Informationally Robust Implementation

Stephen Morris

28th Jerusalem School in Economic Theory 2017

This Lecture

Some informationally robust analogues of first two lectures:

1. Full Implementation in General Environments
2. Allocation of a Single Good with Interdependent Values

Complete Information Environment

- society $\{1, .., n\}$
- set of outcomes A
- states of the world Θ
- preferences $u_i(a, \theta)$
- state θ is common knowledge among players

Problem 1: Complete Information Implementation

- designer does not know θ
- chooses mechanism:
  - player i strategies M_i
  - game form $g : M_1 \times .. \times M_n \rightarrow A$
- wants to implement **function** $f : \Theta \rightarrow A$ in Nash equilibrium, i.e., find a mechanism such that, for each θ, all Nash equilibria $(m_1, ..., m_n)$ satisfy

$$g\left(m_1, ..., m_n\right) = f\left(\theta\right)$$

Characterization

DEFINITION. Social choice function f is **Maskin monotonic** if $f\left(\theta'\right) = f\left(\theta\right)$ whenever, for each i and y

$$u_i\left(f\left(\theta\right), \theta\right) \geq u_i\left(y, \theta\right) \Rightarrow u_i\left(f\left(\theta\right), \theta'\right) \geq u_i\left(y, \theta'\right)$$

THEOREM. Maskin monotonicity is necessary and "almost" sufficient for f to be implementable in Nash equilibrium.

Checklist for Complete Information

1. Interpretable necessary condition
2. Proof of (almost) sufficiency using exotic mechanisms
3. Applications addressing multiple equilibria using simpler mechanisms

Incomplete Information

- society $\{1, .., n\}$
- set of outcomes A
- player i types Θ_i
- preferences $u_i\left(a, \left(\theta_i, \theta_{-i}\right)\right)$

Incomplete Information

- designer still does not know $\theta = (\theta_1, ..., \theta_n)$
- the mechanism is still
  - player i messages M_i
  - game form $g : M_1 \times .. \times M_n \rightarrow A$

Problem 2: Bayesian Implementation

- there is a common prior $\pi^* \in \Delta(\Theta)$
- strategy for player i is now a mapping $s_i : \Theta_i \to M_i$
- a strategy profile $(s_1, .., s_n)$ is a *Bayes Nash equilibrium* if

$$\begin{aligned} & \sum_{\theta} \pi^*(\theta)\, u_i\left(g\left(s_i(\theta_i), s_{-i}(\theta_{-i})\right), (\theta_i, \theta_{-i})\right) \\ \geq \; & \sum_{\theta} \pi^*(\theta)\, u_i\left(g\left(s_i'(\theta_i), s_{-i}(\theta_{-i})\right), (\theta_i, \theta_{-i})\right) \end{aligned}$$

  for all i and $s_i' : \Theta_i \to M_i$
- designer wants to implement social choice function $f : \Theta_1 \times .. \times \Theta_n \to A$ in Bayes Nash equilibrium: all Bayes Nash equilibria $(s_1, ..., s_n)$ satisfy

$$g\left(s_1(\theta_1), ..., s_n(\theta_n)\right) = f(\theta_1, ..., \theta_n)$$

Bayesian Incentive Compatability

Definition

Social choice function f satisfies Bayesian incentive compatability if

$$\begin{aligned} & \sum_{\theta_{-i}} \pi^* \left(\theta_{-i} | \theta_i\right) u_i \left(f\left(\theta_i, \theta_{-i}\right), \left(\theta_i, \theta_{-i}\right)\right) \\ \geq & \sum_{\theta_{-i}} \pi^* \left(\theta_{-i} | \theta_i\right) u_i \left(f\left(\theta_i', \theta_{-i}\right), \left(\theta_i, \theta_{-i}\right)\right) \end{aligned}$$

for all i, θ_i and θ_i'.

By the revelation principle, Bayesian incentive compatibility is already necessary for partial implementation

Definition

Social choice function satisfies Bayesian monotonicity if......

Characterization

Theorem

Bayesian incentive compatibility and Bayesian monotonicity are necessary and "almost" sufficient for f to be implementable in Bayes Nash equilibrium.

Checklist for Bayesian Implementation

1. Interpretable necessary condition **NOT REALLY**
2. Proof of (almost) sufficiency using exotic mechanisms **YES**
3. Applications addressing multiple equilibria using simpler mechanisms **only exchange economies?**

Robust Implementation

- as before but replace common prior $\pi^* \in \Delta(\Theta_1 \times ... \times \Theta_n)$
- instead assume type space:
  - player i types T_i
  - mappings
    1. preferences $\widehat{\theta}_i : T_i \to \Theta_i$
    2. beliefs $\widehat{\pi}_i : T_i \to \Delta(T_{-i})$
- require Bayesian implementation on all type spaces ($\approx$ universal type space)
- yesterday's lecture: equivalent to (belief-free) rationalizable implementation on all payoff type spaces

