theta) &= \int_0^\theta a(\theta')d\theta' - t(0). \end{aligned}$$

Proof: "Only if" Incentive compatibility constraint (7) can be written in the form

$$\theta \in \arg\max_{\theta'} U\left(\theta' : \theta\right), \text{ for all } \theta \in [0,1].$$

The necessary first and second order conditions for the maximum are

$$\begin{aligned} \frac{d}{d\theta'} U\left(\theta' : \theta\right)\Big|_{\theta=\theta'} &= \theta a'(\theta) - t'(\theta) = 0, \text{ for all } \theta \in [0,1], \quad (9) \\ \frac{d^2}{(d\theta')^2} U\left(\theta' : \theta\right)\Big|_{\theta=\theta'} &= \theta a''(\theta) - t''(\theta) \leq 0, \text{ for all } \theta \in [0,1]. \quad (10) \end{aligned}$$

Since FOC (9) must hold as identity for all θ, we can take a derivative of both sides to get

$$a'(\theta) + \theta a''(\theta) - t''(\theta) = 0.$$

Plugging this into SOC (10), we have

$$a'(\theta) \geq 0, \text{ for all } \theta \in [0, 1]. \tag{11}$$

Thus, by (9), also $t'(\theta) \geq 0$ for all θ. Consequently *any* incentive compatible pricing scheme supplies more to and extracts higher transfer from higher types. FOC (9) holds if and only if, for all $\theta \in [0, 1]$,

$$\frac{d}{d\theta} U(\theta : \theta) = a(\theta). \tag{12}$$

Equivalently, by the Envelope Theorem, for all $\theta \in [0,1]$,

$$\begin{aligned} U(\theta : \theta) &= \int_0^\theta \frac{d}{d\theta'} U\left(\theta' : \theta'\right) d\theta' - t(0) \qquad (13) \\ &= \int_0^\theta a(\theta') d\theta' - t(0). \end{aligned}$$

"If". Take θ and, say, $\theta' < \theta$. Now, by (12),

$$\begin{aligned} U(\theta : \theta) - U\left(\theta' : \theta'\right) &= \int_{\theta'}^{\theta} \frac{d}{dy} U(y : y)\, dy \\ &= \int_{\theta'}^{\theta} a(y) dy \\ &\geq \int_{\theta'}^{\theta} a(\theta') dy \\ &= (\theta - \theta') a(\theta'). \end{aligned}$$

Hence

$$\begin{aligned} U(\theta:\theta) &\geq U\left(\theta':\theta'\right)+(\theta-\theta')a(\theta') \\ &= \theta a(\theta')-t(\theta') \\ &= U\left(\theta:\theta'\right), \end{aligned}$$

as required.

- Now we turn back to the optimizing problem of the seller. Since $a \geq 0$ we have

$$\frac{d}{d\theta}U(\theta:\theta)=\frac{d}{d\theta}\left(\int_0^\theta a(\theta')d\theta'-t(0)\right)=a(\theta)\geq 0,$$

thus $U(\theta:\theta)$ is an increasing function of θ. This implies that the individual rationality constraint can bind only at $\theta = 0$. Thus constraint (8) can be replaced with a single condition

$$-t(0) \geq 0. \tag{14}$$

- By integrating by parts, we have (recall that p is the density of cdf P and $P(1) = 1$)

$$
\begin{aligned}
& \int_0^1 \left(\int_0^\theta a(\theta')d\theta' \right) p(\theta)d\theta \\
= & \left(\int_0^1 a(\theta)d\theta \right) \left(\int_0^1 p(\theta)d\theta \right) - \int_0^1 a(\theta) \left(\int_0^\theta p(\theta')d\theta' \right) d\theta \\
= & \int_0^1 a(\theta)(1 - P(\theta))d\theta. \qquad (15)
\end{aligned}
$$

Thus, by using (13) and (15), the ex ante expected payoff of the buyer is

$$\begin{aligned}\int_0^1 U(\theta:\theta)\,p(\theta)d\theta &= \int_0^1 \left(\int_0^\theta a(\theta')d\theta' - t(0)\right) p(\theta)d\theta \qquad (16)\\ &= \int_0^1 a(\theta)\left[1-P(\theta)-t(0)p(\theta)\right]d\theta \\ &= \int_0^1 a(\theta)\left(\frac{1-P(\theta)}{p(\theta)}\right)p(\theta)d\theta - t(0).\end{aligned}$$

- Ratio $p(\theta)/(1-P(\theta))$ is the *hazard function* of cdf P. Hazard functions are well studied in the literature. In most cases (e.g. when P is normal, uniform, exponential, Poisson...) the hazard function is *increasing*. We *assume* that this holds also for P:

$$\frac{d}{d\theta}\left(\frac{1-P(\theta)}{p(\theta)}\right) < 0, \text{ for all } \theta.$$

Term

$$J(\theta) = \theta - \frac{1 - P(\theta)}{p(\theta)}$$

is called agent's i's *virtual payoff.* With increasing hazard rate, the virtual payoff is an increasing function of θ:

$$J'(\theta) > 0, \text{ for all } \theta.$$

Note that there might well be $\theta' > 0$ such that $J(\theta') = 0$. The concept of virtual payoff will prove important.

