Aggregate Matchings

Federico Echenique    SangMok Lee    Matthew Shum

California Institute of Technology

NBER Market Design Conference – Oct. 8, 2010
Revealed Preference Theory

- Individual behavior: Consumers, General Decision Makers.
- Applications: Consumption, Psychiatric Patients, Kids, Rats, Pigeons...
This paper

Revealed preference theory for matching markets:

When are observed matchings compatible with the theory of two-sided matching? (or what are the empirical implications of matching theory)
Revealed Preference Theory

Why is it useful?

- Test the theory.
- Estimate parameters.
Why is it useful?

Test for TU vs. NTU.
Revealed Preference Theory

Why is it useful?

Test for TU vs. NTU.

1. Marriage in Chicago vs. marriage in Berkeley
2. Other markets w/money but imperfect transfers (utility frontier).
Revealed Preference Theory

Why is it useful?

Test for TU vs. NTU.

1. Marriage in Chicago vs. marriage in Berkeley
2. Other markets w/money but imperfect transfers (utility frontier).

Most work on mkt. design uses NTU. Econometric work uses TU.
Why is it hard?
Revealed Preference Theory

Why is it hard?

- Standard revealed preference:
  Alice buys tomatoes when carrots are available
  \[ (T \succ_A C). \]
Why is it hard?

- Standard revealed preference:
  Alice buys tomatoes when carrots are available
  \[ (T \succ_A C). \]

- Two sided decision:
  Alice chooses Tomás over Carlos
  \[ (T \succ_A C) \text{ or } (C \text{ prefers its match to } A). \]
Revealed Preference Theory

Why is it hard?

- Important problem: rationalizing preferences can explain revealed preference and “available sets” (budgets).
Why is it hard?

- Important problem: rationalizing preferences can explain revealed preference and “available sets” (budgets).
- Hence direction of revealed preference is affected by the hypothesized rationalizing preferences.
- Literature mostly deals with the problem by assuming transferable utility.
Reconcile:

- Theory of stable *individual* matchings.
- Data on *aggregate* matchings.
What we do.

\[
\begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{pmatrix}
\]

Echenique – Lee – Shum

Aggregate Matchings
What we do.

\[
\begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{pmatrix}
\quad \begin{pmatrix}
1 & 8 & 0 & 0 \\
0 & 4 & 3 & 0 \\
7 & 3 & 0 & 0 \\
0 & 0 & 9 & 5
\end{pmatrix}
\]
## Marriage Data (Michigan)

<table>
<thead>
<tr>
<th>Age</th>
<th>12-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-50</th>
<th>51-94</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-20</td>
<td>231</td>
<td>47</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>21-25</td>
<td>329</td>
<td>798</td>
<td>156</td>
<td>32</td>
<td>11</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>26-30</td>
<td>71</td>
<td>477</td>
<td>443</td>
<td>136</td>
<td>27</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>31-35</td>
<td>11</td>
<td>148</td>
<td>249</td>
<td>196</td>
<td>83</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>36-40</td>
<td>2</td>
<td>41</td>
<td>105</td>
<td>144</td>
<td>114</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td>41-50</td>
<td>0</td>
<td>15</td>
<td>42</td>
<td>118</td>
<td>121</td>
<td>162</td>
<td>25</td>
</tr>
<tr>
<td>51-94</td>
<td>0</td>
<td>2</td>
<td>11</td>
<td>11</td>
<td>35</td>
<td>137</td>
<td>158</td>
</tr>
</tbody>
</table>
Question:

- Given an “aggregate matching table” (data), when are there preferences for individuals s.t. the matching is stable?
- In other words, what are the testable implications of stability for aggregate matchings.
Main results

Revealed preference exercise:
Main results

Revealed preference exercise:

▶ Characterization of rationalizable agg. match.
▶ Characterization under TU: strictly more restrictive.

Ex:
Main results

Revealed preference exercise:

▶ Characterization of rationalizable agg. match.
▶ Characterization under TU: strictly more restrictive.

Ex:

```
5  3  1
0  7  8
9  4  0
```
Main results

Revealed preference exercise:

- Characterization of rationalizable agg. match.
- Characterization under TU: strictly more restrictive.

Ex:

```
<table>
<thead>
<tr>
<th>5</th>
<th>3</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
```
Main results

Econometric estimation strategy:
   ▶ Moment inequalities
   ▶ Set identification parameters in “index” utility model.
   ▶ Empirical illustration to US marriage data.