Rationalizable Implementation

- Iterated Deletion Solution Concept:
  - initialize $R_i^0(\theta_i) = M_i$
  - for $k \geq 0$, inductive define $R_i^{k+1}(\theta_i)$ to be the set of messages such that there exists $\nu_i \in \Delta(M_{-i} \times \Theta_{-i})$ such that

$$
\begin{aligned}
&(1)\ \nu_i(m_{-i}, \theta_{-i}) > 0 \Rightarrow m_j \in R_j^k(\theta_j) \text{ for all } j \\
&(2)\ m_i \in \arg\max_{m_i'} \sum_{m_{-i}, \theta_{-i}} \nu_i(m_{-i}, \theta_i)\, u_i\left(g\left(m_i', m_{-i}\right), (\theta_i, \theta_{-i})\right)
\end{aligned}
$$

  - $R_i^\infty(\theta_i) = \bigcap_{k \geq 1} R_i^k(\theta_i)$
- PROPOSITION: social choice function f is robustly implemented if $m \in R^\infty(\theta) \Rightarrow g(m) = f(\theta)$

Characterization

Theorem

Robust monotonicity is necessary and "almost" sufficient for f to be robustly implementable.

Maskin Monotonicity and Whistle-Blowing

Social choice function f is **Maskin monotonic** if $f\left(\theta'\right) = f\left(\theta\right)$ whenever, for each i and y,

$$u_i\left(f\left(\theta\right), \theta\right) \geq u_i\left(y, \theta\right) \Rightarrow u_i\left(f\left(\theta\right), \theta'\right) \geq u_i\left(y, \theta'\right)$$

- Equivalent definition: Social choice function f is Maskin monotonic if whenever $f\left(\theta'\right) \neq f\left(\theta\right)$, there exist i and y such that $u_i\left(y, \theta'\right) > u_i\left(f\left(\theta\right), \theta'\right)$ and $u_i\left(f\left(\theta\right), \theta\right) \geq u_i\left(y, \theta\right)$
- Breaking it down:
  - a *deception* is a mis-report $\theta \rightarrow \theta'$
  - a deception is *acceptable* if $f\left(\theta'\right) = f\left(\theta\right)$
  - the *reward set* for player i at θ is the set of outcomes y that do not break the good equilibrium, i.e., $\left\{y \middle| u_i\left(f\left(\theta\right), \theta\right) \geq u_i\left(y, \theta\right)\right\}$
  - a deception $\theta \rightarrow \theta'$ is *refutable* if there exists a whistle-blower i and an outcome y in the state θ reward set, such that $u_i\left(y, \theta'\right) > u_i\left(f\left(\theta\right), \theta'\right)$
  - a social choice function satisfies Maskin monotonicity if every unacceptable deception is refutable

Robust Monotonicity Definition

- A *deception* is a profile $\beta = (\beta_i)_{i=1}^{n}$ with each $\beta_i : \Theta_i \longrightarrow 2^{\Theta_i} / \varnothing_i$
- A deception is *acceptable* if $\theta' \in \beta(\theta) \Rightarrow f(\theta') = f(\theta)$
- The *reward set* for player i at type profile θ'_{-i} is

$$Y_i\left(\theta'_{-i}\right) = \left\{ y \;\middle|\; \begin{array}{l} \text{for all } \theta''_i \text{, either } f\left(\theta''_i, \theta'_{-i}\right) = y \\ \text{or } u_i\left(f\left(\theta''_i, \theta'_{-i}\right), \left(\theta''_i, \theta'_{-i}\right)\right) > u_i\left(y, \left(\theta''_i, \theta'_{-i}\right)\right) \end{array} \right\}$$

Robust Monotonicity Definition

- A deception is *strictly refutable* if there exists a whistle blower i and $\theta_i' \in \beta_i(\theta_i)$ such that for all θ_{-i}' and $\psi_i \in \Delta\left(\beta_{-i}^{-1}(\theta_{-i}')\right)$ there exists $y \in Y_i\left(\theta_{-i}'\right)$ such that

$$\begin{aligned} & \sum_{\theta_{-i} \in \beta_{-i}^{-1}\left(\theta_{-i}'\right)} \psi_i(\theta_{-i}) u_i(y, (\theta_i, \theta_{-i})) \\ > & \sum_{\theta_{-i} \in \beta_{-i}^{-1}\left(\theta_{-i}'\right)} \psi_i(\theta_{-i}) u_i\left(f\left(\theta_i', \theta_{-i}'\right), (\theta_i, \theta_{-i})\right) \end{aligned}$$

- Social choice function satisfies *strict robust monotonicity* if every unacceptable deviation is strictly refutable.

