Lecture 9.1

# Bayesian Game Induced by a Mechanism

$$M = \langle N=[n], \mathcal{X}, (\Theta_i)_{i\in N}, (S_i)_{i\in N}, P\in\Delta(\Theta), \underline{\underline{g : \times_{i=1}^{n} S_i \to \mathcal{X}}}, u_i : \underline{\underline{\mathcal{X}\times\Theta}} \to \mathbb{R} \rangle$$

induces a Bayesian game

$$\Gamma_M = \langle N=[n], (\Theta_i)_{i\in N}, (S_i)_{i\in N}, P, U_i : \times_{i=1}^{n} S_i \times \Theta \to \mathbb{R} \rangle$$

$$U_i(\underline{(s_i)_{i\in N}}, \underline{\underline{(\theta_i)_{i\in N}}}) = u_i(g((s_i)_{i\in N}), (\theta_i)_{i\in N})$$

# Implementation of a Social Choice Function

$$f : \Theta \longrightarrow \mathcal{X}$$

**Definition:** We say that an indirect mechanism $M = \left((S_i)_{i \in N}, g(\cdot)\right)$ implements a social choice function $f : \Theta \to \mathcal{X}$ if there exists an "equilibrium" $(s_i^*(\cdot))_{i \in N}$ in the induced Bayesian game $\Gamma_M$ such that

$$\forall (\theta_1, \ldots, \theta_n) \in \Theta, \quad g\left(s_1^*(\theta_1), \ldots, s_n^*(\theta_n)\right) = f(\theta_1, \ldots, \theta_n)$$

If the equilibrium in the above definition is a very weakly dominant strategy equilibrium, then we say that the mechanism M implements the social choice function f in dominant strategy equilibrium.

In this case, the mechanism M is called dominant strategy incentive compatible (DSIC) with f.

If the equilibrium is a pure strategy Bayesian Nash equilibrium, then we say that the mechanism M implements the social choice function f in Bayesian Nash equilibrium. In this case, we call M to be Bayesian Incentive Compatible (BIC) with f.

Examples:

Buying auction, i.e. one buyer and $n$ potential sellers.

(i) $f_{fp} : \Theta_0 \times \dots \times \Theta_n \longrightarrow \mathcal{X}$ $\qquad f_{fp}(\theta_0, \dots, \theta_n) = (a_0, \dots, a_n, p_0, \dots, p_n)$

defined as follows:

if $\theta_0 < \min_{i \in [n]} \theta_i$ : $a_i = 0$, $p_i = 0$ $\forall\ i \in \{0, 1, \dots, n\}$

otherwise: $\quad a_0 = 1$, $a_j = -1$ where $j \in \operatorname{argmin}_{i \in [n]} \{\theta_i\}$

$a_i = 0 \quad \forall\ i \in [n] \setminus \{j\}$.

$p_0 = \theta_j$, $p_j = -\theta_j$, $p_i = 0 \quad \forall\ i \in [n] \setminus \{j\}$.

(2) $f_{sp} : \Theta_0 \times \cdots \times \Theta_n \longrightarrow \mathcal{X}$ $\qquad f_{sp}(\theta_0, \dots, \theta_n) = (a_0, a_1, \dots, a_n, p_0, p_1, \dots, p_n)$.

defined as:

if $\theta_0 < \min_{i \in [n]} \theta_i$ : $\quad a_0 = 0,\ a_i = 0,\ p_0 = 0,\ p_i = 0 \quad \forall i \in \{1, \dots, n\}$

otherwise

$$a_0 = 1, \quad a_j = -1 \quad \text{where } j \in \arg\min_{i \in [n]} \{\theta_i\}$$

$$a_i = 0 \quad \forall i \in [n] \setminus \{j\}$$

$$p_0 = \min_{i \in [n] \setminus \{j\}} \theta_i, \quad p_j = -p_0, \quad p_i = 0 \quad \forall i \in [n] \setminus \{j\}$$

We have seen that the direct mechanism implements $f_{sp}$ in dominant strategies. However the direct mechanism $f_{fp}$ does not implement the social choice function even in Bayesian Nash equilibrium!

Lecture 9.2

The direct mechanism for the first price auction is not even Bayesian Incentive compatible.