- By (16), the expected revenue to the monopoly can be written

$$\begin{aligned} \int_0^1 t(\theta)p(\theta)d\theta &= \int_0^1 \left[\theta a(\theta) - U(\theta : \theta)\right] p(\theta)d\theta + t(0) \quad (17) \\ &= \int_0^1 a(\theta)J(\theta)p(\theta)d\theta + t(0). \end{aligned}$$

- Thus, given function J, the optimization problem (6) reduces to finding $a : [0,1] \to \mathbb{R}_+$ that solves

$$\max_{\substack{a:[0,1]\to\mathbb{R}_+ \\ t(0)\in\mathbb{R}.}} \left\{ \int_0^1 \left[a(\theta)J(\theta) - c(a(\theta))\right] p(\theta) d\theta + t(0) \right\}, \quad (18)$$

s.t. $a'(\theta) \geq 0$ for all θ, and $t(0) \leq 0$.

At the optimum $t(0) = 0$. Try function $a : [0,1] \to \mathbb{R}_+$ that maximizes term $a(\theta)J(\theta) - c(a(\theta))$ pointwisely. If such a is nondecreasing, then it solves (18).

- Given θ, solve the maximization problem

$$\max_{a\geq 0} \left\{ aJ(\theta) - c(a) \right\}. \quad (19)$$

Identify $a^{**}(\theta) \geq 0$ that satisfies the first order condition

$$J(\theta) = c'(a^{**}(\theta)).$$

Repeating the procedure for all $\theta \in [0, 1]$ one can identify $\underline{\theta} \in [0, 1]$ such that $a^{**}(\underline{\theta}) = 0$. Thus the solution to the problem (19) draws function $a^{**} : [0, 1] \to \mathbb{R}_+$ such that

$$\begin{aligned} J(\theta) &= c'(a^{**}(\theta)), \text{ for all } \theta \in (\underline{\theta}, 1], \\ a^{**}(\theta) &= 0, \text{ for all } \theta \in [0, \underline{\theta}], \end{aligned}$$

Function a^{**} constitutes the solution to the general optimization problem (18) if it is nondecreasing. By totally differentiating, and noting that c is differentiable and convex, we get

$$\frac{da^{**}(\theta)}{d\theta} = \frac{J'(\theta)}{c''(a^{**}(\theta))} > 0, \text{ for all } \theta \in [\underline{\theta}, 1].$$

Thus the necessary condition (**??**) is met by a^{**}. We conclude that a^{**} *is the seller's expected payoff maximizing scheme* under incomplete information.

- The difference between the first best (4) and the second best (19) solutions under the former marginal cost equals marginal utility θ whereas under the latter marginal cost equals virtual utility $J(\theta)$. Since $J(\theta) < \theta$ for all θ, the second best solution *reduces* production and is hence *inefficient*. Also the payoff for the monopolist is lower under each θ.

Parametrized Example Let $c(a) = ca^2$ and $P(\theta) = \theta$ for all $\theta \in [0, 1]$. Then

$$
\begin{aligned}
a^{**}(\theta) &= \begin{cases} \frac{1}{c}(\theta - \frac{1}{2}), & \text{for all } \theta \in (\frac{1}{2}, 1], \\ 0, & \text{for all } \theta \in [0, \frac{1}{2}], \end{cases} \\
\theta a^{**}(\theta) - t^{**}(\theta) &= \frac{1}{2c}(\theta - 1)\theta + \frac{1}{8c} > 0, \text{ for all } \theta \in (\frac{1}{2}, 1].
\end{aligned}
$$

$n-$**buyers**

- Let there be n buyers of a single good of a single seller. A trading mechanism, or *an auction*, specifies the rules of the trading procedure. The seller wants to formulate the most profitable auction mechanism.

- Buyer i's reservation valuation θ_i is his private information, and distributed on $\left[\underline{\theta}_i, \overline{\theta}_i\right]$ according to cdf P (twice continuously differentiable)..