Other results...
Average marriages across 51 states

<table>
<thead>
<tr>
<th>State</th>
<th>Avg Marriages</th>
<th>Std Dev</th>
<th>5%ile</th>
<th>25%ile</th>
<th>75%ile</th>
<th>95%ile</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>858</td>
<td>189</td>
<td>30</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1142</td>
<td>2261</td>
<td>495</td>
<td>121</td>
<td>35</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>253</td>
<td>1436</td>
<td>1349</td>
<td>388</td>
<td>111</td>
<td>36</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>56</td>
<td>401</td>
<td>762</td>
<td>560</td>
<td>203</td>
<td>76</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>120</td>
<td>303</td>
<td>378</td>
<td>290</td>
<td>142</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>155</td>
<td>250</td>
<td>325</td>
<td>431</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>23</td>
<td>45</td>
<td>86</td>
<td>296</td>
<td>461</td>
<td>461</td>
</tr>
</tbody>
</table>
Average marriages across 51 states

<table>
<thead>
<tr>
<th>858</th>
<th>189</th>
<th>30</th>
<th>9</th>
<th>3</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1142</td>
<td>2261</td>
<td>495</td>
<td>121</td>
<td>35</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>253</td>
<td>1436</td>
<td>1349</td>
<td>388</td>
<td>111</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>56</td>
<td>401</td>
<td>762</td>
<td>560</td>
<td>203</td>
<td>76</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>120</td>
<td>303</td>
<td>378</td>
<td>290</td>
<td>142</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>155</td>
<td>250</td>
<td>325</td>
<td>431</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>23</td>
<td>45</td>
<td>86</td>
<td>296</td>
<td>461</td>
</tr>
</tbody>
</table>
Average marriages across 51 states

<table>
<thead>
<tr>
<th></th>
<th>858</th>
<th>1142</th>
<th>253</th>
<th>56</th>
<th>16</th>
<th>8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>189</td>
<td></td>
<td></td>
<td></td>
<td>401</td>
<td>120</td>
<td>54</td>
<td>10</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>2261</td>
<td>1436</td>
<td>762</td>
<td>303</td>
<td>155</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>495</td>
<td>1349</td>
<td>560</td>
<td>378</td>
<td>250</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>121</td>
<td>388</td>
<td>762</td>
<td>290</td>
<td>325</td>
<td>86</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>35</td>
<td>111</td>
<td>203</td>
<td>142</td>
<td>431</td>
<td>296</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>12</td>
<td>36</td>
<td>76</td>
<td>142</td>
<td>53</td>
<td>461</td>
</tr>
</tbody>
</table>

Echenique – Lee – Shum
Aggregate Matchings
Average marriages across 51 states

<table>
<thead>
<tr>
<th>Average marriages across 51 states</th>
</tr>
</thead>
<tbody>
<tr>
<td>858 189 30 9 3 1 0</td>
</tr>
<tr>
<td>1142 2261 495 121 35 12 1</td>
</tr>
<tr>
<td>253 1436 1349 388 111 36 2</td>
</tr>
<tr>
<td>56 401 762 560 203 76 4</td>
</tr>
<tr>
<td>16 120 303 378 290 142 9</td>
</tr>
<tr>
<td>8 54 155 250 325 431 53</td>
</tr>
<tr>
<td>2 10 23 45 86 296 461</td>
</tr>
</tbody>
</table>
## Average Marriages Across 51 States

<table>
<thead>
<tr>
<th></th>
<th>858</th>
<th>1142</th>
<th>253</th>
<th>56</th>
<th>16</th>
<th>8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>189</td>
<td>2261</td>
<td>1436</td>
<td>401</td>
<td>495</td>
<td>1349</td>
<td>762</td>
<td>401</td>
</tr>
<tr>
<td>30</td>
<td>495</td>
<td>1349</td>
<td>762</td>
<td>560</td>
<td>388</td>
<td>303</td>
<td>303</td>
</tr>
<tr>
<td>9</td>
<td>121</td>
<td>388</td>
<td>378</td>
<td>290</td>
<td>378</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>111</td>
<td>290</td>
<td>560</td>
<td>203</td>
<td>325</td>
<td>325</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>36</td>
<td>142</td>
<td>76</td>
<td>76</td>
<td>461</td>
<td>461</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Echenique – Lee – Shum*  
Aggregate Matchings
### Average marriages across 51 states

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>858</td>
<td>189</td>
<td>30</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1142</td>
<td>2261</td>
<td>495</td>
<td>121</td>
<td>35</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>253</td>
<td>1436</td>
<td>1349</td>
<td>388</td>
<td>111</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>56</td>
<td>401</td>
<td>762</td>
<td>560</td>
<td>203</td>
<td>76</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>120</td>
<td>303</td>
<td>378</td>
<td>290</td>
<td>142</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>155</td>
<td>250</td>
<td>325</td>
<td>431</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>23</td>
<td>45</td>
<td>86</td>
<td>296</td>
<td>461</td>
</tr>
</tbody>
</table>
An *aggregate matching market* is described by a triple $\langle M, W, >\rangle$, where:

- $M$ and $W$ are disjoint, finite sets. We call the elements of $M$ *types of men* and the elements of $W$ *types of women*.
- $> = ((>_m)_{m \in M}, (>_w)_{w \in W})$ is a profile of strict preferences: for each $m$ and $w$, $>_m$ is a linear order over $W \cup \{m\}$ and $>_w$ is a linear order over $M \cup \{w\}$.
An *aggregate matching market* is described by a triple \( \langle M, W, > \rangle \), where:

- \( M \) and \( W \) are disjoint, finite sets. We call the elements of \( M \) *types of men* and the elements of \( W \) *types of women*.
- \( > = ((>_m)_{m \in M}, (>_w)_{w \in W}) \) is a profile of strict preferences: for each \( m \) and \( w \), \( >_m \) is a linear order over \( W \cup \{m\} \) and \( >_w \) is a linear order over \( M \cup \{w\} \).

**Note:** identical preferences within type.

We show that relaxing this assumption in our framework leads to a vacuous theory.
An aggregate matching is a $K \times L$ matrix $X = (X_{ij})$ with $X_{ij} \in \mathbb{N}$. 
An aggregate matching is a $K \times L$ matrix $X = (X_{ij})$ with $X_{ij} \in \mathbb{N}$.

An aggregate matching $X$ is canonical if $X_{ij} \in \{0, 1\}$.
An aggregate matching is a $K \times L$ matrix $X = (X_{ij})$ with $X_{ij} \in \mathbb{N}$.

An aggregate matching $X$ is canonical if $X_{ij} \in \{0, 1\}$.

A canonical matching $X$ is a simple matching if for each $i$ there is at most one $j$ with $X_{ij} = 1$, and for each $j$ there is at most one $i$ with $X_{ij} = 1$. 
Model

- $X$ is *individually rational* if

\[ X_{ij} > 0 \Rightarrow w_j >_{m_i} m_i \text{ and } m_i >_{w_j} w_j. \]
X is *individually rational* if

\[ X_{ij} > 0 \Rightarrow w_j > \_m_i \_m_i \text{ and } m_i > \_w_j \_w_j. \]

\((m_i, w_j)\) is a *blocking pair* if \( \exists \)

- \( w_k \in \mathcal{W} \text{ with } X_{ik} > 0, \text{ and } m_l \in \mathcal{M} \text{ with } X_{jl} > 0, \)
- s.t. \( w_j > \_m_i \_w_k \text{ and } m_i > \_w_j \_m_l. \)
X is *individually rational* if

\[ X_{ij} > 0 \Rightarrow w_j > m_i \text{ and } m_i > w_j. \]

\((m_i, w_j)\) is a *blocking pair* if \( \exists \)

\[
\begin{align*}
    & w_k \in W \text{ with } X_{ik} > 0, \text{ and } m_l \in M \text{ with } X_{jl} > 0, \\
    & \text{s.t. } w_j > m_i \text{ and } m_i > w_j. \\
\end{align*}
\]

X is *stable* if it is individually rational and there are no blocking pairs for X.
Given $X$, construct a canonical aggregate matching $X^c$ by:

- $X^c_{ij} = 0$ when $X_{ij} = 0$ and
- $X^c_{ij} = 1$ when $X_{ij} > 0$.

**Observation**

An aggregate matching $X$ is stable if and only if $X^c$ is stable.
Example: simple vs. aggregate matching

Let \( \langle M, W, \rangle \) with \( M = \{ m_1, m_2, m_3 \} \), \( W = \{ w_1, w_2, w_3 \} \), and

<table>
<thead>
<tr>
<th></th>
<th>( m_1 )</th>
<th>( m_2 )</th>
<th>( m_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( w_3 )</td>
<td></td>
</tr>
<tr>
<td>( w_2 )</td>
<td>( w_3 )</td>
<td>( w_1 )</td>
<td></td>
</tr>
<tr>
<td>( w_3 )</td>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( w_1 )</th>
<th>( w_2 )</th>
<th>( w_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_2 )</td>
<td>( m_3 )</td>
<td>( m_1 )</td>
<td></td>
</tr>
<tr>
<td>( m_3 )</td>
<td>( m_1 )</td>
<td>( m_2 )</td>
<td></td>
</tr>
<tr>
<td>( m_1 )</td>
<td>( m_2 )</td>
<td>( m_3 )</td>
<td></td>
</tr>
</tbody>
</table>
The following simple matchings are stable:

\[
X^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}
\]

Sum of \(X^1\) and \(X^2\):

\[
\hat{X} = X^1 + X^2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}
\]

\((m_1, w_2)\) is a blocking pair.
\(\langle M, W, >\rangle\) defines a graph \((V, E)\) where

- \(V\) is the set of pairs \((i, j)\)
- \(((i, j), (k, l)) \in E\) if
  - \(w_l >_m w_j\) and \(m_i > w_l m_k\) or
  - \(w_j >_m w_l\) and \(m_k > w_j m_i\).