## Revelation Principal for Dominant Strategy Incentive Compatibility:

Let $f : \times_{i \in [n]} \Theta_i \rightarrow X$ be any social choice function. Suppose an indirect mechanism $((S_i)_{i \in [n]}, g(\cdot))$ implements $f$ in Dominant strategy equilibrium. Then the direct mechanism $\left((\Theta_i)_{i\in[n]}, f(\cdot)\right)$ implements $f$ in Dominant Strategy equilibrium.

<u>Proof:</u> Since $M = \left((S_i)_{i\in[n]}, g(\cdot)\right)$ implements $f$ in DSE, there exists a very weakly dominant strategy equilibrium $\left(s_i^*(\cdot)\right)_{i\in[n]}$ such that

$\forall\ (\theta_1, \ldots, \theta_n) \in \bigtimes_{i=1}^{n} \Theta_i$

$$g\left(s_i^*(\theta_i)\right)_{i\in[n]} = f\left((\theta_i)_{i\in[n]}\right) \qquad (1)$$

$\forall i \in [n], \forall \theta_i \in \Theta_i, \forall a \in S_i, \forall \underline{a}_{-i} \in \underline{S}_{-i}, \forall \underline{\theta}_{-i} \in \Theta_{-i}$

$$u_i\left(g\left(\left(s_i^*(\theta_i), \underline{a}_{-i}\right)\right), (\theta_i)_{i\in[n]}\right) \geqslant u_i\left(g\left(\underline{(a, \underline{a}_{-i})}\right), (\theta_i)_{i\in[n]}\right) \quad \text{—— (2)}$$

<u>To show:</u> $\left(f, (\Theta_i)_{i\in[n]}\right)$ is DSIC.

$\forall i \in [n], \forall \theta_i \in \Theta_i, \forall \underline{\theta}_{-i} \in \Theta_{-i}$ $\quad \forall \theta_i' \in \Theta_i \; \forall \underline{\theta}_{-i}' \in \Theta_{-i}$

$$u_i\left(f(\theta_i, \underline{\theta}_{-i}'), (\theta_j)_{j\in[n]}\right) \geqslant \underline{u_i\left(f(\theta_i', \underline{\theta}_{-i}'), (\theta_j)_{j\in[n]}\right)}$$

$$u_i\left( f(\theta_i, \underline{\theta}'_{-i}), (\theta_j)_{j\in[n]} \right)$$

$$= u_i\left( g\left( s_i^*(\theta_i), \underline{s}_{-i}^*(\underline{\theta}'_{-i})\right), (\theta_j)_{j\in[n]} \right) \quad \text{[from equation (1)]}$$

$$\geq u_i\left( g\left( s_i^*(\theta'_i), \underline{s}_{-i}^*(\underline{\theta}'_{-i})\right), (\theta_j)_{j\in[n]} \right) \quad \left[\text{from equation (2), put } a = s_i^*(\theta'_i),\ \underline{a}_{-i} = \underline{s}_{-i}^*(\underline{\theta}'_{-i})\right]$$

$$= u_i\left( f(\theta'_i, \underline{\theta}'_{-i}), (\theta_j)_{j\in[n]} \right) \quad \text{[from equation (1)]}$$

$\boxtimes$

# Revelation Principle for BIC mechanisms

Let $f : \underset{i \in [n]}{\times} \Theta_i \longrightarrow \mathcal{X}$ be any social choice function. Suppose an indirect mechanism implements $f$ in Bayesian Nash equilibrium. Then the direct mechanism $\left( (\Theta_i)_{i \in [n]}, f(\cdot) \right)$ also implements $f(\cdot)$ in Bayesian Nash equilibrium.

Lecture 9.3

Main take-away from revelation principle:

"the search for indirect mechanism to implement a social choice function is trivial"

Next big question of mechanism design: characterize the set of all social choice functions which are implementable in DSE / BNE.

