Behavioral Mechanism Design

Lecture 1/2

Paul J. Healy
healy.52@osu.edu

The Ohio State University
Department of Economics

UA IBE
October 2008

Outline

1. Lecture 1: Theory
   1. Mechanism Design History (Briefly)
   2. Mechanism Design Theory
   3. Behavioral M.D. Methodology
2. Lecture 2: Applications
   1. Stable Public Goods Mechanisms
   2. Prediction Markets
   3. Contracting within NASA

The Socialist Calculation Debate

- Marx & Friends (1800s-1918)
  - The labor theory of value
  - To revolt or not (and other annoying details)
  - How to deal with the incentive issues?
- von Mises, von Hayek & Friends (1920–1945)
  - Capital prices give info on productivity & scarcity
  - Example: oil vs. wine
  - Consumption-goods prices give info on welfare
  - Prices coordinate decentralized production plans
  - Opponents: P.O. = sol'n to system of eq's
  - Response: Too much info needed to solve the system

The Socialist Calculation Debate

- Oscar Lange & Abba Lerner (1934-1938)
  - Market Socialism: Admits prices are needed
  - Social planner 'simulates' markets (tatonnemont)
  - Prices are just accounting tools (incentives??)
  - Problem: no financial markets to assess investment
  - Lange: Stalin-Roosevelt & Polish head-of-state

Theory all rested on Pareto, Edgeworth, etc.

Leonid Hurwicz

- Born 1918, right before the Bolshevik Revloution
- 1938: Law degree at Warsaw. No econ degree.
- 1938: LSE under Hayek, met with Mises
- 1941: RA to Samuelson & Lange
- 1942-46: Research associate under Marschak & Koopmans
- 1959: First publication on mechanism design theory
- 1961-2008: Minnesota

Mechanism Design Setup

- Outcome space: X
- Preferences over X: $u_i(x|\theta_i)$
- Type space: $\Theta = \times_{i=0}^{n}\Theta_i$ (0 = planner)
- Message space: $M = \times_{i=0}^{n} M_i$
- Adjustment rule: $m_i^t = \alpha_i(m^{t-1}, \theta_i)$
- Steady-state message: $m^* = \alpha(m^*, \theta)$
- Equilibrium correspondence: $\mu : \Theta \twoheadrightarrow M$
  $\mu(\theta) \subseteq M$ identifies steady-state messages
- Outcome function: $g : M \rightarrow X$
- MECHANISM: (M, g)
- ASSUMED: α (and/or μ)

The Competitive Mechanism

- X = all possible net trade vectors
- M_i $(i > 0)$ = excess demands for each good
- M_0 = prices for each good
- $\alpha_0(m^{t-1})$ = period-t tatonnemont prices
- $\alpha_i(m^{t-1})$ = excess demand given period-$t-1$ prices
- $\mu_0(\theta)$ = Walrasian price vector
- $\sum_i \mu_i(\theta) = \vec{0}$
- $g(m^*) = (m_1^*, \ldots, m_n^*)$
- Satisfactory (Pareto optimal outcomes)
- Privacy-preserving
- Informationally efficient (ask Mark)

The Static Approach

- Typically remove the planner (no pvt info, no adjusting)
- μ = set of (Nash/DomStrat/Bayes)-equilibrium strategies of the induced game:
  - Players: $\{1, \ldots, n\}$
  - Strategies: M_i
  - Utility: $u_i(g(m)|\theta_i)$
- Predicted outcomes: $g(\mu(\theta)) \subseteq X$

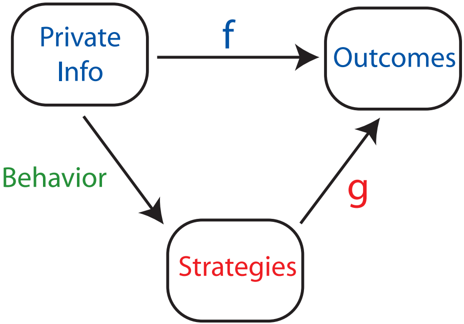
Social Choice Correspondences

- What is the objective of the social planner??
- Abstract notion: $f(\theta) \subseteq X$ (for any $\theta \in \Theta$)
- f = social choice correspondence (SCC) or function (SCF)
- Examples:
  - Pareto optimal x
  - Planer's most-preferred x
  - 'Equitable' x (egalitarian, envy-free, etc.)

