Robust Mechanism Design

Stephen Morris

Atlanta Winter Meetings of the Econometric Society
January 2010

Mechanism Design

GOOD NEWS:
Mechanism Design has succeeding in developing

- a rich consistent theoretical framework for thinking about mechanism design...
- insights about practical mechanism design problems....

Mechanism Design

GOOD NEWS:
Mechanism Design has succeeding in developing

- a rich consistent theoretical framework for thinking about mechanism design...
- insights about practical mechanism design problems....

BAD NEWS:

- Connections between the two are sometimes nebulous...
- "Simple", user-friendly, mechanisms used in preference to "optimal mechanisms"

This Talk

Review some recent work by Dirk Bergemann and myself on "Robust Mechanism Design:" finding mechanisms that work for relaxed informalational assumptions.

Wilson Doctrine

"Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent that it assumes other features to be common knowledge, such as one agent's probability assessment about another's preferences or information. I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems. Only by repeated weakening of common knowledge assumptions will the theory approximate reality." Wilson (1987)

This Talk

Review some recent work by Dirk Bergemann and myself on "Robust Mechanism Design:"

- finding mechanisms that work for relaxed assumptions about beliefs and higher order beliefs (robust partial implementation)
  - "Robust Mechanism Design," *Econometrica* 2005
- full implementing objectives: all equilibria deliver the right outcomes (robust full implementation)
  - "Robust Implementation in Direct Mechanisms," *Review of Economic Studies* 2009

Outline

1. Solution concepts:
   1.1 robust partial implementation = ex post partial implementation ONLY IN SEPARABLE ENVIRONMENTS
   1.2 robust full implementation = implementation in (belief-free) rationalizable strategies
2. Substance: full robust implementation if and only if not too much interdependence in agents' types
3. Further results in this agenda and directions for future research

General Problem Studied in this Talk

- agents $i \in \mathcal{I} = \{1, 2, ..., I\}$
- i's "payoff type" $\theta_i \in \Theta_i$
- payoff type profile $\theta \in \Theta = \Theta_1 \times ... \times \Theta_I$
- social outcome $y \in Y$
- utility function $u_i : Y \times \Theta \to \mathbb{R}$ (allows "interdependent preferences")
- Correspondence $F : \Theta \to 2^Y \setminus \varnothing$
  - $F(\theta) \subseteq Y$ are the set of outcomes acceptable to the planner

Leading Example: Public Good and Quasi-Linear Preferences

- payoff types $\Theta_i = [0,1]$
- social outcomes $Y = [0,1] \times \mathbb{R}^I$
- typical element $y = (q, x) = (q, (x_1, ..., x_I))$
- utility function
  $u_i(y, \theta) = u_i((q, (x_1, .., x_I)), \theta) = v_i(\theta).q + x_i$
  - $v_i(\theta) = \theta_i + \gamma \sum_{j \neq i} \theta_j$
  - $\gamma \geq 0$ (non-negative interdependence)
  - $\gamma = 0$: private values

Leading Example: Public Good and Quasi-Linear Preferences

- planner maximizes $\left(\sum_i v_i(\theta)\right) q - \frac{1}{2}q^2$
  - $q^*(\theta) = (1 + \gamma(I-1)) \sum_{i=1}^{I} \theta_i$
- two classic social choice correspondences

$$
\begin{aligned}
F^*(\theta) &= \{(q,x) \in Y \,|\, q = q^*(\theta)\} \\
F^{**}(\theta) &= \left\{(q,x) \in F^*(\theta) \,\middle|\, \sum_{i=1}^{I} x_i = \overline{x}\right\}
\end{aligned}
$$

Dealing with Incomplete Information

- Consider private values ($\gamma = 0$)
- Dominant strategies
  - Classic results for quasi-linear environments, efficiency (F^*) is achievable (VCG), efficiency and budget balance (F^{**}) is not.
- Classic "Bayesian" Approach
  - Assume a common prior $p \in \Delta(\Theta)$ and study Bayes-Nash equilibrium
  - Under weak ass'ns, F^{**} is achievable

