Matching: The Theory

Muriel Niederle
Stanford and NBER

September 26, 2011

Studying and doing Market Economics

In "Jonathan Strange and Mr. Norrel", Susanna Clarke describes an England around 1800, with magic societies, though not a lot of magic.
When asked why there is not more magic, the president of the York society of magicians replies that the question was wrong, "It presupposes that magicians have some sort of duty to do magic - which is clearly nonsense. [...] Magicians,.., study magic which was done long ago. Why would anyone expect more?"

Market design expects of economists more than just studying economics, rather the aim is to do economics.

Market Design

Design is both a verb and a noun, and we'll approach market design both as an activity and as an aspect of markets that we study.

Design also comes with a responsibility for detail. Designers can't be satisfied with simple models that explain the general principles underlying a market; they have to be able to make sure that all the detailed parts function together. Market design papers often have a detailed description of the market's unique and distinguishing features in their paper

Matching as part of Market Design

Responsibility for detail requires the ability to deal with complex institutional features that may be omitted from simple models.

Game theory, the part of economics that studies the "rules of the game," provides a framework with which design issues can be addressed.

But dealing with complexity will require new tools, to supplement the analytical toolbox of the traditional theorist (Computations, Experiments).

Game Theory, experimentation, and computation, together with careful observation of historical and contemporary markets (with particular attention to the market rules), are complementary tools of Design Economics.

Computation helps us find answers that are beyond our current theoretical knowledge.

Experiments play a role

- In diagnosing and understanding market failures, and successes
- In designing new markets
- In communicating results to policy makers

The Economist as Engineer

A rough analogy may help indicate how the parts of this course hang together. Consider the design of suspension bridges. Their simple physics, in which the only force is gravity, and all beams are perfectly rigid, is beautiful and indispensable.

But bridge design also concerns metal fatigue, soil mechanics, and the sideways forces of waves and wind. Many questions concerning these complications can't be answered analytically, but must be explored using physical or computational models.

These complications, and how they interact with that part of the physics captured by the simple model, are the concern of the engineering literature. Some of this is less elegant than the simple model, but it allows bridges designed on the same basic model to be built longer and stronger over time, as the complexities and how to deal with them become better understood.

A Flash overview of some topics

Lessons from market failures and successes

To achieve efficient outcomes, marketplaces need make markets sufficiently

- Thick

Enough potential transactions available at one time

- Uncongested

Enough time for offers to be made, accepted, rejected...

- Safe

Safe to act straightforwardly on relevant preferences

Some kinds of transactions are repugnant...
This can be an important constraint on market design

- Medical labor markets
  - NRMP in 1995 (thickness, congestion, incentives)
  - Gastroenterology in 2006 (thickness, incentives)
    - Is reneging on early acceptances repugnant?
- School choice systems:
  - New York City since Sept. 2004 (congestion & incentives)
  - Boston since Sept. 2006 (incentives)
    - Repugnant: exchange of priorities (particularly sibling priorities)
- American market for new economists
  - Scramble ((thickness)
  - Signaling (congestion)
- Kidney exchange (thickness, congestion, incentives)
  - New England and Ohio (2005)
  - National US (2007?)
    - Repugnant: monetary markets

Introduction to the theory of Two-Sided Matching

To see which results are robust, we'll look at some increasingly general models. Even before we look at complex design problems, we can get a head start at figuring out which are our most applicable results by doing this sort of theoretical sensitivity analysis.

Discrete models

- One to one matching: the "marriage" model
- many to one matching (with simple preferences) : the "college admissions" model
- many to one matching with money and complex (gross substitutes) preferences

These lectures follow the Roth and Sotomayor book, and thoerems are numbered as in the book.

These lectures follow the Roth and Sotomayor book, and thoerems are numbered as in the book.

Other resources on the web: Al Roth: Market Design Blog: http://marketdesigner.blogspot.com/

One to One Matching: the Marriage Model

Players: Men: $M = \{m_1, .., m_n\}$, Women: $W = \{w_1, .., w_p\}$.
The market is two-sided: Man m_i can only have preferences over the set of $W \cup \{m_i\}$.
Similarly for women's preferences.

Preferences: (complete and transitive):

$$P(m_i) = w_k, w_l, ..., m_i, w_j ... \qquad [w_k \succ_{m_i} w_l]$$

If m_i prefers to remain single rather than to be matched to w_j, i.e. if $m_i \succ_{m_i} w_j$, then w_j is said to be *unacceptable* to m_i.

If an agent is not indifferent between any two acceptable mates, or between being matched and unmatched, we'll say he/she has strict preferences. Some of the theorems we prove will only be true for strict preferences. Indifferences: $P(m_i) = w_k, [w_l, w_m], ..., m_i,$ where m_i is indifferent between w_l and w_m.

Stable Outcome

An outcome of the game is a matching $\mu : M \cup W \to M \cup W$ such that

- $w = \mu(m) \iff \mu(w) = m$
- $\mu(w) \in M \cup \{w\}$ and $\mu(m) \in W \cup \{m\}$. (two-sided).

Stable Outcome

An outcome of the game is a matching $\mu : M \cup W \to M \cup W$ such that

- $w = \mu(m) \iff \mu(w) = m$
- $\mu(w) \in M \cup \{w\}$ and $\mu(m) \in W \cup \{m\}$. (two-sided).

A matching μ is

- blocked by an individual k : k prefers being single to being matched with $\mu(k)$, i.e. $k \succ_k \mu(k)$.

Stable Outcome

An outcome of the game is a matching $\mu : M \cup W \to M \cup W$ such that

- $w = \mu(m) \iff \mu(w) = m$
- $\mu(w) \in M \cup \{w\}$ and $\mu(m) \in W \cup \{m\}$. (two-sided).

A matching μ is

- blocked by an individual k : k prefers being single to being matched with $\mu(k)$, i.e. $k \succ_k \mu(k)$.
- blocked by a pair of agents (m, w) if they each prefer each other to their current outcome, i.e.
  - $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$

Stable Outcome

An outcome of the game is a matching $\mu : M \cup W \to M \cup W$ such that

- $w = \mu(m) \iff \mu(w) = m$
- $\mu(w) \in M \cup \{w\}$ and $\mu(m) \in W \cup \{m\}$. (two-sided).

A matching μ is

- blocked by an individual k : k prefers being single to being matched with $\mu(k)$, i.e. $k \succ_k \mu(k)$.
- blocked by a pair of agents (m, w) if they each prefer each other to their current outcome, i.e.
    - $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$

A matching μ is *stable* if it isn't blocked by any individual or a pair of agents.

