Robust Mechanism Design

Stephen Morris

Atlanta Winter Meetings of the Econometric Society
January 2010
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 informational assumptions.
"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)
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 $i \in I = \{1, 2, ..., I\}$
- $i$'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 $u_i : Y \times \Theta \rightarrow \mathbb{R}$ (allows "interdependent preferences")
- Correspondence $F : \Theta \rightarrow 2^Y \setminus \emptyset$
  - $F(\theta) \subseteq Y$ 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
  $$u_i (y, \theta) = u_i ((q, (x_1, ..., x_I)), \theta) = v_i (\theta) \cdot q + x_i$$
  - $v_i (\theta) = \theta_i + \gamma \sum_{j \neq i} \theta_j$
  - $\gamma \geq 0$ (non-negative interdependence)
  - $\gamma = 0$: private values

Robust Mechanism Design
Leading Example: Public Good and Quasi-Linear Preferences

- planner maximizes \( \left( \sum_i v_i(\theta) \right) q - \frac{1}{2} q^2 \)

- \( q^*(\theta) = (1 + \gamma (l - 1)) \sum_{i=1}^l \theta_i \)

- two classic social choice correspondences

\[
F^*(\theta) = \left\{ (q, x) \in Y \mid q = q^*(\theta) \right\}
\]

\[
F^{**}(\theta) = \left\{ (q, x) \in F^*(\theta) \mid \sum_{i=1}^l x_i = \bar{x} \right\}
\]
Early Debate

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 practice, the "prior" is not known to the planner
  - if social choice correspondence must be implemented without knowing the prior, dominant strategies is required
An early debate about the Bayesian approach

- Critique of Bayesian approach from "practical people"
  - in practice, the "prior" is not known to the planner
  - if social choice correspondence must be implemented without knowing the prior, dominant strategies is required

- 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
Early Debate

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

\[
\mathcal{T} = \left\{ T_i, \hat{\theta}_i, \hat{\pi}_i \right\}_{i=1}^I
\]
Mechanism

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

- each \( M_i \) is a message set
- outcome function \( g : M \rightarrow Y \)
Equilibrium

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

- $i$'s strategy: $\sigma_i : T_i \rightarrow \Delta (\mathcal{M}_i)$
- strategy profile $\sigma$ is a (Bayesian Nash) equilibrium if $\sigma_i (m_i|t_i) > 0$ implies $m_i$ is in

$$\arg \max_{m_i'} \sum_{t_i, t_{-i}, m_{-i}} \hat{\pi}_i [t_i] (t_{-i}) \left( \prod_{j \neq i} \sigma_j (m_j|t_j) \right) u_i \left( g (m'_i, m_{-i}), \hat{\theta} (t) \right)$$
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 $\sigma$ of $(\mathcal{T}, \mathcal{M})$ satisfying

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F(\hat{\theta}(t))$$

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

$$\sigma(m|t) > 0 \Rightarrow g(m) \in F(\hat{\theta}(t))$$
Question 1: Robust Partial Implementation

**DEFINITION.** Social choice function $f : \Theta \rightarrow Y$ is ex post incentive compatible (EPIC) if

$$u_i (f (\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) \geq u_i (f (\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i}))$$

for all $i, \theta_i, \theta_{-i}$ and $\theta'_i$.

**DEFINITION.** Social choice correspondence $F$ is attainable in ex post equilibrium if there exists social choice function $f$ such that $f$ is EPIC and $f (\theta) \in F (\theta)$ for all $\theta$. 
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 $\sigma$ of $(\mathcal{T}, \mathcal{M})$, with

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

(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 $\theta_i = \hat{\theta}_i(t_i)$.

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

\[ Y = Y_0 \times Y_1 \times \ldots \times Y_I; \]
\[ \tilde{u}_i : Y_0 \times Y_i \times \Theta \to \mathbb{R} \]
\[ u_i ((y_0, y_1, \ldots, y_I), \theta) = \tilde{u}_i (y_0, y_i, \theta) \text{ for all } i \text{ and } y \]
\[ f_0 : \Theta \to Y_0 \]
\[ F(\theta) = f_0(\theta) \times F_1(\theta) \times \ldots \times F_I(\theta) \]

- 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 $\sigma$ of $(\mathcal{T}, \mathcal{M})$, with

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

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 $\theta_{-i}$. There must exist $g^{i,\theta_{-i}} : \Theta_i \to Y_i$ such that

$$\tilde{u}_i ( f_0 (\theta_i, \theta_{-i}), g (\theta_i), (\theta_i, \theta_{-i})) \geq \tilde{u}_i ( f_0 (\theta'_i, \theta_{-i}), g (\theta'_i), (\theta_i, \theta_{-i}))$$

for all $i, \theta_i, \theta'_i$.