\(X\) is stable iff

\[
(((i, j), (k, l)) \in E \Rightarrow X_{ij}X_{kl} = 0. \tag{1}
\]

Otherwise (ie. \(X_{ij} = X_{kl} = 1\)), either \((i, j)\) or \((k, j)\) is blocking pair.
Stability – Example

3 men and women:

\[
\begin{array}{ccc|ccc}
> m_1 & > m_2 & > m_3 & > w_1 & > w_2 & > w_3 \\
\hline
w_1 & w_2 & w_3 & m_2 & m_3 & m_1 \\
w_2 & w_3 & w_1 & m_3 & m_1 & m_2 \\
w_3 & w_1 & w_2 & m_1 & m_2 & m_3 \\
\end{array}
\]

Graph:
### Stability – Example

3 men and women:

<table>
<thead>
<tr>
<th>( \succ m_1 )</th>
<th>( \succ m_2 )</th>
<th>( \succ m_3 )</th>
<th>( \succ w_1 )</th>
<th>( \succ w_2 )</th>
<th>( \succ w_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( w_3 )</td>
<td>( m_2 )</td>
<td>( m_3 )</td>
<td>( m_1 )</td>
</tr>
<tr>
<td>( w_2 )</td>
<td>( w_3 )</td>
<td>( w_1 )</td>
<td>( m_3 )</td>
<td>( m_1 )</td>
<td>( m_2 )</td>
</tr>
<tr>
<td>( w_3 )</td>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( m_1 )</td>
<td>( m_2 )</td>
<td>( m_3 )</td>
</tr>
</tbody>
</table>

**Stable matching:**

```
\[
\begin{array}{ccc}
1 & 1 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
\end{array}
\]
```
An **antiedge** is a pair \((i, j), (k, l)\) with \(i \neq k \in M; \ j \neq l \in W\) s.t. \(X_{ij} = X_{kl} = 1\).

Then \(X\) is stable iff

\[(ij), (kl) \text{ is anti-edge} \Rightarrow \begin{cases} d_{ilj}d_{lik} = 0 \\ d_{jki}d_{kjl} = 0 \end{cases} \quad (2)\]

Define: \(d_{ilj} = 1 (w_l >_m w_i, w_j)\)
Structure of Aggregate Stable Matchings

$X$ dominates $X'$ if

$$X_{ij} = 0 \Rightarrow X'_{ij} = 0.$$  

**Proposition**

Let $X$ be a stable aggregate matching. If $X'$ is an aggregate matching, and $X$ dominates $X'$, then $X'$ is stable.

So all stable matchings are described by set of maximal stable matchings.
(Trivial) Algorithm for maximal stable matching.

Given \((V, E)\)

- Enumerate vertices, \(V = \{1, 2, \ldots N\}\).
- \(X^0 = \) identically zero.
- For \(v \in V\), \(X^{v-1}\), define \(X^v\) by changing entry \(v\).
  - \(X^v_v = 1\) if 1 is not violated
  - \(X^v_v = 0\) o/w.
- Let \(X = X^N\).
Proposition

Let $X$ be an individual stable matching.

1. If $K = L = 3$ then $X$ is not a maximal stable matching.

2. If $K > 3$, $L > 3$ and $X$ is a maximal stable matching, then one of the following two possibilities must hold:

   2.1 For all $(i,j)$, the submatching $X^{-(i,j)}$ is a maximal stable matching in the $-(i,j)$ submarket.

   2.2 There is $(h,l)$ with $X_{hl} = 1$, and a maximal stable matching $\tilde{x}$, for which $\tilde{x}_{h,j} = \tilde{x}_{i,l} = 0$ for all $i$ and $j$. 
Given: $M = \{m_1, \ldots, m_K\}$ and $W = \{w_1, \ldots, w_L\}$.

$X$ is *rationalizable* if $\exists$ preference profile $>$ s.t. $X$ is a stable aggregate matching in $\langle M, W, > \rangle$. 
Given $X$:
Define a “lattice graph” $(V, L)$ on the matrix $X$.

- Vertices: $(i, j)$ s.t. $X_{i,j} = 1$
- Edge $(i, j) - (i', j')$ if share a column or a row.
Example

Let $X$ be

\[
\begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
1 & 1 & 0
\end{pmatrix}.
\]

$(V, L)$ is:
**Theorem**

An aggregate matching $X$ is rationalizable if and only if the associated graph $(V, L)$ has not two connected distinct minimal cycles.
Let $X$ be

\[
\begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
1 & 1 & 0
\end{pmatrix}
\]
The following are two minimal cycles that are connected.