Stable Outcome

An outcome of the game is a matching $\mu : M \cup W \to M \cup W$ such that

- $w = \mu(m) \iff \mu(w) = m$
- $\mu(w) \in M \cup \{w\}$ and $\mu(m) \in W \cup \{m\}$. (two-sided).

A matching μ is

- blocked by an individual k : k prefers being single to being matched with $\mu(k)$, i.e. $k \succ_k \mu(k)$.
- blocked by a pair of agents (m, w) if they each prefer each other to their current outcome, i.e.
  - $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$

A matching μ is *stable* if it isn't blocked by any individual or a pair of agents.

A stable matching is efficient and in the core, and in this simple model the set of (pairwise) stable matchings equals the core.

Deferred Acceptance Algorithm

roughly the 1962 Gale-Shapley Version

- 0. If some preferences are not strict, arbitrarily break ties

Deferred Acceptance Algorithm

roughly the 1962 Gale-Shapley Version

- 0. If some preferences are not strict, arbitrarily break ties
- 1 a. Each man m proposes to his 1st choice (if he has any acceptable choices).
  - b. Each woman rejects any unacceptable proposals and, if more than one acceptable proposal is received, "holds" the most preferred (deferred acceptance).

Deferred Acceptance Algorithm

roughly the 1962 Gale-Shapley Version

- 0. If some preferences are not strict, arbitrarily break ties
- 1 a. Each man m proposes to his 1st choice (if he has any acceptable choices).
  - b. Each woman rejects any unacceptable proposals and, if more than one acceptable proposal is received, "holds" the most preferred (deferred acceptance).

. .

- k a. Any man who was rejected at step $k-1$ makes a new proposal to its most preferred acceptable mate who hasn't yet rejected him. (If no acceptable choices remain, he makes no proposal.)
  - b. Each woman holds her most preferred acceptable offer to date, and rejects the rest.

Deferred Acceptance Algorithm

roughly the 1962 Gale-Shapley Version

- 0. If some preferences are not strict, arbitrarily break ties
- 1 a. Each man *m* proposes to his 1st choice (if he has any acceptable choices).
  - b. Each woman rejects any unacceptable proposals and, if more than one acceptable proposal is received, "holds" the most preferred (deferred acceptance).

. .

- k a. Any man who was rejected at step $k-1$ makes a new proposal to its most preferred acceptable mate who hasn't yet rejected him. (If no acceptable choices remain, he makes no proposal.)
  - b. Each woman holds her most preferred acceptable offer to date, and rejects the rest.
- STOP: when no further proposals are made, and match each woman to the man (if any) whose proposal she is holding.

Theorem 2.8(Gale and Shapley)
A stable matching exists for every marriage market.

Theorem 2.8(Gale and Shapley)
A stable matching exists for every marriage market.

Elements of the proof:

- the deferred acceptance algorithm always stops
- the matching it produces is always stable with respect to the strict preferences (i.e. after any arbitrary tie-breaking),
- and with respect to the original preferences.

The Roommate Problem

Suppose the market is not two-sided, does a stable matching always exist?

Agent 1: $2 \succ 3 \succ 1$
Agent 2: $3 \succ 1 \succ 2$
Agent 3: $1 \succ 2 \succ 3$

All agents being alone is not a core matching.
Any matching with 2 students in a room is not stable either.

Stability is theoretically appealing, but does it matter in real life?

Roth (1984) showed that the NIMP algorithm is equivalent to a (hospital-proposing) DA algorithm, so NIMP produces a stable matching.

Priority matching (an unstable system)
Edinburgh, 1967
Birmingham 1966, 1971, 1978
Newcastle 1970's
Sheffield 196x

All matches are no longer in use:
In a priority matching algorithm, a 'priority' is defined for each firm-worker pair as a function of their mutual rankings. The algorithm matches all priority 1 couples and removes them from the market, then repeats for priority 2 matches, priority 3 , etc.
E.g. in Newcastle, priorities for firm-worker rankings were organized by the product of the rankings, (initially) as follows:

1-1, 2-1, 1-2, 1-3, 3-1, 4-1, 2-2, 1-4, 5-1...

After 3 years, 80% of the submitted rankings were pre-arranged 1-1 rankings without any other choices ranked. This worked to the great disadvantage of those who didn't pre-arrange their matches.

Theorem 2.12 (Gale and Shapley)

When all men and women have strict preferences, there always exists an M-optimal stable matching (that every man likes at least as well as any other stable matching), and a W-optimal stable matching.

Furthermore, the matching μ_M produced by the deferred acceptance algorithm with men proposing is the M-optimal stable matching. The W-optimal stable matching is the matching μ_W produced by the algorithm when the women propose.

Theorem 2.12 (Gale and Shapley)

When all men and women have strict preferences, there always exists an M-optimal stable matching (that every man likes at least as well as any other stable matching), and a W-optimal stable matching.

Furthermore, the matching μ_M produced by the deferred acceptance algorithm with men proposing is the M-optimal stable matching. The W-optimal stable matching is the matching μ_W produced by the algorithm when the women propose.

Terminology:
Woman w is achievable for m if there is some stable μ such that $\mu(m) = w$.

Sketch of Proof:

Inductive step: suppose that up to step k of the algorithm, no m has been rejected by an achievable w, and that at step k w rejects m (who is acceptable to w) and (therefore) holds on to some m'.

We show: w is not achievable for m.

Sketch of Proof:

Inductive step: suppose that up to step k of the algorithm, no m has been rejected by an achievable w, and that at step k w rejects m (who is acceptable to w) and (therefore) holds on to some m'.

We show: w is not achievable for m.
Consider μ with $\mu(m) = w$, and $\mu(m')$ achievable for m'. Can't be stable: by the inductive step, (m', w) would be a blocking pair.

Let $\mu \succ_M \mu'$ denote that all men like μ at least as well as μ', with at least one man having a strict preference.

Let $\mu \succ_M \mu'$ denote that all men like μ at least as well as μ', with at least one man having a strict preference.

Then $\succ_M$ is a partial order on the set of matchings, representing the common preferences of the men. Similarly, define $\succ_W$ as the common preference of the women.

Let $\mu \succ_M \mu'$ denote that all men like μ at least as well as μ', with at least one man having a strict preference.

Then $\succ_M$ is a partial order on the set of matchings, representing the common preferences of the men. Similarly, define $\succ_W$ as the common preference of the women.

Theorem 2.13(Knuth)
When all agents have strict preferences, the common preferences of the two sides of the market are *opposed* on the set of stable matchings: Let μ and μ' be stable matchings. Then $\mu \succ_M \mu'$ if and only if $\mu' \succ_W \mu$.
Proof: immediate from definition of stability.