- By the revelation principle, it suffices to focus on Bayesian incentive compatible outcome functions $(a, t)(\cdot)$ where the allocation $a(\theta)$ specifies which of the buyers, if any, receives the good under profile $\theta = (\theta_1, ..., \theta_n)$, and $t(\theta)$ specifies the transfers from the buyers to the seller. That is,.$a_i(\theta) \in \{0, 1\}$ and $\sum_i a_i(\theta) \leq 1$, and $t_i(\theta) \in \mathbb{R}$ for all θ.

- We turn next to the question of what kinds of outcome functions (a, t) are incentive compatible.

- We assume that the utilities take the following linear form:

$$\theta_i a_i - t_i.$$

- Bayesian incentive compatibility requires that all buyers announce their type to (a, t) truthfully if others do so too.

- To check this, denote an arbitrary announcement by player i by θ_i'.

- Denote $p_{-i}(\theta_{-i}) = \Pi_{j \neq i} p_j(\theta_j)$ for any i. Then the expected consumption payoff and transfer given true type θ_i and announcement θ_i' as

$$\begin{aligned} \overline{a}_i\left(\theta_i'\right) &= \int_{\theta_{-i}} p_{-i}(\theta_{-i}) a_i\left(\theta_i', \theta_{-i}\right) d\theta_{-i}, \\ \overline{t}_i\left(\theta_i'\right) &= \int_{\theta_{-i}} p_{-i}(\theta_{-i}) t_i\left(\theta_i', \theta_{-i}\right) d\theta_{-i}. \end{aligned}$$

- Because of the linearity of the payoff functions, we may now write the expected payoff from announcing θ_i' while of type θ_i.

$$\theta_i \overline{a}_i \left(\theta_i'\right) - \overline{t}_i \left(\theta_i'\right).$$

Define also the expected payoff of θ_i when he behaves as θ_i', for any i,

$$U_i \left(\theta_i' : \theta_i\right) = \theta_i \overline{a}_i \left(\theta_i'\right) - \overline{t}_i \left(\theta_i'\right).$$

- Now we are in a position to characterize the inventive compatible mechanisms.

Lemma (Bayesian incentive compatible outcomes) Allocation function $(a, t)(\cdot)$ is Bayesian incentive compatible if and only if for all $i \in I$ and for all θ_i,

$$\overline{a}_i (\theta_i) \text{ is nondecreasing,}$$

$$U_i (\theta_i : \theta_i) = \int_{\underline{\theta}_i}^{\theta_i} \overline{a}_i (y)\, dy - U_i(\underline{\theta}_i) \tag{20}$$

- The proof is essentially the same as in the price discrimination case.

Theorem (Revenue Equivalence): Assume that types θ_i's are independently distributed and that $p_i(\theta_i) > 0$ for all $\theta_i \in \left[\underline{\theta}_i, \overline{\theta}_i\right]$. If two auction mechanisms $(a,t)(\cdot)$ and $(a',t')(\cdot)$ satisfy $a_i(\theta) = a'_i(\theta)$ for all θ and all , and the utility to buyer $\underline{\theta}_i$ is the same in the two auctions for all i, then the expected revenue to the seller is the same in these two auctions.

Proof We rely on the characterization above of the players' utilities in incentive compatible mechanisms. The expected revenue to the seller is the sum of the expected transfers from the buyers. Hence we calculate

$$\begin{aligned}
&\int_{\theta_i} \bar{t}_i(\theta_i)\, p_i(\theta_i) d\theta_i \\
&= \int_{\underline{\theta}_i}^{\overline{\theta}_i} \left[\theta_i \overline{a}_i(\theta_i) - U_i(\theta_i : \theta_i)\, \overline{a}_i(\theta_i)\right] p_i(\theta_i)\, d\theta_i - U_i(\underline{\theta}_i) \\
&= \int_{\underline{\theta}_i}^{\overline{\theta}_i} \overline{a}_i(\theta_i) \left[\theta_i - \left(\frac{1 - P_i(\theta_i)}{p_i(\theta_i)}\right)\right] p_i(\theta_i)\, d\theta_i - U_i(\underline{\theta}_i)
\end{aligned}$$

Opening $\overline{a}_i(\theta_i)$ we have

$$\begin{aligned}
&\int_{\theta_i} \bar{t}_i(\theta_i)\, p_i(\theta_i) d\theta_i \\
&= \int_{\underline{\theta}_1}^{\overline{\theta}_1} \cdots \int_{\underline{\theta}_n}^{\overline{\theta}_n} a_i(\theta) \left(\theta_i - \frac{1 - P_i(\theta_i)}{p_i(\theta_i)}\right) \Pi_{j=1}^{n} p_j(\theta_j)\, d\theta_1 \cdots d\theta_n - U_i(\underline{\theta}_i).
\end{aligned}$$

By summing over i, we get the expected revenue to the seller which depends only on $a_i(\theta)$'s and $U_i(\underline{\theta}_i)$'s.