1 1 1 1

0 1 1 1

1 1 0

1 1 0
<table>
<thead>
<tr>
<th>Age</th>
<th>12-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-50</th>
<th>51-94</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-20</td>
<td>231</td>
<td>47</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>21-25</td>
<td>329</td>
<td>798</td>
<td>156</td>
<td>32</td>
<td>11</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>26-30</td>
<td>71</td>
<td>477</td>
<td>443</td>
<td>136</td>
<td>27</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>31-35</td>
<td>11</td>
<td>148</td>
<td>249</td>
<td>196</td>
<td>83</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>36-40</td>
<td>2</td>
<td>41</td>
<td>105</td>
<td>144</td>
<td>114</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td>41-50</td>
<td>0</td>
<td>15</td>
<td>42</td>
<td>118</td>
<td>121</td>
<td>162</td>
<td>25</td>
</tr>
<tr>
<td>51-94</td>
<td>0</td>
<td>2</td>
<td>11</td>
<td>11</td>
<td>35</td>
<td>137</td>
<td>158</td>
</tr>
</tbody>
</table>
Idea: necessity.

Canonical cycle:
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:

```
1 ----> 1
|     |
|     |
1 ----> 1
```
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:

```
1 --\rightarrow 1
|     |
|     |
1     1
```
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:

\[
\begin{array}{c}
1 \rightarrow 1 \\
\downarrow \\
1 \rightarrow 1
\end{array}
\]
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:

```
1 -> 1
↓   ↓
1 1
```

Echenique – Lee – Shum
Aggregate Matchings
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:

```
  1  --->  1  
   |      |    
 1  |      |  1
   |      |    
 1  <----|---->  1
```
Idea: necessity.

Preferences ⇒ orientation of edges:

\[
\begin{align*}
1 & \rightarrow 1 \\
\uparrow & \quad \downarrow \\
1 & \leftarrow 1 \\
\end{align*}
\]
So a cycle must be oriented as a flow.
Idea: necessity
Idea: necessity

\[
\begin{array}{c}
1 & \rightarrow & 1 \\
\uparrow & & \downarrow \\
1 & \leftarrow & 1 & \rightarrow & 1 \\
\downarrow & & \uparrow \\
1 & \leftarrow & 1 & \rightarrow & 1 \\
\end{array}
\]
Idea: necessity

\[
\begin{align*}
1 & \rightarrow 1 \\
\uparrow & \downarrow \\
1 & \leftarrow 1 \\
\downarrow & \uparrow \\
1 & \rightarrow 1 \\
\end{align*}
\]
Orientation of a minimal path must then point \textit{away} from a cycle.
Idea: necessity

- Orientation of a minimal path must then point \textit{away} from a cycle.
- Two connected cycles $\Rightarrow$ connecting path must point away from both.
Idea: necessity

Subsequent edges in a minimal path must be at a right angle:
Idea: necessity

Two connected cycles ⇒ connecting path must point away from both.
So connected path does (at some point):

1 \rightarrow 1 \uparrow 

1

⇒ no two connected cycles.
Idea: sufficiency

- Given $X$, construct an orientation of $(V, L)$.
- Use orientation to define preferences.
Idea: sufficiency

- Given $X$, construct an orientation of $(V, L)$.
- Use orientation to define preferences.
- Decompose $(V, L)$ in connected components. At most one cycle in each.
Idea: sufficiency

- Given $X$, construct an orientation of $(V, L)$.
- Use orientation to define preferences.
- Decompose $(V, L)$ in connected components. At most one cycle in each.
- Orient cycle as a “flow,” and paths as “flows” pointing away from cycle.
- Uniqueness of cycle within a component ensures transitivity.
Surplus: $\alpha_{i,j} \in \mathbb{R}$.

Surplus generated by matchings of types $i$ and $j$ in $X$ is $X_{i,j}\alpha_{i,j}$. 
X is *TU-rationalizable* by a matrix of surplus \( \alpha \) if \( X \) is unique sol. to:

\[
\begin{align*}
\max_{\tilde{X}} & \quad \sum_{i,j} \alpha_{i,j} \tilde{X}_{i,j} \\
\text{s.t.} & \quad \forall j \sum_i \tilde{X}_{i,j} = \sum_i X_{i,j} \\
& \quad \forall i \sum_j \tilde{X}_{i,j} = \sum_j X_{i,j}
\end{align*}
\]  

(3)
Theorem
An aggregate matching $X$ is TU-rationalizable if and only if the associated graph $(V, L)$ contains no minimal cycles.