The best outcome for one side of the market is the worst for the other.

For any two matchings μ and μ', and for all m and w, define $\nu = \mu \vee_M \mu'$ as the function that assigns each man his more preferred of the two matches, and each woman her less preferred:

- $\nu(m) = \mu(m)$ if $\mu(m) \succ_m \mu'(m)$ and $\nu(m) = \mu'(m)$ otherwise.
- $\nu(w) = \mu(w)$ if $\mu(w) \prec_w \mu'(w)$ and $\nu(w) = \mu'(w)$ otherwise.

Define $\nu = \mu \wedge_M \mu'$ analogously, by reversing the preferences.

For any two matchings μ and μ', and for all m and w, define $\nu = \mu \vee_M \mu'$ as the function that assigns each man his more preferred of the two matches, and each woman her less preferred:

- $\nu(m) = \mu(m)$ if $\mu(m) \succ_m \mu'(m)$ and $\nu(m) = \mu'(m)$ otherwise.
- $\nu(w) = \mu(w)$ if $\mu(w) \prec_w \mu'(w)$ and $\nu(w) = \mu'(w)$ otherwise.

Define $\nu = \mu \wedge_M \mu'$ analogously, by reversing the preferences.

Theorem 2.16 Lattice Theorem (Conway):
When all preferences are strict, if μ and μ' are stable matchings, then the functions
$\nu = \mu \vee_M \mu'$ and $\nu = \mu \wedge_M \mu'$ are also stable matchings.

So if we think of ν as asking men to point to their preferred mate from two stable matchings, and asking women to point to their less preferred mate, the theorem says that

- No two men point to the same woman (this follows from the stability of μ and μ')

So if we think of ν as asking men to point to their preferred mate from two stable matchings, and asking women to point to their less preferred mate, the theorem says that

- No two men point to the same woman (this follows from the stability of μ and μ')
- Every woman points back at the man pointing to her;
  - $\nu(m) = w \Longrightarrow \nu(w) = m$: follows easily from stability.
  - $\nu(w) = m \Longrightarrow \nu(m) = w$: takes a bit more work. (We'll come back to this when we prove the Decomposition Lemma, see next slide).

So if we think of ν as asking men to point to their preferred mate from two stable matchings, and asking women to point to their less preferred mate, the theorem says that

- No two men point to the same woman (this follows from the stability of μ and μ')
- Every woman points back at the man pointing to her;
  - $\nu(m) = w \Longrightarrow \nu(w) = m$: follows easily from stability.
  - $\nu(w) = m \Longrightarrow \nu(m) = w$: takes a bit more work. (We'll come back to this when we prove the Decomposition Lemma, see next slide).
- And the resulting matching is stable. : immediately from the stability of μ and μ'.

Let μ, μ' be stable matchings, and for some $m, w = \mu(m) \succ_m \mu'(m) = w'$.
Stability of μ' implies $\mu'(w) \succ_w \mu(w) = m$.
But how about w'?

Let μ, μ' be stable matchings, and for some $m, w = \mu(m) \succ_m \mu'(m) = w'$.
Stability of μ' implies $\mu'(w) \succ_w \mu(w) = m$.
But how about w'?

The Decomposition Lemma (Corollary 2.21, Knuth):
Let μ and μ' be stable matchings in (M, W, P), with all preferences strict. Let $M(\mu)$ $(W(\mu))$ be the set of men (women) who preferμ to μ', and let $M(\mu')$ $(W(\mu'))$ be those who prefer μ'. Then μ and μ' map $M(\mu')$ onto $W(\mu)$ and $M(\mu)$ onto $W(\mu')$.

Let μ, μ' be stable matchings, and for some $m, w = \mu(m) \succ_m \mu'(m) = w'$.
Stability of μ' implies $\mu'(w) \succ_w \mu(w) = m$.
But how about w'?

The Decomposition Lemma (Corollary 2.21, Knuth):
Let μ and μ' be stable matchings in (M, W, P), with all preferences strict. Let $M(\mu)$ $(W(\mu))$ be the set of men (women) who preferμ to μ', and let $M(\mu')$ $(W(\mu'))$ be those who prefer μ'.
Then μ and μ' map $M(\mu')$ onto $W(\mu)$ and $M(\mu)$ onto $W(\mu')$.

Proof: we've just observed above that $\mu(M(\mu))$ is contained in $W(\mu')$. So $|M(\mu)| \leqslant |W(\mu')|$.
Symmetrically, $\mu'(W(\mu')) \subseteq M(\mu)$, so $|M(\mu)| \geqslant |W(\mu')|$
Since μ and μ' are one-to-one (and since $M(\mu)$ and $W(\mu')$ are finite), both μ and μ' are onto (surjective).

Let μ, μ' be stable matchings, and for some $m, w = \mu(m) \succ_m \mu'(m) = w'$.
Stability of μ' implies $\mu'(w) \succ_w \mu(w) = m$.
But how about w'?

The Decomposition Lemma (Corollary 2.21, Knuth):
Let μ and μ' be stable matchings in (M, W, P), with all preferences strict. Let $M(\mu)$ $(W(\mu))$ be the set of men (women) who preferμ to μ', and let $M(\mu')$ $(W(\mu'))$ be those who prefer μ'.
Then μ and μ' map $M(\mu')$ onto $W(\mu)$ and $M(\mu)$ onto $W(\mu')$.

Proof: we've just observed above that $\mu(M(\mu))$ is contained in $W(\mu')$. So $|M(\mu)| \leqslant |W(\mu')|$.
Symmetrically, $\mu'(W(\mu')) \subseteq M(\mu)$, so $|M(\mu)| \geqslant |W(\mu')|$
Since μ and μ' are one-to-one (and since $M(\mu)$ and $W(\mu')$ are finite), both μ and μ' are onto (surjective).

So, an agent who prefers one stable matching to another is matched at both to a mate with the reverse preferences.

Theorem 2.22

In a market (M, W, P) with strict preferences, the set of people who are single is the same for all stable matchings.

Theorem 2.22
In a market (M, W, P) with strict preferences, the set of people who are single is the same for all stable matchings.

One strategy of proof: What can we say about the number and identity of men and women matched (and hence the number and identity unmatched) at μ_M and at μ_W?
i.e. denoting $M_{\mu_M} = \mu_M(W)$, etc. what can we say about the relative sizes and containment relations of the sets M_{μ_M}, W_{μ_M}, M_{μ_W}, W_{μ_W}?