Sketch of Proof

- Suppose that f is robustly implementable and fix a mechanism that robustly implements it.
- Suppose that deception β is not strictly refutable.
- Let
$$S_i^{\beta}(\theta_i) = \underset{\theta_i' \in \beta_i(\theta_i)}{\cup} R_i^{\infty}(\theta_i')$$
- Because β is not strictly refutable $S_i^{\beta}(\theta_i) \subseteq R_i^{\infty}(\theta_i)$ for all i and θ_i.
- This implies that β is acceptable

Ex Post Incentive Compatability

- social choice function f satisfies strict ex post incentive compatibility (EPIC) if

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

  for all i, θ_{-i}, θ_i and θ_i'; with strictly inequality unless $f \left(\theta_i, \theta_{-i}'\right) = f \left(\theta_i', \theta_{-i}'\right)$ for all θ_{-i}'.
- Robust monotonicity implies strict EPIC
- In private values environments, strict EPIC reduces to strict dominant strategies incentive compability
- To the extent robust monotonicity goes beyond strict EPIC, there must be restrictions on interdependence

Checklist for Robust Implementation

1. Interpretable necessary condition
   - **better than Bayesian monotonicity?**!
   - **= "NOT TOO MUCH INTERDEPENDENCE"**
2. Proof of (almost) sufficiency using exotic mechanisms **YES**
3. Applications addressing multiple equilibria using simpler mechanisms **YES, ALLOCATING A SINGLE GOOD WITH INTERDEPENDENT VALUES**

Second Price Auction (Private Values)

- Allocating a good with private values $\theta_i \in [0, 1]$
- Only weakly dominated strategies: a problem for full implementation
- But can strengthen to strict second price auction to get full implementation:
  - second price auction with probability $1 - \varepsilon$,
    - allocate object to highest bidder
    - winner pays second highest bid
  - strict screening mechanism with probability ε: for each i, with probability $\frac{1}{n} b_i$
    - i gets object
    - pays $\frac{1}{2} b_i$
- truth-telling is a strictly dominant strategy
- thus this (ε-efficient) allocation is robustly implemented

Adding Interdependence of Values

- Agent i has payoff types $\theta_i \in [0,1]$
- Agent i's value is

$$v_i(\theta) = \theta_i + \gamma \sum_{j \neq i} \theta_j$$

- Assume that $\gamma < 1$: single crossing condition from lecture 2

Extended Second Price Auction

- strict extended second price auction:
  - extended second price auction with probability $1 - \varepsilon$,
    - allocate object to highest bidder i
    - winner pays $\max_{j \neq i} b_j + \gamma \sum_{j \neq i} b_j$
  - strict screeing mechanism with probability ε: for each i with probability $\frac{1}{n} b_i$,
    - i gets object
    - pays $\frac{1}{2} b_i + \gamma \sum_{j \neq i} b_j$
- in extended second price auction, winner's payment is independent of his bid and is his willingness to pay at the lowest bid at which he wins

Good News and Bad News

- truth-telling is a strict ex post equilibrium of this "direct" mechanism
- but existence of strict ex post equilibrium does *not* imply robust implementation

Full Implementation

- The direct mechanism robustly implements the (almost) efficient outcome if $|\gamma| < \frac{1}{n-1}$...
- ...because the direct mechanism has linear best response functions

$$a_i = \theta_i - \gamma \sum_{j \neq i} (a_j - \theta_j) .$$

  and we proved last lecture that there is a unique "truth-telling" rationalizable outcome in this game

Failure of Full Implementation

- No mechanism implements the (almost) efficient outcome if $|\gamma| \geq \frac{1}{n-1}$...
- We showed last lecture that every action was rationalizable in this game
- But no other mechanism would work either; two ways to see this:
  1. We can directly verify failure of robust monotonicity
  2. Also ad hoc argument (assuming $\gamma \geq 0$)
     - Consider a type space where whenever a player has payoff type θ_i, he has dogmatic belief that each other player has payoff type
     $$\frac{1}{2} + \frac{1}{\gamma(n-1)}\left(\frac{1}{2} - \theta_i\right).$$
     - We know this is in the interval $[0, 1]$ only because $\gamma \geq \frac{1}{n-1}$
     - His valuation is
     $$\theta_i + \gamma\,(n-1)\left(\frac{1}{2} + \frac{1}{\gamma(n-1)}\left(\frac{1}{2} - \theta_i\right)\right) = \frac{1}{2}\,(1 + \gamma\,(n-1))$$
     independent of his type
     - always a pooling equilibrium

Argument generalizes....

- Consider an environment where players' preferences depend on a statistic of players' payoff types and are single crossing with respect to that statistic...
- covers allocating a good with interdependent properties in lecture 2.
- then robust implementation is possible if and only if it is possible in the payoff type direct mechanism
- robust implementation is possible if and only if there is not too much interdependence

The Common Prior Assumption and Positive and Negative Interdependence

- We allowed for negative interdependence $\gamma \leq 0$. In this case, there are strategic complementarities and imposing the common prior does not make robust implementation easier
- But if $\gamma \geq 0$....
  - there are strategic substitutes
  - arguments from last lecture establish robust implementation easier under the common prior
  - but arguments special