- Consider the implications of this theorem. If we know for example that an auction is efficient in the sense that the bidder with the highest valuation gets the object with probability 1 and the bidder with the lowest possible type gets a zero payoff, then we know that the expected payoffs to all parties are the same as in the dominant strategy second price auction. Hence it will be easy to solve for equilibria in alternative auctions using this technique.

War of Attrition

- A war of attrition is a game of pure timing: Each of the players has a choice of either staying in the game or dropping out at each instant (animal conflict, hence the name).

- Consider first the case of two players $1, 2$.
- If i is the last player remaining, she is awarded a price worth θ_i.
- The θ_i's are assumed to be drawn independently from the same cdf $P(\theta_i)$. Let $p(\cdot)$ denote the corresponding density.
- The cost of staying in the game is 1 per unit of time.
- If $i = 1, 2$ stays until time b_i, then the cost of both players is given by $\min\{b_1, b_2\}$ (only the one staying longer gets the payoff.)

- Let type θ_i of player i stay in the game until time $b(\theta_i)$ (continuously differentiable).

- In a symmetric equilibrium the lowest type will drop out immediately and receive a payoff of 0 and that the outcome of the game is efficient.

- The revenue equivalence theorem tells us that the expected amount of resources spent by each of the players must be the same as in the second price auction..

- Each player wins with probability $P(\theta_i)$.

- The expected value of the other player j's bid conditional on $\theta_j < \theta_i$ is

$$E_{\theta_j}\left[\theta_j : \theta_j < \theta_i\right] = \frac{\int_{\underline{\theta}}^{\theta_i} yp\left(y\right)dy}{P\left(\theta_i\right)},$$

$$E_{\theta_j}\left[b(\theta_j) : \theta_j < \theta_i\right] = \frac{\int_{\underline{\theta}}^{\theta_i} b\left(y\right)p\left(y\right)dy}{P\left(\theta_i\right)}$$

- By equating the expected payments in the war of attrition and the second price auction, we get:

  $$(1 - P\left(\theta_i\right))\,b\left(\theta_i\right) + P\left(\theta_i\right)E[b(\theta_j) : \theta_j < \theta_i] = P\left(\theta_i\right)E[\theta_j : \theta_j < \theta_i].$$

  Thus

  $$b\left(\theta_i\right) - b\left(\theta_i\right)P\left(\theta_i\right) + \int_{\underline{\theta}}^{\theta_i} b\left(y\right)p\left(y\right)dy = \int_{\underline{\theta}}^{\theta_i} yp\left(y\right)dy.$$

Since this holds for all θ_i, we may differentiate to get:

$$b'(\theta_i) - P(\theta_i)b'(\theta_i) - p(\theta_i)b(\theta_i) + b(\theta_i)\,p(\theta_i) = \theta_i p(\theta_i)$$

or

$$b'(\theta_i) = \frac{\theta_i p(\theta_i)}{1 - P(\theta_i)},$$

or

$$b(\theta_i) = \int_{\underline{\theta}}^{\theta_i} y \left(\frac{p(y)}{1 - P(y)} \right) dy.$$

Optimal Auctions

- In the proof of the revenue equivalence theorem, it is shown that the expected revenue from any incentive compatible mechanism depends on the function $a(\theta) = (a_1(\theta), ..., a_n(\theta))$ only. I.e. the expected revenue of the seller is

$$\int_{\underline{\theta}_1}^{\overline{\theta}_1} \cdots \int_{\underline{\theta}_n}^{\overline{\theta}_n} \Pi_{j=1}^n p_j(\theta_j) \sum_{i=1}^n a_i(\theta) \left(\theta_i - \frac{1 - P_i(\theta_i)}{p_i(\theta_i)}\right) d\theta_1 \cdots d\theta_n + \sum_{\iota=1}^n U_i(\underline{\theta}_i).$$

If we want to find the auction that maximizes the revenue to the seller, then we must maximize this expression by choosing the functions a_i appropriately.

- a_i's must all be non-negative and sum up to at most 1 (if the sum is zero, then the seller keeps the object).

- As in (12)

$$\frac{d}{d\theta} U_i(\theta_i : \theta_i) = \overline{a}_i(\theta_i).$$

Incentive compatibility requires that each $\overline{a}_i$ be nondecreasing (= the *monotonicity constraint*). Hence $U_i(\cdot)$ is nondecreasing.

- Voluntary participation by the bidders in the auction requires that $U_i(\theta_i) \geq 0$ for all θ_i. Thus it suffices that $U_i(\underline{\theta}_i) \geq 0$.

- The first obvious point to notice is that it is optimal for the seller to set $U_i(\underline{\theta}_i) = 0$ for all i.

