Aggregate Matchings

Federico Echenique   SangMok Lee   Matthew Shum

California Institute of Technology

Roth-Sotomayor Celebration – May 7th, 2010

What we do:

Revealed preference exercise for matching theory.

Reconcile:

- Theory of stable individual matchings.
- Data on aggregate matchings.

What we do.

vs.

What we do.

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

What we do.

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 8 & 0 & 0 \\ 0 & 4 & 3 & 0 \\ 7 & 3 & 0 & 0 \\ 0 & 0 & 9 & 5 \end{pmatrix}$$

Marriage Data (Michigan)

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

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.

General motivation: two sided decision problems

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

General motivation: two sided decision problems

- Standard revealed preference:
  Alice buys tomatoes when carrots are available
  $\Rightarrow (T \succ_A C)$.
- Two sided decision:
  Alice chooses Tomás over Carlos
  $\Rightarrow (T \succ_A C)$ or (C prefers its match to A).

Broad motivation: two sided decision problems

- Important problem: rationalizing preferences can explain revealed preference and "available sets" (budgets).

Broad motivation: two sided decision problems

- 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.

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:

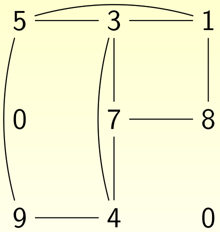
5	3	1
0	7	8
9	4	0

Main results

Revealed preference exercise:

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

Ex:



Main results

Econometric estimation strategy:

- Moment inequalities
- Set identification parameters in "index" utility model.
- Empirical illustration to US marriage data.

Other results

- Stability for aggregate match. is substantially different from individual match.
- Structure of stable aggregate matchings.

Model

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 $M \cup \{w\}$.

Model

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.

Model

- An *aggregate matching* is a $K \times L$ matrix $X = (X_{ij})$ with $X_{ij} \in \mathbf{N}$.

Model

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

Model

- An *aggregate matching* is a $K \times L$ matrix $X = (X_{ij})$ with $X_{ij} \in \mathbf{N}$.
- 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

$$X_{ij} > 0 \Rightarrow w_j >_{m_i} m_i \text{ and } m_i >_{w_j} w_j.$$

Model

- 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 $\exists$
  - $w_k \in W$ with $X_{ik} > 0$, and $m_l \in M$ with $X_{jl} > 0$,
  - s.t. $w_j >_{m_i} w_k$ and $m_i >_{w_j} m_l$.

Model

- 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 $\exists$
  - $w_k \in W$ with $X_{ik} > 0$, and $m_l \in M$ with $X_{jl} > 0$,
  - s.t. $w_j >_{m_i} w_k$ and $m_i >_{w_j} m_l$.
- X is *stable* if it is individually rational and there are no blocking pairs for X.

Model

Given X, construct a canonical aggregate matching X^c by:

- $X_{ij}^c = 0$ when $X_{ij} = 0$ and
- $X_{ij}^c = 1$ 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

m_1	m_2	m_3
w_1	w_2	w_3
w_2	w_3	w_1
w_3	w_1	w_2

w_1	w_2	w_3
m_2	m_3	m_1
m_3	m_1	m_2
m_1	m_2	m_3

Model

The following simple matchings are stable:

$$X^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \; X^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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.

Stability

$\langle M, W, > \rangle$ defines a graph (V, E) where

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

X is stable iff

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

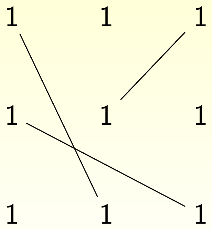
Otherwise (ie. $X_{ij} = X_{kl} = 1$), either (i, j) or (k, j) is blocking pair.

Stability – Example

3 men and women:

$>_{m_1}$	$>_{m_2}$	$>_{m_3}$	$>_{w_1}$	$>_{w_2}$	$>_{w_3}$
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

Graph:

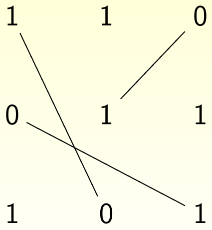


Stability – Example

3 men and women:

$>_{m_1}$	$>_{m_2}$	$>_{m_3}$	$>_{w_1}$	$>_{w_2}$	$>_{w_3}$
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

Stable matching:



Stability – contrapositive

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

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

Define: $d_{ilj} = \mathbb{1}(w_l >_{m_i} w_j)$

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 $X = X^N$.