Implementation & Mechanism Design

- A mechanism (M, g) fully implements a SCC f under μ if, for all $\theta \in \Theta$,
$$g(\mu(\theta)) = f(\theta)$$
- Partial implementation:
$$g(\mu(\theta)) \subseteq f(\theta)$$
or even
$$g(\mu(\theta)) \cap f(\theta) \neq 0$$
- 'Mechanism design' vs. 'implementation'

The Mount-Reiter Diagram



Dominant Strategy Implementation

- Strongest solution concept: dominant strategies
- Gibbard (73) & Satterthwaite (75): Revelation principle
  - Suppose $\Gamma = (M, g)$ has a DSE $\mu_\Gamma(\theta)$ for every θ
  - Consider new mechanism: $\Gamma' = (\Theta, h)$ where
    - Agents announce a type $(m_i \in \Theta_i)$
    - Mechanism picks $h(\theta) = g(\mu_\Gamma(\theta))$
  - Idea: Γ' plays μ_Γ for you
  - μ_Γ is a DSE in $\Gamma \Rightarrow \theta$ is a DSE in Γ'

Multiple Equilibrium Problems

$$X = \{a, b, c, d, e, p, q, r\}$$

Γ	B1	B2	B3
A1	a	d	e
A2	c	b	p
A3	a	b	e

$$\theta_1 = \begin{bmatrix} q \\ a, c, e \\ d, b, p \\ r \end{bmatrix} \qquad \theta_1' = \begin{bmatrix} c, b, p \\ a, d, e \\ q, r \end{bmatrix}$$

$$\theta_2 = \begin{bmatrix} r \\ d, a, e \\ b, c, p \\ q \end{bmatrix} \qquad \theta_2' = \begin{bmatrix} d \\ b, c \\ a \\ e, p, q, r \end{bmatrix}$$

Γ'	$\hat{\theta}_2$	$\hat{\theta}_2'$
$\hat{\theta}_1$	e	b
$\hat{\theta}_1'$	p	b

$$\begin{aligned} f(\theta_1, \theta_2) &= \{a, e\} \\ f(\theta_1', \theta_2) &= \{c, p\} \\ f(\theta_1, \theta_2') &= \{b, d\} \\ f(\theta_1', \theta_2') &= \{b\}. \end{aligned}$$

Γ''	$\hat{\theta}_2$	$\hat{\theta}_2'$
$\hat{\theta}_1$	a	d
$\hat{\theta}_1'$	c	b

Revelation Principle Cont'd

- Revelation principle holds iff $\mu_i(\theta)$ depends only on θ_i
  - DSE & BNE: yes
  - Nash, SPNE, IESDS: no

$$X = \{a, b, c, d\}$$

$$\theta_1 = \begin{bmatrix} a \\ c \\ d \\ b \end{bmatrix} \quad \theta_2 = \begin{bmatrix} c \\ d \\ a \\ b \end{bmatrix} \quad \theta_2' = \begin{bmatrix} b \\ a \\ d \\ c \end{bmatrix}$$

$$f(\theta_1, \theta_2) = \{a\}$$
$$f(\theta_1, \theta_2') = \{d\}$$

Γ	L	R
U	a	b
D	c	d

Γ'	$\hat{\theta}_2$	$\hat{\theta}_2'$
$\hat{\theta}_1$	a	d

Dom. Strat. Results

- f is dictatorial if $\exists i$ s.t. $\forall\theta$

$$f(\theta) \in \arg\max u_i(x|\theta_i)$$

- Gibbard-Satterthwaite: When all strict preferences are admissible, if $|X| \geq 3$ and if f can be implemented in DSE then f is dictatorial.
- Zhou 91 & others: Also true for 'economic' environments

DSE: Quasilinear Prefs

- Quasilinear prefs: $u_i(x, t_i|\theta_i) = v_i(x|\theta_i) + t_i$
- Roberts 79: If $|X| \geq 3$, all quasilinear prefs are admissible, and f can be implemented in DSE then $\exists(\lambda_1, \ldots, \lambda_n) \gneqq 0$ and v_0 s.t.

$$x(\hat{\theta}) = \arg\max_{x \in X} \left[v_0(x) + \sum_i \lambda_i v_i(x|\hat{\theta}_i) \right]$$

and if $\lambda_i > 0$ then

$$t_i(\hat{\theta}) = \frac{1}{\lambda_i} \left(\left[v_0(x(\hat{\theta})) + \sum_{j \neq i} \lambda_j v_j(x(\hat{\theta})|\hat{\theta}_j) \right] + h_i(\hat{\theta}_{-i}) \right)$$