- Second, it is optimal to put $a_i = 1$ if for all j,

$$J_i(\theta_i) = \left(\theta_i - \frac{1 - P_i(\theta_i)}{p_i(\theta_i)}\right) > \max\left\{0, \left(\theta_j - \frac{1 - P_j\left(\theta_j\right)}{p_j\left(\theta_j\right)}\right)\right\},$$

if this is consistent with the monotonicity constraint.

- If the monotone hazard rate condition holds, then J_i is a nondecreasing function. In that case a_i's monotonicity requirement is clearly satisfied.

- Observe that the optimal auctions do not necessarily allocate the object to the bidder who values it the most, and in some cases, the seller may fail to sell altogether (if $\max_i J_i(\theta_i) < 0$).

Bilateral Bargaining

- The adverse selection problem relied on the assumption that only very simplistic contractual arrangements were possible, i.e. that transactions could not be made contingent with observations when the uncertainty is removed. If they would have been allowed, the adverse selection problem could be solved with a more simple contractual arrangement: a warranty.

- In general, any asymmetric information problem with *ex post* observable types can be solved with an arrangement where the *ex ante* incentives are reconciled with the ex post occurrences. This would typically be the case under common values.

- For example, there is a hard enough penalty for an observable lie that would prevent any such attempts in the ex ante stage. Managing the incentives appropriately at the ex ante stage would guarantee efficiency.

- The "Coase Theorem": In the absence of transactions costs, an efficient contractual arrangement would always emerge in free bargaining.

- However, when the asymmetry of information concerns individual's own preferences, efficiency cannot be reached with *any* contractual arrangement.

- Let there be a seller and a buyer of an indivisible object.

- Both of the players' valuations are drawn from $[0, 1]$ interval with cdf's P_s and P_b and densities p_s and p_b, resp..

- Let there be a trading mechanism $(a, t) : [0, 1] \times [0, 1] \longrightarrow \{0, 1\} \times \mathbb{R}$ such that, under $(a_i, t_i)(\theta_s, \theta_b)$ the seller's payoff is

$$t(\theta_s, \theta_b) - a(\theta_s, \theta_b)\theta_a$$

  and the buyer's payoff is

$$a(\theta_s, \theta_b)\theta_b - t(\theta_s, \theta_b).$$

- Denote, for $i = b, s$, and $j \neq i$,

$$\overline{a}_i(\theta_i) = \int_0^1 a(\theta_s, \theta_b) p_j(\theta_j) d\theta_j,$$
$$\overline{t}_i(\theta_i) = \int_0^1 t(\theta_s, \theta_b) p_j(\theta_j) d\theta_j,$$

and

$$U_s\left(\theta_s : \theta_s'\right) = \overline{t}_s(\theta_s') - \overline{a}_s(\theta_s')\theta_s,$$
$$U_b\left(\theta_b : \theta_b'\right) = \overline{a}_b(\theta_b')\theta_b - \overline{t}_b(\theta_b').$$

- Incentive compatibility implies

$$U_s(\theta_s : \theta_s) \geq U_s\left(\theta_s : \theta_s'\right), \text{ for all } \theta_s, \theta_s',$$
$$U_b(\theta_b : \theta_b) \geq U_b\left(\theta_b : \theta_b'\right), \text{ for all } \theta_b, \theta_b'.$$

- Individual rationality means

$$\begin{aligned} U_s(\theta_s : \theta_s) &\geq 0, \text{ for all } \theta_s, \\ U_b\left(\theta_b : \theta_b'\right) &\geq 0, \text{ for all } \theta_b. \end{aligned}$$

- Efficiency requires

$$a = \begin{cases} 1 \text{ if } \theta_b > \theta_s, \\ 0 \text{ if } \theta_b < \theta_s. \end{cases}$$

Theorem (Myerson-Satterhwaite): There is no incentive compatible, individually rational, and efficient mechanism.

Proof: As in (20), incentive compatibility implies

$$U_s(\theta_s : \theta_s) = U_s(1:1) - \int_{\theta_b}^{1} \overline{a}_s(y)\, dy,$$

$$U_b(\theta_b : \theta_b) = \int_{0}^{\theta_b} \overline{a}_b(y)\, dy + U_b(0:0).$$