Model

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.*

Rationalizable Matchings

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 $\langle M, W, > \rangle$.

Rationalizable Matchings

Given X:
Define a "lattice graph" (V, L) on the matrix X.

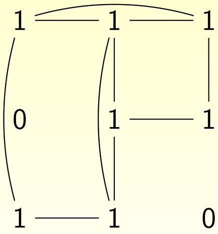
- Vertices: (i, j) s.t. $X_{i,j} = 1$
- Edge $(i, j) - (i', j')$ if share a column or a row.

Example

Let X be

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

(V, L) is:



Rationalizable Matchings

Theorem

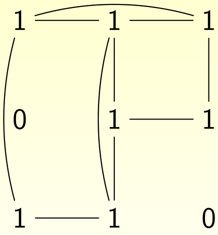
An aggregate matching X is rationalizable if and only if the associated graph (V, L) has not two connected distinct minimal cycles.

Rationalizable Matchings

Let X be

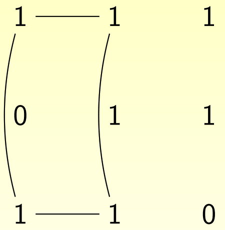
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

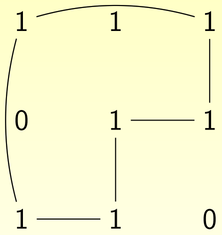
(V, L) is:



Rationalizable Matchings

The following are two minimal cycles that are connected.





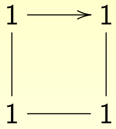
Idea: necessity.

Canonical cycle:

$$\begin{array}{ccc} 1 & \text{---} & 1 \\ | & & | \\ 1 & \text{---} & 1 \end{array}$$

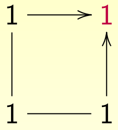
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:



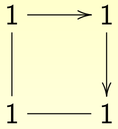
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:



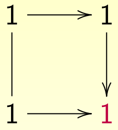
Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:



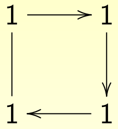
Idea: necessity.

Preferences ⇒ orientation of edges:



Idea: necessity.

Preferences $\Rightarrow$ orientation of edges:



Idea: necessity.

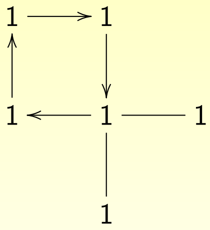
Preferences $\Rightarrow$ orientation of edges:

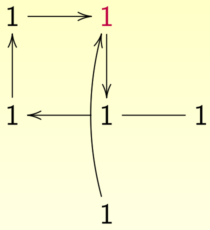
$$\begin{array}{ccc} 1 & \longrightarrow & 1 \\ \uparrow & & \downarrow \\ 1 & \longleftarrow & 1 \end{array}$$

Idea: necessity

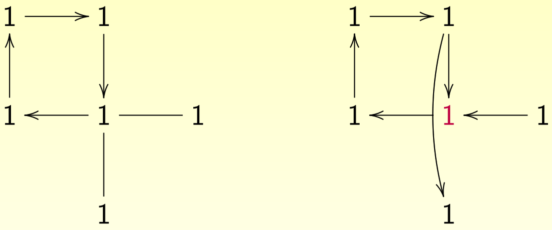
So a cycle must be oriented as a flow.

Idea: necessity

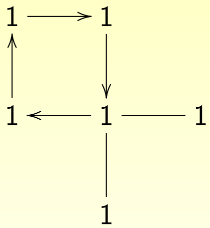


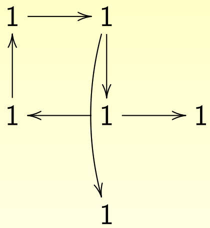


Idea: necessity



Idea: necessity





Idea: necessity

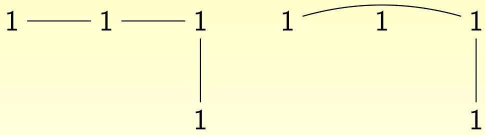
- Orientation of a minimal path must then point away from a cycle.