An early debate about the Bayesian approach

- Critique of Bayesian approach from "practical people"
  - in practise, the "prior" is not known to the planner
  - if social choice correspondence must be implemented without knowing the prior, dominant strategies is required
  - Dasgupta, Hammond + Maskin (1979), Ledyard (1978, 1979), Groves + Ledyard (1987)

An early debate about the Bayesian approach

- Critique of Bayesian approach from "practical people"
  - in practise, the "prior" is not known to the planner
  - if social choice correspondence must be implemented without knowing the prior, dominant strategies is required
  - Dasgupta, Hammond + Maskin (1979), Ledyard (1978, 1979), Groves + Ledyard (1987)
- Critique of the critique from "Bayesian purists":
  - if the agents know the prior, you can always ask them to reveal it as part of the mechanism (folk argument, Choi and Kim (1999))
  - the Ledyard etc.. argument assumes the same mechanism is used as the prior changes

An early debate about the Bayesian approach

- Critique of Bayesian approach from practical people
- Critique of the critique from Bayesian purists
- Our position:
  - "Wrong" formalism: should not be maintaining (implicit) assumption of common knowledge of a common prior....
  - alternative formalism gives more nuanced answers:
    - sometimes fully Bayesian approach gives dominant strategies even allowing reports of beliefs
    - sometimes it doesn't

Type Spaces

- i's *type* is $t_i \in T_i$
- t_i includes description of
  - payoff type: $\widehat{\theta}_i : T_i \rightarrow \Theta_i$
    - $\widehat{\theta}_i(t_i)$ is i's payoff type of t_i
  - beliefs about types of other players: $\widehat{\pi}_i : T_i \rightarrow \Delta(T_{-i})$
    - $\widehat{\pi}_i(t_i)$ is i's belief type of t_i
- *type space* is a collection

$$\mathcal{T} = \left\{ T_i, \widehat{\theta}_i, \widehat{\pi}_i \right\}_{i=1}^{I}$$

Mechanism

A mechanism $\mathcal{M}$ is a collection $\left((M_i)_{i=1}^{I}, g\right)$

- each M_i is a message set
- outcome function $g : M \to Y$

Equilibrium

type space $\mathcal{T}$ and mechanism $\mathcal{M}$ define incomplete information game $(\mathcal{T}, \mathcal{M})$

- i's strategy: $\sigma_i : T_i \to \Delta(M_i)$
- strategy profile σ is a (Bayesian Nash) equilibrium if $\sigma_i(m_i|t_i) > 0$ implies m_i is in

$$\arg\max_{m_i'} \sum_{t_{-i}, m_{-i}} \widehat{\pi}_i [t_i](t_{-i}) \left(\prod_{j \neq i} \sigma_j (m_j | t_j) \right) u_i \left(g\left(m_i', m_{-i}\right), \widehat{\theta}(t) \right)$$

Questions - Informal

1. ROBUST PARTIAL IMPLEMENTATION: Does there exist a mechanism such that, whatever agents' beliefs and higher order beliefs about other agents' types, there EXISTS an equilibrium that is consistent with the social choice correspondence?
2. ROBUST FULL IMPLEMENTATION: Does there exist a mechanism such that, whatever agents' beliefs and higher order beliefs about other agents' types, ALL equilibria are consistent with the social choice correspondence?

Questions - Formal

1. ROBUST PARTIAL IMPLEMENTATION: Does there exist a mechanism $\mathcal{M}$ such that for all type spaces $\mathcal{T}$, there EXISTS an equilibrium σ of $(\mathcal{T}, \mathcal{M})$ satisfying

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F\left(\widehat{\theta}(t)\right)?$$

2. ROBUST FULL IMPLEMENTATION: Does there exist a mechanism $\mathcal{M}$ such that for all type spaces $\mathcal{T}$, ALL equilibria σ of $(\mathcal{T}, \mathcal{M})$ satisfy

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F\left(\widehat{\theta}(t)\right)?$$

Question 1: Robust Partial Implementation

DEFINITION. Social choice *function* $f : \Theta \rightarrow Y$ is ex post incentive compatible [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'.