As with the one buyer case,

$$\begin{aligned} & \int_0^1 t_s(\theta_s)\, dP_s(\theta_s) \\ = & \int_0^1 \left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)} \right) \overline{a}_s(\theta_s) + \theta_s \overline{a}_s(\theta_s) \right] dP_s(\theta_s) + U_s(1) \\ = & \int_0^1 \overline{a}_s(\theta_s) \left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)} \right) + \theta_s \right] dP_s(\theta_s) + U_s(1). \end{aligned}$$

and

$$
\begin{aligned}
& \int_0^1 t_b\left(\theta_b\right) dP_b(\theta_b) \\
= \; & \int_0^1 \left[\theta_b \overline{a}_b\left(\theta_b\right) - \left(\frac{1 - P_b\left(\theta_b\right)}{p_b\left(\theta_b\right)}\right) \overline{a}_b\left(\theta_b\right)\right] dP_b\left(\theta_b\right) - U_b\left(0\right) \\
= \; & \int_0^1 \overline{a}_b\left(\theta_b\right) \left[\theta_b - \left(\frac{1 - P_b\left(\theta_b\right)}{p_b\left(\theta_b\right)}\right)\right] dP_b\left(\theta_b\right) - U_b\left(0\right)
\end{aligned}
$$

Since money is not thrown away,

$$
\begin{aligned}
0 &= \int_0^1 \overline{t}_s(\theta_s)\, dP_s(\theta_s) - \int_0^1 \overline{t}_b(\theta_b)\, dP_b(\theta_b) \qquad (21) \\
&= \int_0^1 \overline{a}_s(\theta_s)\left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)}\right) + \theta_s\right] dP_s(\theta_s) + U_s(1) \\
&\quad - \int_0^1 \overline{a}_b(\theta_b)\left[\theta_b - \left(\frac{1 - P_b(\theta_b)}{p_b(\theta_b)}\right)\right] dP_b(\theta_b) + U_b(0) \\
&= \int_0^1 \int_0^1 a(\theta_s, \theta_b)\left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)}\right) + \left(\frac{1 - P_b(\theta_b)}{p_b(\theta_b)}\right) + \theta_s - \theta_b\right] dP_b(\theta_b)\, dP_s(\theta_s) \\
&\quad + U_s(1) + U_b(0)
\end{aligned}
$$

By efficiency, this is equivalent to

$$\int_0^1 \int_{\theta_s}^1 \left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)} \right) + \left(\frac{1 - P_b(\theta_b)}{p_b(\theta_b)} \right) + \theta_s - \theta_b \right] dP_b(\theta_b)\, dP_s(\theta_s)$$
$$+ U_s(1) + U_b(0)$$

$$> \int_0^1 \int_{\theta_s}^1 \left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)} \right) + \left(\frac{1 - P_b(\theta_b)}{p_b(\theta_b)} \right) \right] dP_b(\theta_b)\, dP_s(\theta_s)$$
$$+ U_s(1:1) + U_b(0:0).$$

By individual rationality, this is at least

$$\int_0^1 \int_{\theta_s}^1 \left[\left(\frac{P_s(\theta_s)}{p_s(\theta_s)} \right) + \left(\frac{1 - P_b(\theta_b)}{p_b(\theta_b)} \right) \right] dP_b(\theta_b)\, dP_s(\theta_s)$$

Since the terms inside the brackets are positive, so is the integral. But this violates (21).

- Thus incentives and participation are in a fundamental conflict with efficiency.

  $\Longrightarrow$ The Coase Theorem fails (unless incentive constraint is counted as a transaction cost).

- Note that, by the revelation principle, it does *not* help to allow an additional round of bargaining once an inefficient outcome realizes (why??).

3 Efficiency concepts

- Suppose a benevolent social designer wants what is best for the society: he wants to implement an efficient decision that cannot be changed for anything more efficient.

- If information is incomplete, and dispersed among the individuals in the society, then the designer cannot simply implement the efficient allocation. He has to use an incentive compatible (and, perhaps, individually rational) mechanism for that purpose.

- Thus the designer's the choice set is the class of incentive compatible mechanisms. Which mechanism should a benevolent planner choose?

- There is an immediate choice problem. If the types of the players are their private information, then it is unclear *which* types payoffs the designer is maximizing. The whole concept of efficiency (or Pareto optimality) is no longer obviously defined.

- Homström and Myerson (1983) lay down a taxonomy of different efficiency concepts, depending on at which time the efficiency (Prateto) comparisons are conducted.

Ex ante efficiency An incentive compatible direct mechanism $g : \Theta \to A$ is ex ante efficient if there is no incentive compatible mechanism $f : \Theta \to A$ such that the expected value of b is at least that of a to all for all $i \in I$, and higher for some i.

Interim efficiency An incentive compatible direct mechanism $g : \Theta \to A$ is interim efficient if there is no incentive compatible mechanism $f : \Theta \to A$ such that the expected value of f conditional on θ_i is at least that of g to all $\theta_i \in \Theta_i$ for all $i \in I$, and higher for some θ_i.