Corollary
If an aggregate matching $X$ is TU-rationalizable, then it is rationalizable.
Estimation

Parametrized preferences:

\[ u_{ij} = Z_{ij} \beta + \varepsilon_{ij}, \quad (4) \]

\[ d_{ijk} \equiv 1(u_{ij} \geq u_{ik}). \]
Recall:

An **antiedge** is a pair \((i, j), (k, l)\) with \(i \neq k \in M; \ j \neq l \in W\) s.t. \(X_{ij} = X_{kl} = 1\).

Then \(X\) is stable iff

\[(ij), (kl) \text{ is anti-edge } \Rightarrow \begin{cases} d_{ilj}d_{lik} = 0 \\ d_{jki}d_{kjl} = 0 \end{cases} \] (5)
\[ Pr((ij), (kl) \text{ antiedge}) \leq (1 - Pr(d_{ilj}d_{lik} = 1))(1 - Pr(d_{jkl}d_{kjl} = 1)) = Pr(d_{ilj}d_{lik} = 0, d_{jkl}d_{kjl} = 0). \]
\[ Pr((ij), (kl) \text{ antiedge}) \leq (1 - Pr(d_{ij}d_{ik} = 1))(1 - Pr(d_{jki}d_{kjl} = 1)) = Pr(d_{ij}d_{ik} = 0, d_{jki}d_{kjl} = 0). \]

Gives a moment inequality:
\[ \mathbb{E} \left[ \mathbb{I}((ij), (kl) \text{ antiedge}) - Pr(d_{ij}d_{ik} = 0, d_{jki}d_{kjl} = 0; \beta) \right] \leq 0. \]

\[ g_{ijkl}(X_t; \beta) \]

The identified set is defined as
\[ B_0 = \{ \beta : \mathbb{E}g_{ijkl}(X_t; \beta) \leq 0, \forall i, j, k, l \}. \]
Sample analog

\[ \frac{1}{T} \sum_t 1((ij), (kl) \text{ is antiedge in } X_t) - 1 \]

\[ + \Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta) \]

\[ = \frac{1}{T} \sum_t g_{ijkl}(X_t; \beta). \]
Problem: condition in the theorem is violated. Hence no preferences (no betas) rationalize data.
Problem: condition in the theorem is violated. Hence no preferences (no betas) rationalize data.

We relax the model (\(\exists\) other solutions).
Estimation – Relaxation of the model

A blocking pair may not form.

$$\delta_{ijkl} = P(\text{types (}i, j\text{), (}k, l\text{) communicate}).$$

Idea: a BP forms only when types (i, j), (k, l) communicate. Then stability condition becomes:

$$\left( \begin{array}{c} (ij), (kl) \text{ is anti-edge} \\
(ij), (kl) \text{ meet} \end{array} \right) \Rightarrow \begin{cases} d_{ij}d_{ik} = 0 \\
d_{ji}d_{jk} = 0 \end{cases}$$

Modified moment inequality:

$$Pr((ij), (kl) \text{ antiedge}) \ast \delta_{ijkl} \leq Pr(d_{ij}d_{ik} = 0, d_{ji}d_{jk} = 0; \beta)$$

Assume: two events are independent.
Estimation – Relaxation of the model

We put some structure on “communication probabilities”. We allow $\delta_{ijkl}$ to vary across antiedges $(ij), (kl)$, depending on the number of $(ij)$ and $(kl)$ couples:

$$
\delta_{ijkl} = \min \left\{ 2 \cdot \frac{|X_{T_i}^{M}, T_j^{W}|}{|X|} \cdot \frac{|X_{T_k}^{M}, T_l^{W}|}{|X|}, 1 \right\}.
$$

where $\gamma > 0$ is a tuning parameter (higher is more restrictive).

$\delta^t_{ijkl}$ set to the number of potential blocking pairs which can form between $(ij)$ and $(kl)$ couples, as a proportion of total number of potential couples in the population $|X|^2$.

Essentially: we weigh/smooth anti-edges by $\#$ agents involved.

here
Specification of Utilities

-Men: 

\[ \text{Utility}^{m,w} = \beta_1 |\text{Age}^m - \text{Age}^w|^ - + \beta_2 |\text{Age}^m - \text{Age}^w|^ + + \varepsilon^{m,w} \]

-Women: 

\[ \text{Utility}^{w,m} = \beta_3 |\text{Age}^m - \text{Age}^w|^ - + \beta_4 |\text{Age}^m - \text{Age}^w|^ + + \varepsilon^{w,m} \]

Interpretation of preference parameters:

- \( \beta_1 (\beta_3) > 0 \): when wife older, men (women) prefer larger age gap; men prefer older women, women prefer younger men

- \( \beta_2 (\beta_4) > 0 \): when husband older, men (women) prefer larger age gap; men prefer younger women, women prefer older men
We describe the identified set for different values of $\gamma$.

if $\gamma$ is too high $\Rightarrow$ identified set $= \emptyset$.

if $\gamma$ is too low $\Rightarrow$ identified set is everything.