# Useful Properties of Social Choice Functions

## (1) Ex-post efficiency / Pareto optimality / Efficiency :-

A social choice function $f : \times_{i \in [n]} \Theta_i \longrightarrow \mathcal{X}$ is called ex-post efficient if, for every type profile, the out-come chosen is pareto-optimal. That is

$\forall (\theta_1, \ldots, \theta_n) \in \times_{i \in [n]} \Theta_i$, there does not exist any outcome $x \in \mathcal{X}$ such that

$$u_i(x, (\theta_1, \ldots, \theta_n)) \geq u_i(f(\theta_1, \ldots, \theta_n), (\theta_1, \ldots, \theta_n)) \quad \forall i \in [n]$$

and there exists a player $j \in [n]$ such that

$$u_j(x, (\theta_1, \ldots, \theta_n)) > u_j(f(\theta_1, \ldots, \theta_n), (\theta_1, \ldots, \theta_n)).$$

(2) **Non-Dictatorship:** A player $d \in [n]$ is called a dictator if $\forall (\theta_1, \ldots, \theta_n) \in \bigtimes_{i=1}^{n} \Theta_i$

$$u_d(f(\theta_1, \ldots, \theta_n), (\theta_1, \ldots, \theta_n)) \geq u_d(x, (\theta_1, \ldots, \theta_n)) \quad \forall x \in \mathcal{X}$$

A social choice function is called a dictatorship if there exists a dictator. Otherwise the social choice function is called non-dictatorship.

③ Individual Rationality: Let $\bar{u}_i(\theta_i)$ be the utility of player $i$ when it does not parcipate in the mechanism and its type is $\theta_i \in \Theta_i$.

(a) Ex-post Individual Rationality: A social choice function $f$ is called ex-post IR if

$\forall i \in [n]$,

$$u_i\left(f(\theta_i, \underline{\theta}_{-i}), (\theta_1, \ldots, \theta_n)\right) \geq \bar{u}_i(\theta_i) \quad \forall (\theta_i, \underline{\theta}_{-i}) \in \Theta$$

(b) Interim Individual Rationality: A social choice function $f$ is called interim IR if

$\forall i \in [n]$, $\forall \theta_i \in \Theta_i$

$$\mathbb{E}\left[u_i\left(f(\theta_i, \underline{\theta}_{-i}), (\theta_1, \ldots, \theta_n)\right) \middle| \theta_i\right] \geq \bar{u}_i(\theta_i) \quad \forall \theta_i \in \Theta_i$$

---

Prior $\underline{\underline{P}} \in \Delta\left(\bigtimes_{i=1}^{n} \Theta_i\right)$

(c) Ex-Ante Individual Rationality: A social choice function $f$ is ex-ante IR if

$$\forall i \in [n], \quad \mathbb{E}\left[u_i\left(f(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})\right)\right] \geq \mathbb{E}\left[\bar{u}_i(\theta_i)\right]$$

# Gibbard (1973) - Satterwaite (1977) Theorem

**Lecture 9.4**

For any player $i \in [n]$ and any type $\theta_i \in \Theta_i$, we get a partial order $R_i$ on the set $\mathcal{X}$ of outcomes.

$x, y \in \mathcal{X}$, $\quad x \, R_i^{\theta_i} \, y \overset{\text{def}^n}{\equiv} u_i(x, \theta_i) \geq u_i(y, \theta_i)$

$R_i^{\theta_i}$ is called the rational preference relation of player $i$ when its type is $\theta_i$.

We call a preference relation to be strict if that rational preference relation is a linear order/complete order. We denote the set of all possible strict rational preference relations on $\mathcal{X}$ by $\mathcal{L}(\mathcal{X})$.

Recall: $f : \underset{i \in [n]}{\times} \Theta_i \longrightarrow \mathcal{X}$

Social choice function (G.S Theorem) $f : \mathcal{L}(\mathcal{X})^n \longrightarrow \mathcal{X}$

**Unanimity:** A social choice function $f : \mathcal{L}(X)^n \longrightarrow X$ is called unanimous if

$\forall\, P_1, \ldots, P_n \in \mathcal{L}(X)$ such that the best outcomes in all $P_1, \ldots, P_n$ are the same, say $x$.

we have $f(P_1, \ldots, P_n) = x$

Ex-post efficiency $\Rightarrow$ unanimity.

Theorem: Let $f : \mathcal{L}(\mathcal{X})^n \longrightarrow \mathcal{X}$ be a social choice function such that

(i) We have at least 3 outcomes. That is $|\mathcal{X}| \geq 3$.

(ii) $f$ is unanimous.

(iii) Every player has a strict rational preference relation.

Then $f$ is dominant strategy incentive compatible if and only if $f$ is a dictatorship.

## Way-Outs from GS Impossibility

(1) Assume "more structure" on outcomes and "more structure" on the utility functions of the players.