Idea: necessity

- Orientation of a minimal path must then point away from a cycle.
- Two connected cycles $\Rightarrow$ connecting path must point away from both.

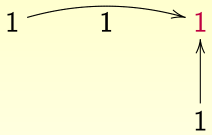
Idea: necessity

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):



⇒ 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.

TU rationalization

Surplus: $\alpha_{i,j} \in \mathbb{R}$.

Surplus generated by matchings of types i and j in X is $X_{i,j}\alpha_{i,j}$.

TU rationalization

X is *TU-rationalizable* by a matrix of surplus α if X is unique sol. to:

$$\max_{\tilde{X}} \sum_{i,j} \alpha_{i,j} \tilde{X}_{i,j}$$
$$s.t. \begin{cases} \forall j \ \sum_i \tilde{X}_{i,j} = \sum_i X_{i,j} \\ \forall i \ \sum_j \tilde{X}_{i,j} = \sum_j X_{i,j} \end{cases} \tag{3}$$

TU rationalization

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:

$$u_{ij} = Z_{ij}\beta + \varepsilon_{ij}, \tag{4}$$

$$d_{ijk} \equiv \mathbf{1}(u_{ij} \geq u_{ik}).$$

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

$$(ij),(kl) \text{ is anti-edge } \Rightarrow \begin{cases} d_{ilj}d_{lik} = 0 \\ d_{jki}d_{kjl} = 0 \end{cases} \tag{5}$$

Estimation

$$
\begin{aligned}
Pr((ij),(kl)\text{ antiedge}) &\leq (1 - Pr(d_{ilj}d_{lik} = 1))(1 - Pr(d_{jki}d_{kjl} = 1)) \\
&= Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0).
\end{aligned}
$$

Estimation

$$\begin{aligned} Pr((ij),(kl)\text{ antiedge}) &\leq (1 - Pr(d_{ilj}d_{lik} = 1))(1 - Pr(d_{jki}d_{kjl} = 1)) \\ &= Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0). \end{aligned}$$

Gives a moment inequality:

$$\mathbb{E}\underbrace{\left[\mathbb{1}((ij),(kl)\text{ antiedge}) - Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta)\right]}_{g_{ijkl}(X_t;\beta)} \leq 0.$$

The identified set is defined as

$$\mathbb{B}_0 = \left\{\beta : \mathbb{E}g_{ijkl}(X_t;\beta) \leq 0,\ \forall i,j,k,l\right\}.$$

Estimation

Sample analog

$$\frac{1}{T}\sum_t \mathbb{1}((ij),(kl) \text{ is antiedge in } X_t) - 1 \\ + Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta) \\ = \frac{1}{T}\sum_t g_{ijkl}(X_t;\beta).$$

Estimation

Problem: condition in the theorem is violated.
Hence no preferences (no betas) rationalize data.

Estimation

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.

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

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

Estimation – Relaxation of the model

Stability inequalities become:

$$\left(\begin{array}{c} (ij),(kl) \text{ is anti-edge} \\ (ij),(kl) \text{ meet} \end{array} \right) \Rightarrow \left\{ \begin{array}{l} d_{ilj}d_{lik} = 0 \\ d_{jki}d_{kjl} = 0 \end{array} \right.$$

Assume: two events are independent.

Estimation – Relaxation of the model

Modified moment inequality:

$$Pr((ij),(kl)\text{ antiedge}) \leq \frac{Pr(d_{ilj}d_{lik}=0, d_{jki}d_{kjl}=0;\beta)}{\delta_{ijkl}}$$

Estimation – Relaxation of the model

We suppose δ_{ijkl} depends on the number of agents in each type.

$$\delta^t_{ijkl} = 1 - (1-p)^{x_{T_i^M, T_j^W} + x_{T_k^M, T_l^W}}$$

where p is the prob. that two agents meet.

Estimation – Relaxation of the model

As a result: we are weighting anti-edges by how many agents are involved.