DEFINITION. Social choice correspondence F is attainable in ex post equilibrium if there exists social choice function f such that f is EPIC and $f\left(\theta\right) \in F\left(\theta\right)$ for all θ.

Question 1: Robust Partial Implementation

PROPOSITION. If F is attainable in ex post equilibrium, then there exists a mechanism $\mathcal{M}$ such that for all type spaces $\mathcal{T}$, there exists an equilibrium σ of $(\mathcal{T}, \mathcal{M})$, with

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F\left(\widehat{\theta}(t)\right).$$

(TRIVIAL) PROOF. Consider the mechanism where each agent reports only his payoff type. Consider "truthtelling strategy" profile $\sigma_i^*(\theta_i|t_i) = 1$ if $\theta_i = \widehat{\theta}_i(t_i)$.
CONVERSE IS FALSE: can show by example that incentive constraints can be relaxed by having agents report more than their payoff types.

Separable SCC and environment

- Separability is satisfied if

$$
\begin{aligned}
Y &= Y_0 \times Y_1 \times \ldots \times Y_I; \\
\widetilde{u}_i &: Y_0 \times Y_i \times \Theta \rightarrow \mathbb{R} \\
u_i\left(\left(y_0, y_1, \ldots, y_I\right), \theta\right) &= \widetilde{u}_i\left(y_0, y_i, \theta\right) \text{ for all } i \text{ and } y \\
f_0 &: \Theta \rightarrow Y_0 \\
F(\theta) &= f_0(\theta) \times F_1(\theta) \times \ldots \times F_I(\theta)
\end{aligned}
$$

- Satisfied if F single valued, quasi-linear environment w/o budget restrictions (e.g., F^*)
- Not satisfied, e.g., in quasi-linear environment with budget restrictions (e.g., F^{**})

Question 1: Robust Partial Implementation

PROPOSITION. If (1) F is attainable in ex post equilibrium, then (2) there exists a mechanism $\mathcal{M}$ such that for all type spaces $\mathcal{T}$, there exists an equilibrium σ of $(\mathcal{T}, \mathcal{M})$, with

$$\sigma\left(m|t\right) > 0 \Rightarrow g\left(m\right) \in F\left(\widehat{\theta}\left(t\right)\right).$$

If the environment is separable, then the converse is true (i.e., (2) $\Rightarrow$ (1))

Idea of Proof

Consider type space where agents $-i$ are known to be θ_{-i}.
There must exist $g^{i,\theta_{-i}} : \Theta_i \to Y_i$ such that

$$\widetilde{u}_i\left(f_0\left(\theta_i, \theta_{-i}\right), g\left(\theta_i\right), \left(\theta_i, \theta_{-i}\right)\right) \geq \widetilde{u}_i\left(f_0\left(\theta_i', \theta_{-i}\right), g\left(\theta_i'\right), \left(\theta_i, \theta_{-i}\right)\right)$$

for all i, θ_i, θ_i'.
Now let

$$f\left(\theta\right) = \left(f_0\left(\theta\right), g_1^{1,\theta_{-1}}\left(\theta_{-1}\right), .., g_i^{i,\theta_{-i}}\left(\theta_{-i}\right), .., g_I^{I,\theta_{-I}}\left(\theta_{-I}\right)\right)$$

Public Good Example

Efficiency is always attainable

A Key "Epistemic" Result: Informal

A message m_i can be sent by an agent with payoff type θ_i in an equilibrium on some type space if and only if m_i is "incomplete information rationalizable" for θ_i in the following sense:

1. First, suppose that every payoff type θ_i could send any message m_i
2. Delete those messages m_i that are not best responses to some conjecture over payoff type - message pairs of the opponents that have not yet been deleted.
3. Repeat step 2 until you converge