— Quasilinear setting.

(2) Be satisfies with Bayesian incentive compatibility.

(3) "Computational barrier" can the manipulation problem be NP-hard?

# Quasi-Linear Setting / Environment

The outcomes are not arbitrary. In particular, an outcome is a tuple $(k, t_1, \ldots, t_n)$. The first component $k$ is an allocation which belongs to a set $\mathcal{K}$ of set of allocations. $t_i$ is the money received by player $i$.

The set of outcomes:

$$X = \left\{ (k, t_1, \ldots, t_n) \;\middle|\; k \in \mathcal{K},\ \underbrace{\sum_{i=1}^{n} t_i \leq 0}_{\text{since no external supply of money.}} \right\}$$

The utility function has the following structure.

$$\begin{aligned} u_i(x, \theta_1, \ldots, \theta_n) &= u_i(k, t_1, \ldots, t_n, \theta_1, \ldots, \theta_n) \\ &= u_i(k, t_1, \ldots, t_n, \theta_i) \\ &= v_i(k, \theta_i) + t_i \end{aligned}$$

$v_i(k, \theta_i)$ is the valuation of allocation $k$ to player $i$ when its type is $\theta_i$.

$$\text{utility} = \text{valuation} + \text{payment}.$$

A social choice function $f(\theta_1, \ldots, \theta_n)$ has the following structure:-

$$f(\theta_1, \ldots, \theta_n) = \left( k(\theta_1, \ldots, \theta_n), \ \left( t_i(\theta_1, \ldots, \theta_n) \right)_{i \in [n]} \right)$$

The social choice function consists of two functions.

(i) Allocation function: $k : \bigtimes_{i=1}^{n} \Theta_i \longrightarrow \mathcal{K}$

(ii) payment functions: $t_j : \bigtimes_{i=1}^{n} \Theta_i \longrightarrow \mathbb{R}$

Q: How does any efficient social choice function look like in a quasi-linear environment?

Allocative Efficiency (AE): An allocation function is called allocatively efficient if

$\forall\ (\theta_1, \ldots, \theta_n) \in \bigtimes_{i=1}^{n} \Theta_i$, we have

$$k(\theta_1, \ldots, \theta_n) \in \underset{k \in \mathcal{K}}{\operatorname{argmax}} \sum v_i(k, \theta_i)$$

equivalently,

$$\sum_{i=1}^{n} v_i\left(k(\theta_1, \ldots, \theta_n), \theta_i\right) = \max_{k \in \mathcal{K}} \sum_{i=1}^{n} v_i(k, \theta_i)$$

<u>Budget Balanced (BB)</u>: The payment functions $(t_i(\cdot))_{i \in [n]}$ are called strongly budget balanced (SBB) if

$$\forall (\theta_1, \ldots, \theta_n) \in \bigtimes_{i=1}^{n} \Theta_i ,$$

$$\sum_{i=1}^{n} t_i(\theta_1, \ldots, \theta_n) = 0$$

The payment functions $(t_i(\cdot))_{i \in [n]}$ are called weakly budget balanced if

$$\forall (\theta_1, \ldots, \theta_n) \in \bigtimes_{i=1}^{n} \Theta_i$$

$$\sum_{i=1}^{n} t_i(\theta_1, \ldots, \theta_n) \leq 0$$

Observation: If we have at least two players, then no social choice function is dictatorial in a quasi-linear environment.

Proof: Let $n (\geq 2)$ be the number of players. If possible, let us assume that there exists a social choice function $f(\cdot)$ where a player, say player $d$, is a dictator.

Let $\varepsilon > 0$ be any positive real number, $\theta$ any type profile. Suppose $f(\theta) = (k(\theta), (t_d(\theta), t_{-d}(\theta)))$

Consider the outcome,

$$x = \left( k(\theta), \left(t_j(\theta) - \epsilon\right)_{j \in [n], j \neq d}, t_d(\theta) + (n-1)\epsilon \right)$$

$$u_d(x, \theta_d) = \underbrace{v_d(k(\theta)) + t_d} + (n-1)\epsilon$$

$$= u_d(f(\theta), \theta_d) + (n-1)\epsilon$$

$$> u_d(f(\theta), \theta_d)$$

This contradicts our assumption that $d$ is the dictator of the social choice function $f$. $\blacksquare$