μ_M	$\lvert M_{\mu_M} \rvert$	$\lvert W_{\mu_M} \rvert$
μ_W	$\lvert M_{\mu_W} \rvert$	$\lvert W_{\mu_W} \rvert$

A little disgression: There is an easy graphic representation of the last theorems: Once single, always single and the decomposition lemma (strict preferences).

A little disgression: There is an easy graphic representation of the last theorems: Once single, always single and the decomposition lemma (strict preferences).

Let there be two stable matchings μ and μ'. Let $G(\mu, \mu') = (V, E)$ be a bi-choice graph, with $V = N$ the set of agents (works also for the roommate problem). E consists of 3 types of edges: Let $i, j \in N$.

- $E1 : i \longrightarrow j$ if $j = \mu(i) \succ_i \mu'(i)$

A little disgression: There is an easy graphic representation of the last theorems: Once single, always single and the decomposition lemma (strict preferences).

Let there be two stable matchings μ and μ'. Let $G(\mu, \mu') = (V, E)$ be a bi-choice graph, with $V = N$ the set of agents (works also for the roommate problem). E consists of 3 types of edges: Let $i, j \in N$.

- $E1 : i \longrightarrow j$ if $j = \mu(i) \succ_i \mu'(i)$
- $E2 : i \Longrightarrow j$ if $j = \mu'(i) \succ_i \mu(i)$

A little disgression: There is an easy graphic representation of the last theorems: Once single, always single and the decomposition lemma (strict preferences).

Let there be two stable matchings μ and μ'. Let $G(\mu, \mu') = (V, E)$ be a bi-choice graph, with $V = N$ the set of agents (works also for the roommate problem). E consists of 3 types of edges: Let $i, j \in N$.

- $E1 : i \longrightarrow j$ if $j = \mu(i) \succ_i \mu'(i)$
- $E2 : i \Longrightarrow j$ if $j = \mu'(i) \succ_i \mu(i)$
- $E3 : i \rightsquigarrow j$ if $j = \mu(i) \sim_i \mu'(i)$ (a loop if $i = j$)

A little disgression: There is an easy graphic representation of the last theorems: Once single, always single and the decomposition lemma (strict preferences).

Let there be two stable matchings μ and μ'. Let $G(\mu, \mu') = (V, E)$ be a bi-choice graph, with $V = N$ the set of agents (works also for the roommate problem). E consists of 3 types of edges: Let $i, j \in N$.

- $E1 : i \longrightarrow j$ if $j = \mu(i) \succ_i \mu'(i)$
- $E2 : i \Longrightarrow j$ if $j = \mu'(i) \succ_i \mu(i)$
- $E3 : i \rightsquigarrow j$ if $j = \mu(i) \sim_i \mu'(i)$ (a loop if $i = j$)

Lemma:
Consider $G(\mu, \mu')$. Let $i \in N$. Then agent $i's$ component of $G(\mu, \mu')$ is either:

- $i \rightsquigarrow j$ for some agent j (possibly $j = i$)
- a directed even cycle of 4 or more agents, where edges of 1 of 2 lines alternate.

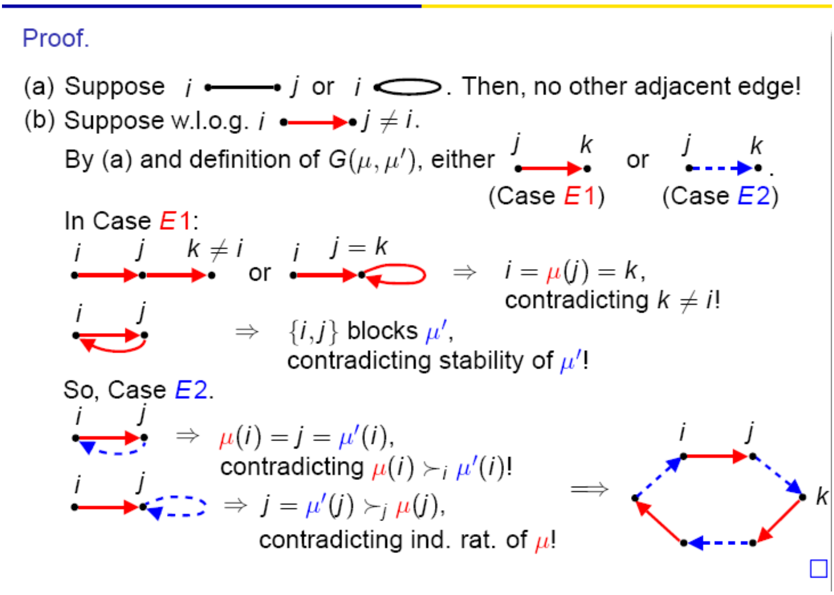


Figure: Slide from B. Klaus

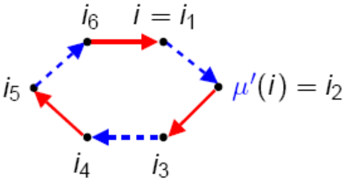
Theorem

Lonely wolves

μ and μ' have the same set of single agents, i.e., $\mu(i) = i \Leftrightarrow \mu'(i) = i$.

Proof.

Suppose w.l.o.g. $\mu(i) = i$ but $\mu'(i) \neq i$. Then,



$$\Rightarrow \begin{cases} i_6 \neq i \\ \mu(i_6) = i, \text{ i.e., } \mu(i) = i_6 \end{cases}$$

$$\Downarrow$$

$$\mu(i) \neq i \Rightarrow \text{contradiction!}$$

Figure: Slide from B. Klaus

Lemma

Decomposability

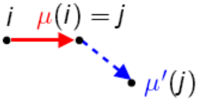
Let $\mu(i) = j$. Then,

(a) $\mu(i) \succ_i \mu'(i)$ *implies* $\mu'(j) \succ_j \mu(j)$ *and*

(b) $\mu'(i) \succ_i \mu(i)$ *implies* $\mu(j) \succ_j \mu'(j)$.

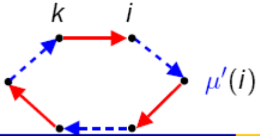
Proof.

(a) Suppose $j = \mu(i) \succ_i \mu'(i)$. Then, lonely wolf theorem: $j, \mu'(i) \neq i$. Moreover,



$\Rightarrow \mu'(j) \succ_j \mu(j)$.

(b) Suppose $\mu'(i) \succ_i \mu(i) = j$. Then, lonely wolf theorem: $j, \mu'(i) \neq i$. Moreover,



$\Rightarrow$ $i = \mu(k)$, i.e., $j = k$ and $\mu(j) \succ_j \mu'(j)$.