Idea: choose high $\gamma$ to “discipline” our estimates:

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\beta_1$ min</th>
<th>$\beta_1$ max</th>
<th>$\beta_2$ min</th>
<th>$\beta_2$ max</th>
<th>$\beta_3$ min</th>
<th>$\beta_3$ max</th>
<th>$\beta_4$ min</th>
<th>$\beta_4$ max</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>-2.00</td>
<td>2.00</td>
<td>-2.00</td>
<td>2.00</td>
<td>-2.00</td>
<td>2.00</td>
<td>-2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>28</td>
<td>-2.00</td>
<td>1.60</td>
<td>-2.00</td>
<td>2.00</td>
<td>-2.00</td>
<td>1.60</td>
<td>-2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>29</td>
<td>-2.00</td>
<td>0.40</td>
<td>-2.00</td>
<td>1.80</td>
<td>-2.00</td>
<td>0.40</td>
<td>-2.00</td>
<td>1.80</td>
</tr>
<tr>
<td>30</td>
<td>-2.00</td>
<td>-0.80</td>
<td>-2.00</td>
<td>0.60</td>
<td>-2.00</td>
<td>-0.85</td>
<td>-2.00</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Some guidance for interpreting identified sets

- More anti-edges below the diagonal, where $\text{age}^m > \text{age}^w$. So focus on $\beta_2, \beta_4$ (lower triangular preferences)
- More “downward-sloping” anti-edges than “upward-sloping”

- Downward-sloping anti-edge:
  $$(i, j) \rightarrow (i, l)$$
  $$(k, j) \rightarrow (k, l)$$

- Upward-sloping anti-edge:
  $$(k, j) \rightarrow (k, l)$$
  $$(i, j) \rightarrow (i, l)$$

Thus, stability “implies” antipodal preferences:
Some guidance for interpreting identified sets

- More anti-edges below the diagonal, where $age^m > age^w$. So focus on $\beta_2, \beta_4$ (lower triangular preferences)
- More “downward-sloping” anti-edges than “upward-sloping”

Downward-sloping anti-edge:

\[
(i, j) \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow (i, l)
\]

Upward-sloping anti-edge:

\[
(k, j) \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow (k, l)
\]

Thus, stability “implies” antipodal preferences:
if women prefer older men, then men prefer older women
Some guidance for interpreting identified sets

- More anti-edges below the diagonal, where $age^m > age^w$. So focus on $\beta_2, \beta_4$ (lower triangular preferences)
- More “downward-sloping” anti-edges than “upward-sloping”

\[
\begin{align*}
\text{Downward-sloping anti-edge:} & \quad (i, j) \leftarrow (i, l) \\
\text{Upward-sloping anti-edge:} & \quad (k, j) \rightarrow (k, l)
\end{align*}
\]

Thus, stability “implies” antipodal preferences:
if women prefer older men, then men prefer older women
if women prefer younger men, then men prefer younger women
Joint identified sets

Some guidance for interpreting identified sets

- More anti-edges below the diagonal, where $age^m > age^w$. So focus on $\beta_2, \beta_4$ (lower triangular preferences)
- More “downward-sloping” anti-edges than “upward-sloping”

Downward-sloping anti-edge:

\[
(i, j) \leftarrow (i, l) \quad (k, j) \quad (k, l)
\]

Upward-sloping anti-edge:

\[
(k, j) \rightarrow (k, l) \quad (i, j) \quad (i, l)
\]

Thus, stability “implies” antipodal preferences:
- if women prefer older men, then men prefer older women
- if women prefer younger men, then men prefer younger women

Do we see this in identified sets? Consider slices of identified set
When women prefer older men ($\beta_4 = 1$)

$$(\beta_3, \beta_4) = -2, 1$$

$$(\beta_3, \beta_4) = 0, 1$$

$$(\beta_3, \beta_4) = 1, 1$$

Then $\beta_2 < 0$ (y-axis): men prefer older women
When women prefer younger men ($\beta_4 = -2$)

$$(\beta_3, \beta_4) = -2, -2$$  
$$(\beta_3, \beta_4) = 0, -2$$  
$$(\beta_3, \beta_4) = 1, -2$$

Then $\beta_2 > 0$ primarily (men prefer younger women)
When women indifferent \((\beta_4 = 0)\)

\[
\begin{align*}
(\beta_3, \beta_4) &= -2,0 \\
(\beta_3, \beta_4) &= 0,0 \\
(\beta_3, \beta_4) &= 1,0
\end{align*}
\]

Then \(\beta_2 < 0\): men prefer older women (ie. more equal-aged spouse, given that men are older). Men’s “true colors”? 
Confidence sets: $\gamma = 28$.