- Groves (1973) mechanisms: $\lambda_i = 1$ & $v_0 \equiv 0$
- Green & Laffont

VCG Mechanism

- Groves mechs need not be balanced
  - Ex: $h_i(\hat{\theta}_{-i}) = \$1,000,000 \forall i$
- Clarke 71: use

$$h_i(\hat{\theta}_{-i}) = -\max_{x \in X}\left[v_0(x) + \sum_{j \neq i} \lambda_j v_j(x|\hat{\theta}_j)\right]$$

- Guarantees $t_i(\hat{\theta}) \leq 0$
- Vickrey 1961

Issues with VCG

- Generally not fully efficient (not balanced)
- Often not individually rational (right notion??)
- Distribution of taxes can be highly inequitable
- The equilibrium may be 'weak' & other NE may exist
- May be difficult to communicate preferences
- Agents may have reasons to conceal their type
- Assumes no bankruptcy/liquidity constraints
- Finding $\arg\max_{x \in X} \sum_i v_i(x|\hat{\theta}_i)$
- Shill bidding can be advantageous
- Collusion can be profitable
- Dynamic allocation problems

Bayesian Implementation

- Agents' types are drawn from $\rho_i(\theta_i)$
- Each ρ_i is common knowledge (incl. the planner)
- General preferences: see Jackson 91
- Quasilinear prefs: d'Aspremont & Gerard-Varet
  - Let $W_i(\theta_i) = E_{\theta_{-i}}[\sum_{j \neq i} v_j(x(\theta_i, \theta_{-i})|\theta_j)]$
  - Transfer:
$$t_i(\theta) = W_i(\theta_i) - \frac{1}{n-1} \sum_{j \neq i} W_j(\theta_j)$$
  - Sum of transfers = zero

Nash Implementation

- Partial implementation is easy:
- Each agent reports $m_i = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$
- If $m_1 = m_2 = \cdots = m_n$ then select $f(m_1)$
- If any disagreement, shoot everyone
- 'Truth-telling' is an equilibrium, giving $f(\theta)$
- Consistent lies are also equilibria, giving $f(\theta')$
- The main problem is these bad equilibria

Maskin Monotonicity

- Let $L_i(x, \theta_i) = \{y \in X : u_i(x|\theta_i) \geq u_i(y|\theta_i)$
- f is monotonic if $x \in f(\theta)$ and $L_i(x|\theta_i) \subseteq L_i(x|\theta_i')\forall i$ implies $x \in f(\theta')$.
- If f can be fully implemented in NE then f is monotonic
- Suppose f is not monotonic. Then $\exists x, \theta$ with $x \in f(\theta)$ and $L_i(x|\theta_i) \subseteq L_i(x|\theta_i')\forall i$ but $x \notin f(\theta')$
- But if x is getting better for everyone, then the alternatives players could get by deviating aren't any better under θ'
- Thus, NE under θ are NE under θ'
- Full implementation requires that $x \in f(\theta')$

Maskin's Mechanism

- $m_i = (\hat{\theta}^i, \hat{x}^i, k^i)$
- If all m_i agree and $\hat{x} \in f(\hat{\theta})$ then pick $\hat{x}$
- If one person j differs:
  - If j's deviation is good for himself (by $\hat{\theta}^i$), pick $\hat{x}^i$
  - If j's deviation is bad for himself (by $\hat{\theta}^i$), pick $\hat{x}^j$
- If two or more disagree, pick $\hat{x}^{\arg\max_i k^i}$
- Clearly, 'truth-telling' is a NE
- Bad equilibria eliminated (ex: coordinated lie)

Groves-Ledyard 1977

$m_i \in \mathbb{R}^1$

$y(m) = \sum_i m_i$

$t_i(m) = q_i(m_{-i})y(m) + g_i(m)$, where

$$q_i(m_{-i}) = \frac{\kappa}{n}$$

$$g_i(m) = \frac{\gamma}{2}\left[\frac{n-1}{n}(m_i - \mu_{-i})^2 - \sigma^2_{-i}\right]$$

$\mu_{-i} = \frac{1}{n-1}\sum_{j\neq i} m_j$ $\qquad$ $\sigma^2_{-i} = \frac{1}{n-2}\sum_{j\neq i}(m_j - \mu_{-i})^2$

Notes:

- Implements PO outcomes
- Not Lindahl $\Rightarrow$ not IR (see Hurwicz 79)