$\square$

Example of such a graph $G(\mu, \mu')$:

$1 \circlearrowleft 1$
$2 \sim 3$
$4 \sim 5$

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm...)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm...)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Hence all of these women, $\mu(M)$, would have been matched under μ_M. That is $\mu_M(\mu(M)) = M$.

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm. . .)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Hence all of these women, $\mu(M)$, would have been matched under μ_M. That is $\mu_M(\mu(M)) = M$.

Hence all of M would have been matched under μ_M and $\mu_M(M) = \mu(M)$.

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm. . .)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Hence all of these women, $\mu(M)$, would have been matched under μ_M. That is $\mu_M(\mu(M)) = M$.

Hence all of M would have been matched under μ_M and $\mu_M(M) = \mu(M)$.

But since all of M are matched under μ_M any woman who gets a proposal in the last step of the algorithm at which proposals were issued has not rejected any acceptable man, i.e. the algorithm stops as soon as every woman in $\mu_M(M)$ has an acceptable proposal.

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm. . .)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Hence all of these women, $\mu(M)$, would have been matched under μ_M. That is $\mu_M(\mu(M)) = M$.

Hence all of M would have been matched under μ_M and $\mu_M(M) = \mu(M)$.

But since all of M are matched under μ_M any woman who gets a proposal in the last step of the algorithm at which proposals were issued has not rejected any acceptable man, i.e. the algorithm stops as soon as every woman in $\mu_M(M)$ has an acceptable proposal.

So such a woman must be single at μ (since every man prefers μ to μ_M), which contradicts the fact that $\mu_M(M) = \mu(M)$.

Theorem 2.27 Weak Pareto optimality for the men: There is no individually rational matching μ (stable or not) such that $\mu \succ_m \mu_M$ for all $m \in M$.

Proof (using the deferred acceptance algorithm...)

If μ were such a matching it would match every man m to some woman w who had rejected him in the algorithm in favor of some other man m' (i.e. even though m was acceptable to w).

Hence all of these women, $\mu(M)$, would have been matched under μ_M. That is $\mu_M(\mu(M)) = M$.

Hence all of M would have been matched under μ_M and $\mu_M(M) = \mu(M)$.

But since all of M are matched under μ_M any woman who gets a proposal in the last step of the algorithm at which proposals were issued has not rejected any acceptable man, i.e. the algorithm stops as soon as every woman in $\mu_M(M)$ has an acceptable proposal.

So such a woman must be single at μ (since every man prefers μ to μ_M), which contradicts the fact that $\mu_M(M) = \mu(M)$.

Example 2.31 (μ_M not strongly Pareto optimal for the men)

$M = \{m1, m2, m3\}$,
$P(m1) = w2, w1, w3$
$P(m2) = w1, w2, w3$
$P(m3) = w1, w2, w3$

$W = \{w1, w2, w3\}$
$P(w1) = m1, m2, m3$
$P(w2) = m3, m1, m2$
$P(w3) = m1, m2, m3$

$\mu_M = ([m1, w1], [m2, w3], [m3, w2]) = \mu_W$

But note that $\mu \succ_M \mu_M$ for $\mu = ([m1, w2], [m2, w3], [m3, w1])$

Example 2.31 (μ_M not strongly Pareto optimal for the men)

$M = \{m1, m2, m3\}$, $W = \{w1, w2, w3\}$

$P(m1) = w2, w1, w3$ $P(w1) = m1, m2, m3$

$P(m2) = w1, w2, w3$ $P(w2) = m3, m1, m2$

$P(m3) = w1, w2, w3$ $P(w3) = m1, m2, m3$

$$\mu_M = ([m1, w1], [m2, w3], [m3, w2]) = \mu_W$$

But note that $\mu \succ_M \mu_M$ for $\mu = ([m1, w2], [m2, w3], [m3, w1])$

Note: This is also shows that while in the whole game (M, W, P) a stable matching μ may be unique, this may not be the case for a subset (M', W', P), even as $M' = \mu(W')$.

Lemma 3.5 Blocking Lemma (Gale and Sotomayor)
Let μ be any individually rational matching with respect to strict preferences P and let M' be all men who prefer μ to μ_M. If M' is nonempty there is a pair (m, w) which blocks μ such that m is in $M \backslash M$' and w is in $\mu(M')$.

Lemma 3.5 Blocking Lemma (Gale and Sotomayor)
Let μ be any individually rational matching with respect to strict preferences P and let M' be all men who prefer μ to μ_M. If M' is nonempty there is a pair (m, w) which blocks μ such that m is in $M \backslash M$' and w is in $\mu(M')$.

Proof : Case I: $\mu(M') \neq \mu_M(M')$: Choose $w \in \mu(M') \backslash \mu_M(M')$, say $w = \mu(m')$. Then $m' : w \succ_{m'} \mu_M(m')$ so $w :$
$\mu_M(w) = m \succ_w m'$.
But $m \notin M'$, since $w \notin \mu_M(M')$, hence $m : w \succ_m \mu(m)$ (since preferences are strict), so (m, w) blocks μ.

Case II: $\mu_M(M') = \mu(M') = W'$. Let w be the woman in W' who receives the last proposal from an acceptable member of M' in the deferred acceptance algorithm. Since all w in W' have rejected acceptable men from M', w had some man m engaged when she received this last proposal. We claim $(m, w)$is the desired blocking pair.

Case II: $\mu_M(M') = \mu(M') = W'$. Let w be the woman in W' who receives the last proposal from an acceptable member of M' in the deferred acceptance algorithm. Since all w in W' have rejected acceptable men from M', w had some man m engaged when she received this last proposal. We claim $(m, w)$is the desired blocking pair.

First, m is not in M' for if so, after having been rejected by w, he would have proposed again to a member of W' contradicting the fact that w received the last such proposal. But m prefers w to his mate under μ_M and since he is no better off under μ, he prefers w to $\mu(m)$.

Case II: $\mu_M(M') = \mu(M') = W'$. Let w be the woman in W' who receives the last proposal from an acceptable member of M' in the deferred acceptance algorithm. Since all w in W' have rejected acceptable men from M', w had some man m engaged when she received this last proposal. We claim $(m, w)$is the desired blocking pair.

First, m is not in M' for if so, after having been rejected by w, he would have proposed again to a member of W' contradicting the fact that w received the last such proposal. But m prefers w to his mate under μ_M and since he is no better off under μ, he prefers w to $\mu(m)$.

On the other hand, m was the last man to be rejected by w so she must have rejected her mate under μ before she rejected m and hence she prefers m to $\mu(w)$, so (m, w) blocks μ.