(a) $(\beta_3, \beta_4) = -2, 1$
(b) $(\beta_3, \beta_4) = 0, 1$
(c) $(\beta_3, \beta_4) = 1, 1$
(d) $(\beta_3, \beta_4) = -2, 0$
(e) $(\beta_3, \beta_4) = 0, 0$
(f) $(\beta_3, \beta_4) = 1, 0$

Echenique – Lee – Shum | Aggregate Matchings
For TU model, define the surplus obtained by matching of a type $i$ man with type $j$ woman as:

$$\alpha_{ij} = u_{ij} + u_{ji} = (\beta_1 + \beta_3)|\text{Age}_m - \text{Age}_w|^- + (\beta_2 + \beta_4)|\text{Age}_m - \text{Age}_w|^+$$

We work from the pairwise stability condition: for every anti-edge $(ij), (kl)$, we have

$$(ij), (kl) \text{ antiedge} \Rightarrow \alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}.$$

This leads to the moment inequality

$$Pr((ij), (kl) \text{ antiedge}) \leq Pr(\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}; \theta).$$
This condition derived via optimality. Let $W_{ij}$ be the number of $(ij)$ couples; an anti-edge is a pair $(ij), (kl)$ such that $W_{ij} > 0, W_{kl} > 0$. Consider an alternative matching $W'$ where a pair of $(ij), (kl)$ couples are swapped:

$$
W'_{ij} = W_{ij} - 1 \quad W'_{kl} = W_{kl} - 1
$$

$$
W'_{il} = W_{il} + 1 \quad W'_{kj} = W_{kj} + 1.
$$

By optimality of $W$, this swapping must lower surplus:

$$
\alpha_{ij} W_{ij} + \alpha_{il} W_{il} + \alpha_{kj} W_{kj} + \alpha_{kl} W_{kl}
\geq \alpha_{ij} W'_{ij} + \alpha_{il} W'_{il} + \alpha_{kj} W'_{kj} + \alpha_{kl} W'_{kl}
= \alpha_{ij} (W_{ij} - 1) + \alpha_{il} (W_{il} + 1) + \alpha_{kj} (W_{kj} + 1) + \alpha_{kl} (W_{kl} - 1)
\Rightarrow \alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}.
$$

For same reason, we introduce communication probability:

$$
\Pr((ij), (kl) \text{ antiedge}) \leq \Pr(\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}; \theta) / \delta_{(ij),(kl)}.
$$
Identified set: TU model.

Id set takes form $K_1 \leq \sum_{i=1}^{4} \beta_i \leq K_2$. 
Conclusions

- First theoretical characterization of stable aggregate matchings.
- Testable implications of stability for aggregate matchings. TU nested in NTU: could test more formally? (Model comparison with partially identified models)
- New estimation approach ← based on moment inequalities implied by stability. Multiplicity of stable matchings: partial identification of preference parameters
- Estimation using US marriage data: “antipodal” preferences Stripped down: add in additional covariates.
Sample moment inequalities

For antiedge \((ij), (kl)\):

\[
\frac{1}{T} \sum_t g_{ijkl}(X_t; \beta) = \frac{1}{T} \sum_t \left\{ \mathbb{1}(((ij),(kl) \text{ is antiedge in } X_t) \right\} \\
- \frac{1}{\delta_{ijkl}} \cdot (1 - \Pr(d_{ilj} = 1; \beta_{1,2}) \Pr(d_{lik} = 1; \beta_{3,4})) \\
(1 - \Pr(d_{jki} = 1; \beta_{3,4}) \Pr(d_{kjl} = 1; \beta_{1,2}))
\]

for all combinations of pairs, \((i,j)\) and \((k,l)\).
Why individual-level preference heterogeneity fails

- **Individual-level stability:**

\[ \sum_k x_{i,k}d_{ikj} + \sum_k x_{k,j}d_{jki} + x_{i,j} \geq 1 \]

where each \( x \in \{0, 1\} \).

- Ind-specific i.i.d. preference shock (Dagsvik, Choo-Siow):

\[ P(d_{ijk} = 1) = P(d_{i'j'k'} = 1) : \]

for all \((i, i') \in t_i^M\), \((j, j') \in t_j^M\), \((k, k') \in t_k^M\).

- Aggregating up to type-level (cf. Proof Claim 4.1) we get

\[ 2 \left| t_i^M \right| \left| t_j^W \right| \geq \left| t_i^M \right| \left| t_j^W \right| (1 - E[x_{ij}]) \Rightarrow 2 \geq (1 - E[x_{ij}]) \]

which is trivially satisfied.