Now let

$$f (\theta) = \left( f_0 (\theta), g^{1,\theta_1}_{-1} (\theta_{-1}), \ldots, g^{i,\theta_{-i}}_{-i} (\theta_{-i}), \ldots, g^{I,\theta_{-I}}_{-I} (\theta_{-I}) \right)$$
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 $\theta_i$ in an equilibrium on some type space if and only if $m_i$ is "incomplete information rationalizable" for $\theta_i$ in the following sense:

1. First, suppose that every payoff type $\theta_i$ 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

\[
\begin{align*}
\text{let } S_i^{M,0} (\theta_i) &= M_i, \\
S_i^{M,k+1} (\theta_i) &= \\
\left\{ m_i \right\} &\quad \exists \mu_i \in \Delta (\Theta_{-i} \times M_{-i}) \text{ s.t.:} \\
(1) &\quad \mu_i (\theta_{-i}, m_{-i}) > 0 \implies m_{-i} \in S_{-i}^{M,k} (\theta_{-i}) \\
(2) &\quad m_i \in \arg \max \sum_{m_{-i}'} \mu_i (\theta_{-i}, m_{-i}) u_i (g (m_{i}', m_{-i}), \theta) \}
\end{align*}
\]

\[
S_i^M (\theta_i) = \bigcap_{k \geq 0} S_i^{M,k} (\theta_i)
\]

\(S_i^M (\theta_i)\) are rationalizable actions of payoff type \(\theta_i\) in \(M\)
PROPOSITION. \( m_i \in S_i^M (\theta_i) \) if and only if there exist

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

Epistemic Foundations: Proof

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

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

\[
m_i \in \arg\max \sum_{m_i'} \mu_{i}^{m_i,\theta_i} (\theta_{-i}, m_{-i}) u_i \left( g \left( m_i', m_{-i} \right), \theta \right)
\]

Construct type space:

- \( T_i = \{ (\theta_i, m_i) \in \Theta_i \times M_i \mid m_i \in S_i^M (\theta_i) \} \)
- \( \hat{\theta}_i (\theta_i, m_i) = \theta_i \)
- \( \hat{\pi}_i ((\theta_i, m_i)) (\theta_j, m_j)_{j \neq i} = \mu_{i}^{m_i,\theta_i} (\theta_{-i}, m_{-i}) \)

Construct equilibrium:

- Type \( (\theta_i, m_i) \) sends message \( m_i \)
Robust Full Implementation: Epistemic Result

**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 $\sigma$ of $(\mathcal{T}, \mathcal{M})$ satisfy

  $$\sigma(m|t) > 0 \Rightarrow g(m) \in F(\hat{\theta}(t))$$

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

  $$m \in S^\mathcal{M}(\theta) \Rightarrow g(m) \in F(\theta)$$
Robust Full Implementation: Substantive Results

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 \( \theta_{-i} \) and reports \( \theta'_{-i} \), with restriction on conjecture \( \lambda_i (\theta'_{-i}, \theta_{-i}) \) that each type \( \theta_j \) of agent \( j \) sends a message in \( \beta^{k-1} (\theta_j) \).
Rationalizability

- agent $i$ conjectures that other agents have type $\theta_{-i}$ and report $\theta'_i$.
- agent $i$ with type $\theta_i$ has best response

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

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

$$\beta^k(\theta_i) = \left[ \underline{\beta}^k(\theta_i), \overline{\beta}^k(\theta_i) \right]$$
Robust Full Implementation: Substantive Results

Rationalizability

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

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

\[ \bar{\beta}^k (\theta_i) = \min \{ 1, \theta_i + \gamma \max_{\theta_i} \sum_{j \neq i} \left( \theta_j - \bar{\beta}^{k-1}(\theta_j) \right) \} \]
Rationalizability

rewriting:

\[ \bar{\beta}^k (\theta_i) = \min \left\{ 1, \theta_i + (\gamma (l - 1))^k \right\}, \]

and likewise

\[ \beta^k (\theta_i) = \max \left\{ 0, \theta_i - (\gamma (l - 1))^k \right\}. \]

thus

\[ \theta'_i \neq \theta_i \Rightarrow \theta'_i \not\in \beta^k (\theta_i) \]

for sufficiently large \( k \), provided that

\[ \gamma < \frac{1}{l - 1} \]
On the other hand...

- but now suppose that $\gamma \geq \frac{1}{l-1}$
- each type $\theta_i$ convinced that others’ types are
  $\theta_j = \frac{1}{2} + \frac{1}{\gamma(l-1)} \left( \frac{1}{2} - \theta_i \right)$
- now $\theta_i + \gamma (l - 1) \left[ \frac{1}{2} + \frac{1}{\gamma(l-1)} \left( \frac{1}{2} - \theta_i \right) \right] = \frac{1}{2} [1 + \gamma (l - 1)]$
- 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 $\Theta_i$ is a compact subset of the real line
- Agent $i$’s preferences depend on $\theta$ through $h_i : \Theta \to \mathbb{R}$
- 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