Walker 1981

$m_i \in \mathbb{R}^1$

$y(m) = \sum_i m_i$

$t_i(m) = q_i(m_{-i})y(m) + g_i(m)$, where

$$q_i(m_{-i}) = \frac{\kappa}{n} + m_{i+1} - m_{i+2}$$

$$g_i(m) \equiv 0$$

Notes:

- Implements Lindahl allocations ($\Rightarrow$ IR)
- Uses only one dimension
- Agents are 'price-taking'

Hurwicz 1979

$m_i = (s_i, z_i) \in \mathbb{R}^2$

$y(m) = \frac{1}{n} \sum_i s_i$

$t_i(m) = q_i(m_{-i})y(m) + g_i(m)$, where

$$q_i(m_{-i}) = \frac{\kappa}{n} + z_{i+1} - z_{i+2}$$

$$g_i(m) = z_i(s_i - s_{i+1})^2 - z_{i+1}(s_{i+1} - s_{i+2})^2$$

Notes:

- Implements Lindahl using two dimensions
- $g_i = 0$ in equilibrium
- Agents are 'price-taking'

Kim 1993

$m_i = (s_i, z_i) \in \mathbb{R}^2$

$y(m) = \sum_i s_i$

$t_i(m) = q_i(m_{-i})y(m) + g_i(m)$, where

$$q_i(m_{-i}) = \tfrac{\kappa}{n} - \sum_{j \neq i} s_j + \tfrac{1}{n} \sum_{j \neq i} z_j$$

$$g_i(m) = -\tfrac{1}{2}\left(z_i - y(m)\right)^2$$

Notes:

- Implements Lindahl using two dimensions
- $g_i = 0$ in equilibrium
- Agents are ‘price-taking’

Y. Chen 2002

$m_i = (s_i, z_i) \in \mathbb{R}^2$

$y(m) = \sum_i s_i$

$t_i(m) = q_i(m_{-i})y(m) + g_i(m)$, where

$$q_i(m_{-i}) = \frac{\kappa}{n} - \gamma \sum_{j \neq i} s_j + \frac{\gamma}{n} \sum_{j \neq i} z_j$$

$$g_i(m) = -\frac{1}{2}\left(z_i - y(m)\right)^2 + \frac{\delta}{2} \sum_{j \neq i} (z_j - y(m))^2$$

Notes:

- Implements Lindahl using two dimensions
- $g_i = 0$ in equilibrium
- Agents are ‘price-taking’

Others...

Other mechanisms:

- de Trenqualye 1989
- Vega-Redondo 1989
- Kim 1996
- Corchon & Wilkie 1996

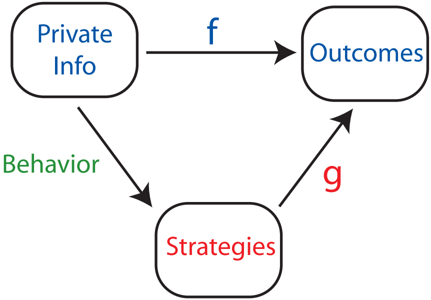
In all of these...

- Implement Lindahl using two dimensions
- $g_i = 0$ in equilibrium
- Agents are ‘price-taking’

Other Solution Concepts

- Moore & Repullo: SPNE
- Palfrey & Srivistava: Undominated NE
- Bergemann & Valimaki, Morris & Shin: Ex-post NE

Behavioral Mechanism Design



BMD: Methodology

When to use BMD?

1. Theory provides too many solutions
   - Multiple candidate mechanisms to compare
2. Theory provides no solutions
   - Use mechanisms from ‘similar’ problems

After identifying candidate mechanisms, test them!

BMD: Methodology

The ‘testbedding’ procedure

- Step 1: Identify candidate mechanism
- Step 2: Experiments, behavioral insights
- Step 3a (short-run): Modify the mechanism directly
- Step 3b (long-run): Modify the design problem and re-solve
- - New candidate mechanism
- Return to Step 2 until satisfactory

Four Applications

1. Public goods mechanisms
   - Long-run goal
   - Theory: too many solutions
   - Comparison tests, behavioral insights
2. Contracting within NASA
   - Short-run goal
   - Theory: too hard (no solutions)
   - Develop candidate mechanism
   - Test & modify
3. Prediction Markets
   - Short- & Long-run goals
   - Theory: ambiguous
   - Comparison tests, behavioral insights
4. Matching Markets (not covered)
   - Short-run goals (so far)
   - Theory: multiple solutions, trade-offs
   - Comparison tests