Ex post efficiency An incentive compatible direct mechanism $g : \Theta \to A$ is interim efficient if there is no incentive compatible mechanism $f : \Theta \to A$ such that the value of $f(\theta)$ is at least that of $g(\theta)$ to all $\theta_i \in \Theta_i$ for all $i \in I$, and higher for some θ_i.

- If mechanism is also individually rational, then that constraint is imposed on the feasible mechansims in the definitions of ex ante and interim efficiency (but what about ex post?).

- Mechanism is ex ante efficient if it solves, for some vector $(\lambda_i)_{i \in I}$ of real numbers, the problem

$$\begin{aligned}
&\max \sum_{\theta} \lambda_i u_i(g(\theta), \theta) p(\theta) \\
&\text{s.t.} \quad \sum_{\theta} u_i(g(\theta), \theta) p(\theta) \geq \sum_{\theta} u_i(g(\theta_i', \theta_{-i}), \theta) p(\theta) \ , \\
&\quad \text{for all } \theta_i, \theta_i' \in \Theta_i, \ \text{for all } i \in I.
\end{aligned}$$

- Mechanism is interim efficient if it solves, for some vector $(\lambda(\theta_i))_{\theta_i \in \Theta_i, i \in I}$ of real numbers, the problem

$$\begin{aligned}
&\max \sum_{\theta} \lambda(\theta_i) u_i(g(\theta), \theta) p(\theta) \\
&\text{s.t.} \quad \sum_{\theta} u_i(g(\theta), \theta) p(\theta) \geq \sum_{\theta} u_i(g(\theta_i', \theta_{-i}), \theta) p(\theta) \ , \\
&\quad \text{for all } \theta_i, \theta_i' \in \Theta_i, \ \text{for all } i \in I \ .
\end{aligned}$$

- Note that incentive constraints are linear in $u_i(a, \theta)$'s.

- Thus the question of finding an incentive efficient mechanism reduces to a linear programming problem which are, in general "easily" solvable.

- Ex ante efficient mechanism is interim efficient but not vice versa (why?).

3.1 Moral Hazard (Holmström 1979)

- In many principal-agent situations the question is not of the private information of the agent but rather to control the agent.

- In a complete information - complete contracts world, this could be easily achieved by imposing the agent a huge penalty in the case of a failure.

- In the incomplete information - incomplete contracts world, the solution is not as simple: the principle needs to design scheme that gives enough incentives for the agent to act appropriately. This is not without costs, however.

- Let the agent choose an effort $a \in \{0, 1\}$ which causes him a private cost a.

- After the agent has chosen a, a revenue $r \in \mathbb{R}$ realizes. The principal may condition the payment t to the agent only on the realized payoff r. Transfer t generates utility $v(t)$ to he agent, where v is concave and $v(0) = 0$.

- The realized revenue is conditional on the action of the agent. Denote by $p(\cdot : a)$ the (discrete, for simplicity) probability distribution over possible revenues after effort a.

- The expected payoff to the pricipal from incentive scheme $t(\cdot)$ is

$$\sum_r (r - t(r))p(r : a)$$

  and for the agent

$$\sum_r (v(t(r)) - a)p(r : a).$$

- The agent has an outside option of value 0.

Contractible a

- Suppose first that the principal can force the agent to choose the action he wants e.g. by restricting the agent's action space.Then he only needs to worry about agent's participation. The principal maximizes his payoff subject to the individual rationality consraint.

$$\max_{t(\cdot)} \sum_r (r - t(r))p\,(r : a)$$
$$\text{s.t.} \quad \sum_r v(t(r))p\,(r : a) \geq a\ .$$

- Construct a Lagrangian

$$\max_{t(\cdot)} \sum_r -t(r)p(r : a) + \lambda \left[\sum_r v(t(r)) - a\right].$$

- Solving this for $t(r)$'s gives,

$$-p(r:a) + \lambda v'(t(r))p(r:a), \text{ for any } r,$$

or

$$\frac{1}{v'(t(r))} = \lambda, \text{ for any } r.$$

At the optimum, the individual rationality constraint binds (why?) and hence $\lambda > 0$.

- If the agent is risk averse, so that v is strictly convex, then the optimal contract is constant: the agent takes no risks.

Non-contractible a

- Suppose now that the principal cannot monitor the agents effort, nor verify it ex post. Now the agent muct be provided incentives to *choose* the desired action. Suppose that the desired action is $a = 1$.