Strategic Behavior

So far, we assumed we know the preferences. But, when we apply such a centralized matching procedure to actual markets, we will have to elicit preferences.

Strategic Behavior

So far, we assumed we know the preferences. But, when we apply such a centralized matching procedure to actual markets, we will have to elicit preferences.
Let's consider strategic behavior in centralized matching mechanisms, in which participants submit a list of stated preferences. By the revelation principle, some of the results will apply to decentralized markets also, in which agents have different sets of strategies.

Strategic Behavior

So far, we assumed we know the preferences. But, when we apply such a centralized matching procedure to actual markets, we will have to elicit preferences.

Let's consider strategic behavior in centralized matching mechanisms, in which participants submit a list of stated preferences. By the revelation principle, some of the results will apply to decentralized markets also, in which agents have different sets of strategies.

Consider a marriage market (M, W, P) whose outcome will be determined by a centralized clearinghouse, based on a list of preferences that players will state ("reveal"). If the vector of stated preferences is Q, the algorithm employed by the clearinghouse produces a matching $h(Q)$. The *matching mechanism* h is defined for all (M, W, Q). If the matching produced is *always* a stable matching with respect to Q, we'll say that h is a stable matching mechanism.

Theorem 4.4 Impossibility Theorem (Roth):
No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent.

Theorem 4.4 Impossibility Theorem (Roth):
No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent.

Consider an example with 2 agents on each side, with true preferences $P = (Pm1, Pm2, Pw1, Pw2)$ as follows:

$m1 : w1, w2$	$w1 : m2, m1$
$m2 : w2, w1$	$w2 : m1, m2$

In this example, what must an arbitrary stable mechanism do? I.e. what is the range of $h(P)$ if h is a stable mechanism?

Theorem 4.4 Impossibility Theorem (Roth):
No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent.

Consider an example with 2 agents on each side, with true preferences $P = (Pm1, Pm2, Pw1, Pw2)$ as follows:

$m1 : w1, w2$	$w1 : m2, m1$
$m2 : w2, w1$	$w2 : m1, m2$

In this example, what must an arbitrary stable mechanism do? I.e. what is the range of $h(P)$ if h is a stable mechanism?
Given $h(P)$, and the restriction that h is a stable mechanism, can one of the players x engage in a profitable manipulation by stating some $P'_x \neq P_x$ such that x prefers $h(P')$ to $h(P)$?

Of course, this kind of proof of the impossibility theorem leaves open the possibility that situations in which some participant can profitably manipulate his preferences are rare. The following result suggests otherwise.

Theorem 4.6
When any stable mechanism is applied to a marriage market in which preferences are strict and there is more than one stable matching, then at least one agent can profitably misrepresent his or her preferences, assuming the others tell the truth. (This agent can misrepresent in such a way as to be matched to his or her most preferred achievable mate under the true preferences at every stable matching under the mis-stated preferences.)

Incentives facing the men when the M-optimal stable mechanism is used

Theorem 4.7 (Dubins and Freedman, Roth)
The mechanism that yields the M-optimal stable matching (in terms of the stated preferences) makes it a dominant strategy for each man to state his true preferences.

Theorem 4.10 (Dubins and Freedman)
Let P be the true preferences of the agents, and let P' differ from P in that some coalition M' of the men mis-state their preferences. Then there is no matching μ, stable for P', which is preferred to μ_M by all members of M'.

Theorem 4.11 (Limits on successful manipulation.) (Demange, Gale, and Sotomayor).
Let P be the true preferences (not necessarily strict) of the agents, and let P' differ from P in that some coalition C of men and women mis-state their preferences. Then there is no matching μ, stable for P', which is preferred to every stable matching under the true preferences P by all members of C.

Note that Theorem 4.11 implies both Theorems 4.7 and 4.10.

Proof of Theorem 4.11:

Suppose some nonempty subset $M' \cup W'$ of men and women mis-state their preferences and are strictly better off under μ, stable w.r.t. P', than under any stable matching w.r.t. P.
μ must be individually rational wrt P, even though unstable.

$\{$Now construct strict preferences $\widetilde{P}$, so that if any agent x is indifferent under P between $\mu(x)$ and some other alternative, then under $\widetilde{P}$ x prefers $\mu(x)$ (but otherwise make no change in the ordering of preferences P). Then (m, w) blocks μ under $\widetilde{P}$ only if (m, w) blocks μ under P. Since every stable matching under $\widetilde{P}$ is also stable under P,
$\mu(m) \succ_m \mu_M(m)$ for every $m \in M'$ and $\mu(w) \succ_w \mu_W(w)$ for every $w \in W'$. (*)
where μ_M and μ_W are the M- and W- optimal stable matchings for $(M, W, \widetilde{P})$.$\}$

If M' is not empty we can apply the Blocking Lemma (3.5) to the market $(M, W, \widetilde{P})$, since by (*) M' is a subset of $\widetilde{M}$ the set of men who prefer μ to μ_M; thus there is a pair (m, w) which blocks μ under $\widetilde{P}$ and so under P such that
$\mu_M(m) \succ_m \mu(m)$ and
$\mu_M(w) \succ_w \mu(w)$ (otherwise w and $\mu(w)$ would block μ_M, since w is in $\mu(\widetilde{M})$ by the blocking lemma).
Clearly m and w are not in $M' \cup W'$ and so are not mis-stating their preferences, so they will also block μ under P', contradicting that μ is stable under P'.
If M' is empty W' is not, and the symmetrical argument applies.

What can we say about equilibrium?

Pure strategy equilibria exist:

Theorem 4.15 (Gale and Sotomayor)

When all preferences are strict, let μ be any stable matching for (M, W, P). Suppose each woman w in $\mu(M)$ chooses the strategy of listing only $\mu(w)$ on her stated preference list of acceptable men (and each man states his true preferences). This is an equilibrium in the game induced by the M-optimal stable matching mechanism (and μ is the matching that results).

Furthermore, every equilibrium mis-representation by the women must nevertheless yield a matching that is stable with respect to the true preferences. (But the proof of this should raise doubts about it's applicability:)

Theorem 4.16 (Roth)

Suppose each man chooses his dominant strategy and states his true preferences, and the women choose any set of strategies (preference lists) P_w' that form an equilibrium for the matching game induced by the M-optimal stable mechanism. Then the corresponding M-optimal stable matching for (M, W, P') is one of the stable matchings of (M, W, P).