Sample moment inequalities

$$\frac{1}{T}\sum_t g_{ijkl}(X_t;\beta) = \Big(\frac{1}{T}\sum_t \mathbb{1}((ij),(kl) \text{ is antiedge in } X_t) * \delta^t_{ijkl}\Big)$$
$$- (1 - Pr(d_{ilj} = 1; \beta_{1,2})Pr(d_{lik} = 1; \beta_{3,4}))$$
$$(1 - Pr(d_{jki} = 1; \beta_{3,4})Pr(d_{kjl} = 1; \beta_{1,2}))$$

for all combinations of pairs, (i, j) and (k, l).

Specification of Utilities

$$Utility^m = \beta_1|Age^m - Age^w|^- + \beta_2|Age^m - Age^w|^+ + \varepsilon^m$$
$$Utility^w = \beta_3|Age^m - Age^w|^- + \beta_4|Age^m - Age^w|^+ + \varepsilon^w$$

Results: Identified set.

We describe the identified set for different values of p.

if p is too high $\Rightarrow$ identified set $= \emptyset$.
if p is too low $\Rightarrow$ identified set is everything.

Idea: choose high p as discipline on our estimates.

Results: Identified set.

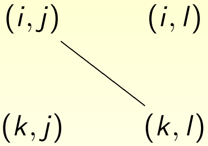
Table: Unconditional Bounds of β.

	β_1		β_2		β_3		β_4	
p	min	max	min	max	min	max	min	max
0.0006	-2	2	-2	2	-2	2	-2	2
0.0007	-2	2	-2	1.6	-2	2	-2	1.2
0.0008	-2	1.2	-1.6	1	-2	1	-1.4	0.6
0.0009	-1.2	0.4	-0.6	0.6	-2	0.4	-0.6	0.4

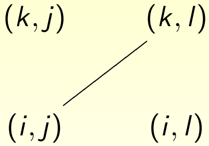
Joint identified sets

- More anti-edges below the diagonal, where $age^m > age^w$.
- More "downward-sloping" anti-edges than "upward-sloping" ones.

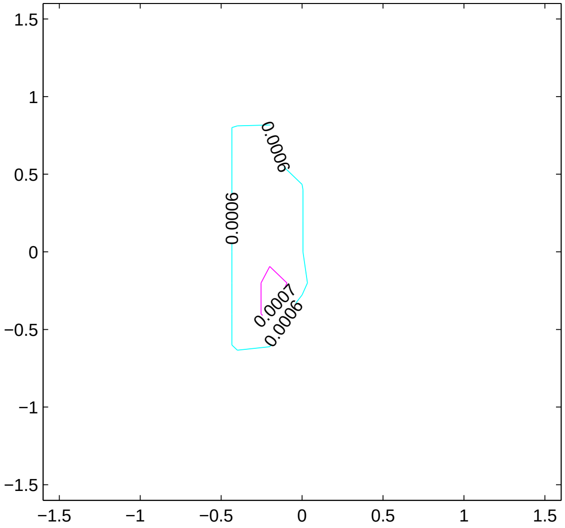
Downward-sloping anti-edge:



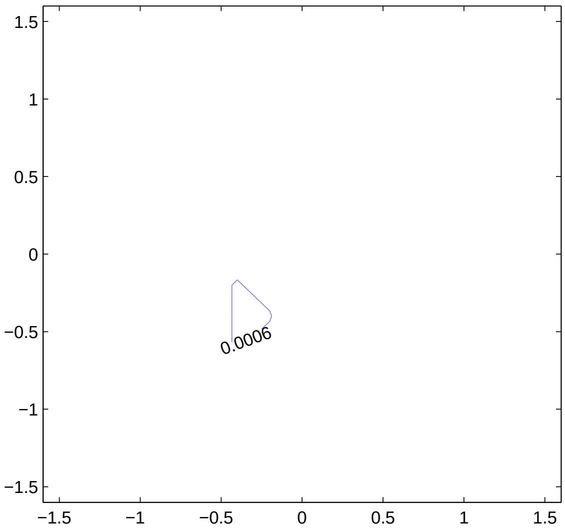
Upward-sloping anti-edge:



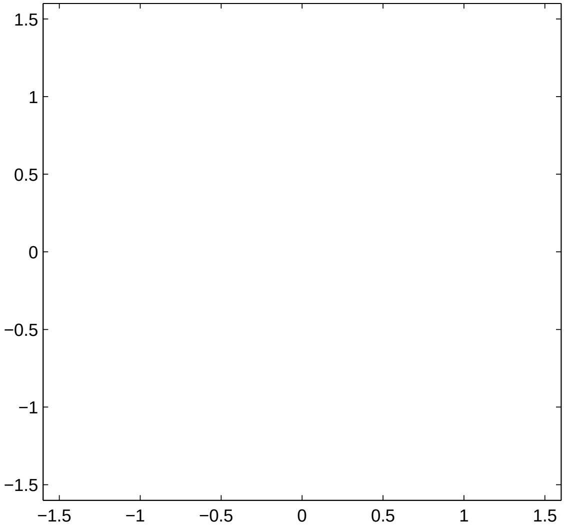
$\beta_3 = -1$ and $\beta_4 = 1$



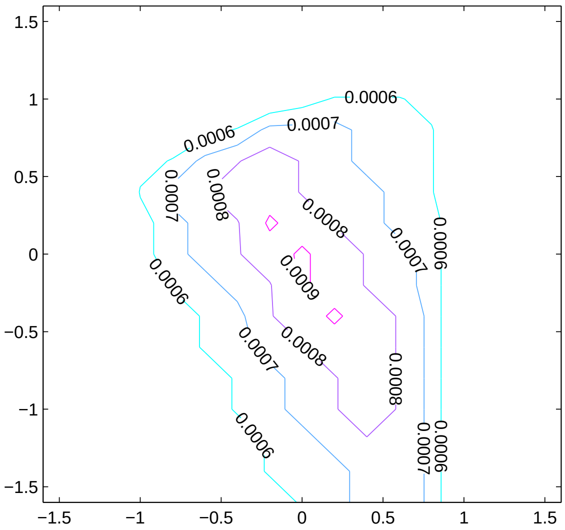
$\beta_3 = 0$ and $\beta_4 = 1$



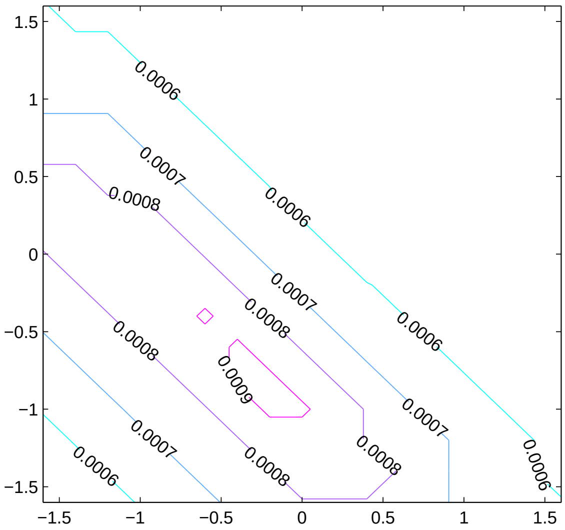
$\beta_3 = 1$ and $\beta_4 = 1$



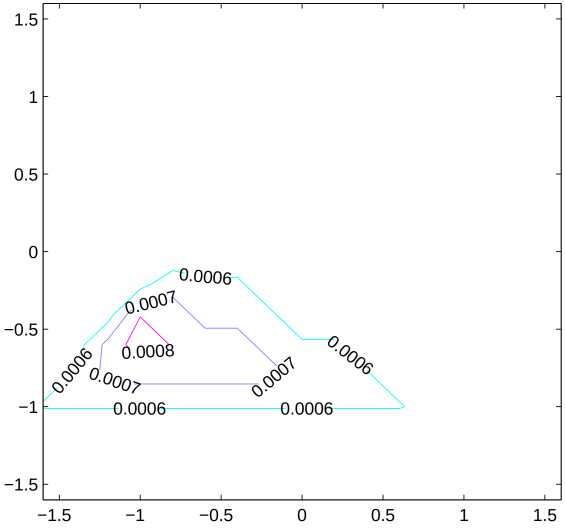
$\beta_3 = -1$ and $\beta_4 = 0$



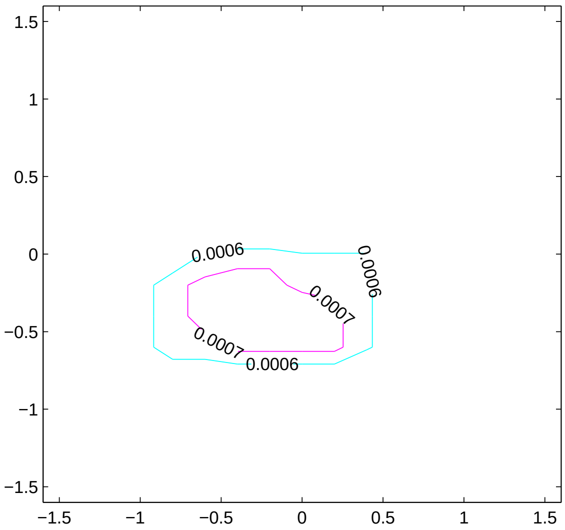
$\beta_3 = 0$ and $\beta_4 = 0$



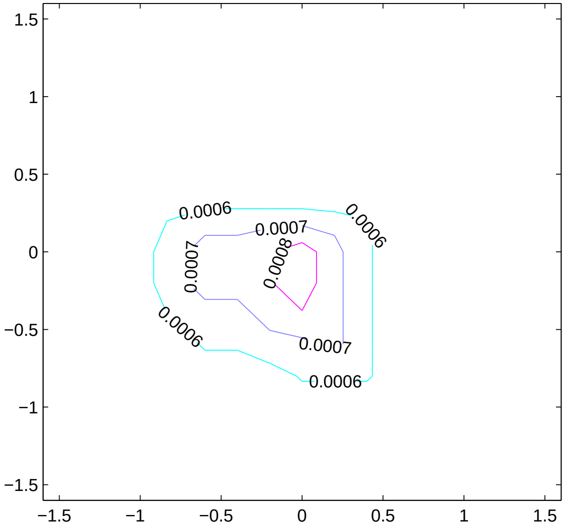
$\beta_3 = 1$ and $\beta_4 = 0$



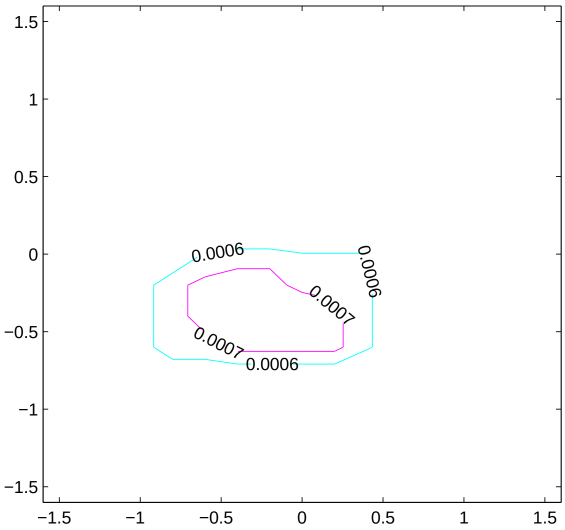
$\beta_3 = -1$ and $\beta_4 = -1$



$\beta_3 = 0$ and $\beta_4 = -1$



$\beta_3 = 1$ and $\beta_4 = -1$



Related literature

TU model:

- Theory: Shapley-Shubik (1971)
- Econometrics: Choo-Siow (2006), Fox (2007), Bajari-Fox (2008)

NTU model:

- Theory: Gale-Shapley (1967), Knuth (1971)
- Econometrics: Dagsvik (2000)

Aggregate NTU matching: no theoretical results; Dagsvik develops estimation techniques.

Conclusions

- First theoretical characterization of stable aggregate matchings.
- Testable implications of stability for aggregate matchings.
- Econometric estimation technique $\longleftarrow$ based on moment inequalities implied by stability.
- Illustration to US marriage data.