This is a belief-free incomplete information version of iterated deletion of strictly dominated strategies

Rationalizable Messages

let $S_i^{\mathcal{M},0}(\theta_i) = M_i$,

$$S_i^{\mathcal{M},k+1}(\theta_i) = \left\{ m_i \;\middle|\; \begin{array}{l} \exists\ \mu_i \in \Delta(\Theta_{-i} \times M_{-i}) \text{ s.t.:} \\ (1)\ \mu_i(\theta_{-i}, m_{-i}) > 0 \Rightarrow m_{-i} \in S_{-i}^{\mathcal{M},k}(\theta_{-i}) \\ (2)\ m_i \in \underset{m_i'}{\arg\max} \sum\limits_{\theta_{-i}, m_{-i}} \mu_i(\theta_{-i}, m_{-i})\, u_i(g(m_i', m_{-i}), \theta) \end{array} \right\}$$

$$S_i^{\mathcal{M}}(\theta_i) = \bigcap_{k \geq 0} S_i^{\mathcal{M},k}(\theta_i)$$

$S_i^{\mathcal{M}}(\theta_i)$ are *rationalizable actions* of payoff type θ_i in $\mathcal{M}$

Epistemic Foundations Result

PROPOSITION. $m_i \in S_i^{\mathcal{M}}(\theta_i)$ if and only if there exist

1. a type space $\mathcal{T}$
2. an equilibrium σ of $(\mathcal{T}, \mathcal{M})$ and
3. a type $t_i \in T_i$, such that
   3.1 $\sigma_i(m_i | t_i) > 0$ and
   3.2 $\widehat{\theta}_i(t_i) = \theta_i$.

Brandenburger and Dekel (1987), Battigalli (1996), Battigalli and Siniscalchi (2003).

Epistemic Foundations: Proof

If $m_i \in S_i^{\mathcal{M}}(\theta_i)$, then $\exists\ \mu_i^{m_i,\theta_i} \in \Delta(\Theta_{-i} \times M_{-i})$ s.t.

$$\mu_i(\theta_{-i}, m_{-i}) > 0 \Rightarrow m_{-i} \in S_{-i}^{\mathcal{M}}(\theta_{-i})$$

$$m_i \in \arg\max_{m_i'} \sum_{\theta_{-i}, m_{-i}} \mu_i^{m_i,\theta_i}(\theta_{-i}, m_{-i})\, u_i\left(g\left(m_i', m_{-i}\right), \theta\right)$$

Construct type space:

- $T_i = \left\{(\theta_i, m_i) \in \Theta_i \times M_i \,\middle|\, m_i \in S_i^{\mathcal{M}}(\theta_i)\right\}$
- $\widehat{\theta}_i(\theta_i, m_i) = \theta_i$
- $\widehat{\pi}_i((\theta_i, m_i))(\theta_j, m_j)_{j \neq i} = \mu_i^{m_i,\theta_i}(\theta_{-i}, m_{-i})$

Construct equilibrium:

- Type (θ_i, m_i) sends message m_i

Equivalence

So the following statements of the robust full implementation question are equivalent:

- Does there exist a mechanism $\mathcal{M}$ such that for all type spaces $\mathcal{T}$, ALL equilibria σ of $(\mathcal{T}, \mathcal{M})$ satisfy

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F\left(\widehat{\theta}(t)\right)?$$

- Does there exist a mechanism $\mathcal{M}$ such that

$$m \in S^{\mathcal{M}}(\theta) \Rightarrow g(m) \in F(\theta)?$$

Example

- $\beta^0(\theta_i) = [0,1]$
- $\beta^k(\theta_i)$ is set of reports i might send for some conjecture $\lambda_i(\theta'_{-i}, \theta_{-i})$ over his opponents' types θ_{-i} and reports θ'_{-i}, with restriction on conjecture $\lambda_i(\theta'_{-i}, \theta_{-i})$ that each type θ_j of agent j sends a message in $\beta^{k-1}(\theta_j)$.