- The principal maximizes his payoff subject to incentive compatibility and individual rationality constraints, and subject to $a = 1$ requirement. Hence

$$\min_{t(\cdot)} \sum_r (r - t(r)) p\,(r : 1)$$
s.t.
$$\sum_r v(t(r)) p\,(r : 1) - 1 \geq \sum_r v(t(r)) p\,(r : 0)\,,$$
$$\sum_r v(t(r)) p\,(r : 1) - 1 \geq 0\,.$$

- Construct a Lagrangian

$$\max_{t(\cdot)} \sum_r -t(r) p\,(r : 1) + \lambda \left[\sum_r v(t(r)) p\,(r : 1) - a\right] + \mu \left[\sum_r v(t(r)) p\,(r : 1) - a - \sum_r v\right.$$

- The FOC gives, for all r,

$$-p\,(r:1)+\lambda v'(t(r))p\,(r:1)+\mu[v'(t(r))p\,(r:1)-v'(t(r))p\,(r:0)]=0,$$

or

$$\frac{1}{v'(t(r))}=\lambda+\mu\left[1-\frac{p\,(r:0)}{p\,(r:1)}\right].$$

At the optimum, the individual rationality and incentive compatibility constraints bind (why?) and hence $\lambda,\mu>0$.

- Note that even if the agent is risk averse, the optimal contract is *not* constant: the optimal $t(r)$ varies in the ratio $p(r:0)/p(r:1)$.

- The optimal contract rewards such r that indicate $a=1$ being committed.

- But nothing guarantees that the ratio $p(r:0)/p(r:1)$ is increasing in r. This implies that the optimal scheme $t(\cdot)$ may not be increasing.

- However, in many cases it is natural to think that $p(r:0)/p(r:1)$ does increase in r. Such condition is called the *monotone likelyhood ratio*.

Principal-agent model with mean variance utility (Holmström-Milgrom 1987)

- Let the agent choose an action $a \in \mathbb{R}_{++}$. Assume that his payoffs are characterized by CARA utility function

$$v(t) = 1 - \exp\{-\rho(t - c(a))\},$$

where ρ is the degree of absolute risk aversion, t is a monetary reward and $c(a)$ is the monatary value of the cost that the agent sacrifices to action a. Assume c is is convex.

- Realization of r is random and defined by $r = a + \varepsilon$ where $\varepsilon \sim N(0, 1)$.

- The principal is risk neitral and maximizes his expected payoff $r - t(r)$. Let us focus on a class of linear incentive schemes

$$t(r) = \alpha + \beta r,$$

where α and β are parametrized to determined: β reflects the degree of incentives in the contract and α the level of basic enumeration to the agent.

- The expected value of $a + \varepsilon - \alpha - \beta(a + \varepsilon)$, the principal's expected payoff, is

$$a(1 - \beta) - \alpha.$$

- The agent must be provided enough incentives to participate. The value of outside option is $0(= 1 - e^0)$. What is the principal's payoff maximizing linear incentive scheme?

- Agent's expected payoff is

$$\frac{1}{\sqrt{2\pi}} \int -\exp\{-\rho(t(a + \varepsilon) - c(a)\} \exp\{-\frac{\varepsilon^2}{2}\} d\varepsilon.$$

Complete the square

$$\begin{aligned} & \rho t(a+\varepsilon)+\frac{\varepsilon^2}{2} \\ = \ & \rho(\alpha+\beta(a+\varepsilon))+\frac{\varepsilon^2}{2} \\ = \ & \rho\alpha+\rho\beta a+\rho\beta\varepsilon+\frac{\varepsilon^2}{2} \\ = \ & \rho(\alpha+\beta a)-\frac{(\rho\beta)^2}{2}+\frac{(\varepsilon-\rho\beta)^2}{2}. \end{aligned}$$

- Thus the agents expected payoff is then

$$-\exp\{-\rho[\alpha+\beta a+\frac{\rho\beta^2}{2}-c(a)]\}.$$

Given α and β, the agent's optimization problem reduces to solving

$$\max_a \{\alpha + \beta a - \frac{\rho\beta^2}{2} - c(a)\}.$$

The FOC gives

$$\beta = c'(a).$$

- The principal's optimization problem reduces to

$$\begin{aligned} & \max_a [a(1-\beta) - \alpha] \\ \text{s.t. } \beta &= c'(a), \\ 0 &\leq \alpha + \beta a - \frac{\rho\beta^2}{2} - c(a). \end{aligned}$$

In the optimum the participation constraint binds. Solving for β and α, the principal's problem simplifies to

$$\max_a [a - \frac{\rho}{2} c'(a)^2 - c(a)].$$

FOC

$$1 - \rho c''(a) c'(a) - c'(a) = 0.$$

Using $\beta = c'(a)$, we have

$$\beta = \frac{1}{1 + \rho c''(a)}.$$

- Thus increase in the degree of of risk aversion ρ decreases the incentives β provided to the agent (increase in the varaince of ε would have a similar impact). Note that β is always less than one (what does it mean?).