Many-to-one matching: The college admissions model

Players: Firms $\{f_1, .., f_n\}$ with number of positions: $q_1, ..., q_n$
Workers: $\{w_1, ..., w_p\}$
Synonyms (sorry:): $F =$ Firms, $C =$ Colleges, $H =$ Hospitals
$W =$ Workers, $S =$ Students.

Preferences over individuals (complete and transitive), as in the marriage model:

$$P(fi) = w3, w2, ...fi... \qquad [w3 \succ_{f_i} w2]$$
$$P(wj) = f2, f4, ...wj...$$

An outcome of the game is a matching: $\mu : F \cup W \longrightarrow F \cup W$ s.t.

- $w \in \mu(f)$ iff $\mu(w) = f$ for all f, w
- $|\mu(f)| \leqslant q_f$
- $\mu(w) \in F \cup \{w\}$.

so f is matched to the set of workers $\mu(f)$.

We need to specify how firms' preferences over matchings, are related to their preferences over individual workers, since they hire groups of workers. The simplest model is
Responsive preferences: for any set of workers $S \subseteq W$ with $|S| < q_i$, and any workers w and w' in $W \backslash S$:
$S \cup w >_{f_i} S \cup w'$ if and only if $w >_{f_i} w'$, and
$S \cup w >_{f_i} S$ if and only if w is acceptable to f_i.
A matching μ is individually irrational (and blocked by the relevant individual) if $\mu(w) = f$ for some w, f such that either the worker is unacceptable to the firm or the firm is unacceptable to the student.
A matching μ is BLOCKED BY A PAIR OF AGENTS (f, w) if they each prefer each other to μ:

- $w >_f w'$ for some w' in $\mu(f)$ or $w >_f f$ if $|\mu(f)| < q_f$
- $f >_w \mu(w)$

As in the marriage model, a matching is (pairwise) stable if it isn't blocked by any individual or pair of agents.

But now that firms employ multiple workers, it might not be enough to concentrate only on pairwise stability. The assumption of responsive preferences allows us to do this, however.
A matching μ is blocked by a coalition A of firms and workers if there exists another matching μ' such that for all workers w in A, and all firms f in A

- $\mu'(w) \in A$
- $\mu'(w) \succ_w \mu(w)$
- $\widetilde{w} \in \mu'(f)$ implies $\widetilde{w} \in A \cup \mu(f)$ (i.e. every firm in A is matched at μ' to new students only from A, although it may continue to be matched with some of its "old" students from μ. (THIS DIFFERS FROM THE STANDARD DEFINITION OF THE CORE...)
- $\mu'(f) \succ_f \mu(f)$

A matching is group stable if it is not blocked by a coalition of any size.

Lemma 5.5: When preferences are responsive, a matching is group stable if and only if it is (pairwise) stable.

Proof: instability clearly implies group instability.
Now suppose μ is blocked via coalition A and outcome μ'. Then there must be a worker w and a firm f such that w is in $\mu'(f)$ but not in $\mu(f)$ such that w and f block μ. (Otherwise it couldn't be that $\mu'(f) \succ_f \mu(f)$, since f has responsive preferences.)

A related marriage market

Replace college C by q_C positions of C denoted by $c_1, c_2, \ldots, c_{q_C}$. Each of these positions has C's preferences over individuals. Since each position c_i has a quota of 1, we do not need to consider preferences over groups of students.

Each student's preference list is modified by replacing C, wherever it appears on his list, by the string $c_1, c_2, \ldots, c_{q_C}$, in that order.

A matching μ of the college admissions problem, corresponds to a matching μ' in the related marriage market in which the students in $\mu(C)$ are matched, in the order which they occur in the preferences $P(C)$, with the ordered positions of C that appear in the related marriage market. (If preferences are not strict, there will be more than one such matching.)

Lemma 5.6: A matching of the college admissions problem is stable if and only if the corresponding matchings of the related marriage market are stable.

(NB: some results from the marriage model translate, but not those involving both stable and unstable matchings. . .)

Geographic distribution:
Theorem 5.12
When all preferences over individuals are strict, the set of students employed and positions filled is the same at every stable matching.
The proof is immediate via the similar result for the marriage problem and the construction of the corresponding marriage problem (Lemma 5.6).
So any hospital that fails to fill all of its positions at some stable matching will not be able to fill any more positions at any other stable matching. The next result shows that not only will such a hospital fill the same number of positions, but it will fill them with exactly the same interns at any other stable matching.

Theorem 5.13 Rural hospitals theorem (Roth '86):
When preferences over individuals are strict, any hospital that does not fill its quota at some stable matching is assigned precisely the same set of students at every stable matching.
(This will be easy to prove after Lemma 5.25)

Comparison of stable matchings in the college admissions model: Overview: suppose one of the colleges, C, evaluates students by their scores on an exam, and evaluates entering classes according to their average score on the exam. (So even when we assume no two students have exactly the same score, so that college C's preferences over individuals are strict, it does not have strict preferences over entering classes, since it is indifferent between two entering classes with the same average score.) Then different stable matchings may give college C different entering classes. However *no two distinct entering classes that college C could have at stable matchings will have the same average exam score.* Furthermore, for any two distinct entering classes that college C could be assigned at stable matchings, we can make the following strong comparison. Aside from the students who are in both entering classes, every student in one of the entering classes will have a higher exam score than any student in the other entering class.

Lemma 5.25 (Roth and Sotomayor)
Suppose colleges and students have strict individual preferences, and let μ and μ' be stable matchings for (S, C, P), such that $\mu(C) \neq \mu'(C)$ for some C. Let $\widetilde{\mu}$ and $\widetilde{\mu}'$ be the stable matchings corresponding to μ and μ' in the related marriage market. If $\widetilde{\mu}(c_i) \succ_C \widetilde{\mu}'(c_i)$ for some position c_i of C then $\widetilde{\mu}(c_i) \succ_C \widetilde{\mu}'(c_i)$ for all positions c_i of C.

Proof: It is enough to show that $\widetilde{\mu}(c_j) \succ_C \widetilde{\mu}'(c_j)$ for all $j > i$. So suppose this is false. Then there exists an index j such that $\widetilde{\mu}(c_j) \succ_C \widetilde{\mu}'(c_j)$, but $\widetilde{\mu}'(c_{j+1}) \succeq_C \widetilde{\mu}(c_{j+1})$. Theorem 5.12 (constant employment) implies $\widetilde{\mu}'(c_j) \in S$. Let $s' = \widetilde{\mu}'(c_j)$. By the decomposition lemma $c_j = \widetilde{\mu}'(s').\succ_{s'} \widetilde{\mu}(s')$.
Furthermore, $\widetilde{\mu}(s') \neq c_{j+1}$, since $s' \succ_C \widetilde{\mu}'(c_{j+1}) \succeq_C \widetilde{\mu}(c_{j+1})$ (where the first of these preferences follows from the fact that for any stable matching $\widetilde{\mu}'$ in the related marriage market, $\widetilde{\mu}'(c_j) \succ_C \widetilde{\mu}'(c_{j+1})$ for all j).
Therefore c_{j+1} comes right after c_j in the preferences of s' in the related marriage problem. So $\widetilde{\mu}$ is blocked via s' and c_{j+1}, contradicting (via Lemma 5.6) the stability of μ.
(This proof also establishes the rural hospitals theorem).