Rationalizability

- agent i conjectures that other agents have type θ_{-i} and report θ'_{-i}
- agent i with type θ_i has best response

$$\theta'_i = \theta_i + \gamma \sum_{j \neq i} \left(\theta_j - \theta'_j \right).$$

linear best response allows us characterize $\beta^k(\theta_i)$:

$$\beta^k(\theta_i) = \left[\underline{\beta}^k(\theta_i), \overline{\beta}^k(\theta_i) \right]$$

Rationalizability

$$\overline{\beta}^{k}(\theta_i) = \min\left\{1, \theta_i + \gamma \max_{\left\{\left(\theta'_{-i},\theta_{-i}\right):\theta'_j\in\beta^k(\theta_j) \text{ for all } j\neq i\right\}} \sum_{j\neq i}\left(\theta_j - \theta'_j\right)\right\}$$

$$= \min\left\{1, \theta_i + \gamma \max_{\theta_{-i}} \sum_{j\neq i}\left(\theta_j - \underline{\beta}^{k-1}(\theta_j)\right)\right\}$$

$$\overline{\beta}^{k}(\theta_i) = \min\{1, \theta_i + \gamma \max_{\theta_{-i}} \sum_{j\neq i}\left(\theta_j - \underline{\beta}^{k-1}(\theta_j)\right)\}$$

Rationalizability

rewriting:

$$\overline{\beta}^{k}\left(\theta_i\right)=\min\left\{1,\theta_i+\left(\gamma\left(I-1\right)\right)^{k}\right\},$$

and likewise

$$\underline{\beta}^{k}\left(\theta_i\right)=\max\left\{0,\theta_i-\left(\gamma\left(I-1\right)\right)^{k}\right\}.$$

thus

$$\theta_i'\neq\theta_i\Rightarrow\theta_i'\notin\beta^{k}\left(\theta_i\right)$$

for sufficiently large k, provided that

$$\gamma<\frac{1}{I-1}$$

On the other hand...

- but now suppose that $\gamma \geq \frac{1}{I-1}$
- *each* type θ_i convinced that others' types are $\theta_j = \frac{1}{2} + \frac{1}{\gamma(I-1)}\left(\frac{1}{2} - \theta_i\right)$
- now $\theta_i + \gamma(I-1)\left[\frac{1}{2} + \frac{1}{\gamma(I-1)}\left(\frac{1}{2} - \theta_i\right)\right] = \frac{1}{2}\left[1 + \gamma(I-1)\right]$
- types cannot be distinguished in direct or any other mechanism....

Efficient Public Good Example Summary

- Robust partial implementation is always possible
- Robust full implementation is possible if and only if $\gamma \leq \frac{1}{I-1}$

In general....

- Each Θ_i is a compact subset of the real line
- Agent i's preferences depend on θ through $h_i : \Theta \longrightarrow \mathbb{R}$
- Preferences are single crossing in $h_i(\theta)$
- Robust implementation is possible in the *direct* mechanism if strict EPIC and the "contraction property" hold.
- Robust implementation is impossible in *any* mechanism if either strict EPIC or the "contraction property" fails.
- "Contraction Property" $=$ not too much interdependence in agents' preferences

The Other Papers

- "Ex Post Implementation" (GEB 08);
  - other ex post equilibria? wrong question?
- "Robust Virtual Implementation" (TE 09);
  - "virtual" loses its bite under robustness; same conclusion in leading public example
- "Robust Implementation in General Mechanisms;"
  - relation to the classic Full Implementation Literature
- Common Prior Assumption (JEEA P&P 08)
  - robust partial implementation results unchanged; robust full implementation changed a lot
- Dynamic Mechanisms (AER P&P 07, Penta, Mueller)
  - lots of extra power

Open Questions

- Non-Separable Case
- Richer Objectives like Revenue Maximization (Ely-Chung 07)
- Intermediate versions of informational robustness (Artemov-Kunimoto-Serrano)
- Robust Predictions of Economic Models