Distributional Preferences

May 7, 2009
Schedule

- Distributional Preferences [1-2]
  - Evidence
  - Models
- Reciprocity [1]
- Beliefs-based Preferences & Psychological Games [2-3]
Modeling Distributional Preferences

Represent Person 0’s preferences by

\[ U_0(\pi_0, \pi_1, \pi_2, \ldots), \]

where \( \pi_k \) is \( k \)'s “material payoff”.

- Supposes that preferences can be represented purely in terms of money or “material resources”
- Excludes reciprocity, procedural justice, intentionality, beliefs-based preferences
- What kind of “disinterested” distributional preferences \( (W_0(\pi_1, \pi_2, \ldots)) \) might we want to consider?
- Some (extreme) examples:
  - \( W_0 = \sum_k \pi_k \quad \Rightarrow \quad \text{surplus-maximizing, efficiency, utilitarianism} \) (if \( \pi_k \) is ‘material hedonic return’)
  - \( W_0 = \min_k \{\pi_k\} \quad \Rightarrow \quad \text{Rawlsian or maximin preferences} \)
  - \( W_0 = -\sum_k (\pi_k - \bar{\pi})^2 \quad \Rightarrow \quad \text{egalitarian preferences} \)
Modeling Distributional Preferences

Broad approach: Person 0’s ‘non-disinterested’ distributional preferences might be represented by:

\[ U_0(\pi_0, \pi_1, \pi_2, \ldots) = (1-k-l)\pi_0 + kW_0(\pi_0, \pi_1, \pi_2, \ldots) + lD_0(\pi_0-\pi_1, \pi_0-\pi_2, \ldots) \]

where \( k, l, k + l \in [0, 1] \).

Interpretation:

- First component:
- Second component:
- Third component:
Modeling Distributional Preferences

Broad approach: Person 0’s ‘non-disinterested’ distributional preferences might be represented by:

\[ U_0(\pi_0, \pi_1, \pi_2, \ldots) = (1-k-l)\pi_0 + kW_0(\pi_0, \pi_1, \pi_2, \ldots) + lD_0(\pi_0 - \pi_1, \pi_0 - \pi_2, \ldots) \]

where \( k, l, k + l \in [0, 1] \).

Interpretation:
- First component: material/self-interested utility
- Second component:
- Third component:
Modeling Distributional Preferences

Broad approach: Person 0’s ‘non-disinterested’ distributional preferences might be represented by:

\[ U_0(\pi_0, \pi_1, \pi_2, \ldots) = (1-k-l)\pi_0 + kW_0(\pi_0, \pi_1, \pi_2, \ldots) + ID_0(\pi_0 - \pi_1, \pi_0 - \pi_2, \ldots), \]

where \( k, l, k + l \in [0, 1] \).

Interpretation:

- First component: material/self-interested utility
- Second component: distributional preferences
- Third component:
Broad approach: Person 0’s ‘non-disinterested’ distributional preferences might be represented by:

\[ U_0(\pi_0, \pi_1, \pi_2, \ldots) = (1-k-l)\pi_0 + kW_0(\pi_0, \pi_1, \pi_2, \ldots) + lD_0(\pi_0-\pi_1, \pi_0-\pi_2, \ldots) \]

where \( k, l, k + l \in [0, 1] \).

Interpretation:

- First component: material/self-interested utility
- Second component: distributional preferences
- Third component: comparison utility
Charness-Rabin Preferences

Charness & Rabin (2002) basically use a simplified version of this approach to model two-person preferences:

$$U_b(\pi_a, \pi_b) = \begin{cases} \rho \pi_a + (1 - \rho)\pi_b & \text{if } \pi_b \geq \pi_a \\ \sigma \pi_a + (1 - \sigma)\pi_b & \text{if } \pi_b \leq \pi_a \end{cases}$$

Interpretation:

<table>
<thead>
<tr>
<th>Preferences</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = \sigma = 0$</td>
<td>Self-Interested</td>
</tr>
<tr>
<td>$\rho = 1, \sigma = 0$</td>
<td>Rawlsian</td>
</tr>
<tr>
<td>$\rho = \sigma = 1/2$</td>
<td>Utilitarian</td>
</tr>
<tr>
<td>$1 \geq \rho \geq \sigma \geq 0$</td>
<td>Social-Welfare Preferences</td>
</tr>
<tr>
<td>$1 \geq \rho \geq 0 \geq \sigma$</td>
<td>Difference Aversion</td>
</tr>
<tr>
<td>$\rho \geq 1 \geq 0 \geq \sigma$</td>
<td>Difference Phobia</td>
</tr>
<tr>
<td>$0 \geq \rho \geq \sigma$</td>
<td>Envy</td>
</tr>
</tbody>
</table>
Experiments

- Experiments that can be used to give parameter estimates:
  - Andreoni & Miller (2002)
  - Charness & Rabin (2002)
  - Fisman, Kariv, & Markovits (2005) – emphasize heterogeneity

- Some of the data: whiteboard

- Evidence of positive $\rho$, e.g. give up $2$ so other can have $3$, avg. $\sim 0.4$

- $\sigma$: certainly $< \rho$, maybe $< 0$—data not consistent (perhaps b/c of heterogeneity)
Data seems to be all over the place

Fehr-Schmidt (1999) use similar model to explain puzzle:

Substantial evidence that fairness motives affect behavior...

But plenty of situations where they don’t matter.

E.g. in competitive experimental markets with complete contracts, trading well-defined homogeneous good: people act like own-money maximizers, converge quickly to equilibrium, even if equilibrium implies unequal distribution of gains-from-trade.