Theorem 5.26: (Roth and Sotomayor)
If colleges and students have strict preferences over individuals, then colleges have strict preferences over those groups of students that they may be assigned at stable matchings. That is, if μ and μ' are stable matchings, then a college C is indifferent between $\mu(C)$ and $\mu'(C)$ only if $\mu(C) = \mu'(C)$.

Proof: via the lemma, and repeated application of responsive preferences.

Theorem 5.27: (Roth and Sotomayor)
Let preferences over individuals be strict, and let μ and μ' be stable matchings for (S, C, P). If $\mu(C) \succ_C \mu'(C)$ for some college C, then $s \succ_C s'$ for all s in $\mu(C)$ and s' in $\mu'(C) \backslash \mu(C)$. That is, C prefers every student in its entering class at μ to every student who is in its entering class at μ' but not at μ.

Proof: Consider the related marriage market and the stable matchings $\widetilde{\mu}$ and $\widetilde{\mu}'$ corresponding to μ and μ'. Let $q_C = k$, so that the positions of C are $c_1, .., c_k$.
First observe that C fills its quota under μ and μ', since, if not, Theorem 5.13 (Rural hospitals) would imply that $\mu(C) = \mu'(C)$.
So $\mu'(C) \backslash \mu(C)$.is a nonempty subset of S, since $\mu(C) \neq \mu'(C)$.
Let $s' = \widetilde{\mu}'(c_j)$ for some position c_j such that s' is not in $\mu(C)$.
Then $\widetilde{\mu}(c_j) \neq \widetilde{\mu}'(c_j)$.

By Lemma 5.25 $\widetilde{\mu}(c_j) \succ_C \widetilde{\mu}'(c_j) = s'$.

The Decomposition Lemma implies $c_j \succ_{s'} \widetilde{\mu}(s')$.
So the construction of the related marriage problem implies $c \succ_{s'} \mu(s')$, since $C \succ_{s'} \mu(s')$, since $\mu(s') \neq C$.
Thus $s \succ_C s'$ for all $s \in \mu(C)$ by the stability of μ, which completes the proof.

So, for considering stable matchings, we have some slack in how carefully we have to model preferences over groups. (This is lucky for design, since it reduces the complication of soliciting preferences from firms with responsive preferences. . .)

The results also have an unusual mathematical aspect, since they allow us to say quite a bit about stable matchings even without knowing all the preferences of potential blocking pairs.

Consider a College C with quota 2 and preferences over individuals $P(C) = s1, s2, s3, s4$. Suppose that at various matchings $1-4$, C is matched to

1.	$\{s1, s4\}$,	3.	$\{s1, s3\}$, and
2.	$\{s2, s3\}$,	4.	$\{s2, s4\}$.

Which matchings can be simultaneously stable for some responsive preferences over individuals?

So long as all preferences over groups are responsive, matchings 1 and 2 cannot both be stable (Lemma 5.25), nor can matchings 3 and 4 (Theorem 5.27).

Strategic questions in the College Admissions model:

Theorem 5.16 (Roth)

A stable matching procedure which yields the student-optimal stable matching makes it a dominant strategy for all students to state their true preferences.

Proof: immediate from the related marriage market

Theorem 5.14 (Roth)

No stable matching mechanism exists that makes it a dominant strategy for all hospitals to state their true preferences.

Proof: consider a market consisting of 3 hospitals and 4 students. H1 has a quota of 2, and both other hospitals have a quota of 1. The preferences are:

$s1 : H3, H1, H2$	$s3 : H1, H3, H2$	$H1 : s1, s2, s3, s4$
$s2 : H2, H1, H3$	$s4 : H1, H2, H3$	$H2 : s1, s2, s3, s4$
		$H3 : s3, s1, s2, s4$

The unique stable matching is $\{[H1, s3, s4], [H2, s2], [H3, s1]\}$

But if H1: instead submitted the preferences $s1, s4$ the unique stable matching is $\{[H1, s1, s4], [H2, s2], [H3, s3]\}$.

A More Complex Market: Matching with Couples

This model is the same as the college admissions model , except the set of workers is replaced by a set of applicants that includes both individuals and couples.

Denote the set of applicants by $A = A_1 \cup C$, where A_1 is the set of (single) applicants who seek no more than one position, and C is the set of couples $\{a_i, a_j\}$ such that a_i is in the set A_2 (of husbands) and a_j is in the set A_3 (of wives), and the sets of applicants A_1, A_2, and A_3 together make up the entire population of individual applicants, $A = A_1 \cup A_2 \cup A_3$.

Each couple $c = \{a_i, a_j\}$ in C has preferences over ordered pairs of positions, i.e. an ordered list of elements of $F \times F$. The first element of this list is some (r_i, r_j) in $F \times F$. which is the couples' first choice pair of jobs for a_i and a_j respectively, and so forth.

Applicants in the set A_1 have preferences over residency programs, and residency programs (firms) have preferences over the individuals in A, just as in the simple model discussed earlier. A matching is a set of pairs in $F \times A$.

Each single applicant, each couple, and each residency program submits to the centralized clearinghouse a Rank Order List (ROL) that is their stated preference ordering of acceptable alternatives.

A matching μ is blocked by a single applicant (in the set A_1), or by a residency program, if μ matches that agent to some individual or residency program not on its ROL. A matching is blocked by an individual couple $\{a_i, a_j\}$ if they are matched to a pair (r_i, r_j) not on their ROL.
A residency program r and a single applicant a in A_1 together block a matching μ precisely as in the college admissions market, if they are not matched to one another and would both prefer to be.

A couple $c = \{a_i, a_j\}$ in A and residency programs r_1 and r_2 in F block a matching μ if the couple prefers (r_1, r_2) to $\mu(c)$, and if either r_1 and r_2 each would prefer to be matched to the corresponding member of the couple, or if one of them would prefer, and the other already is matched to the corresponding couple member. That is, c and (r_1, r_2) block μ if

- $.(r_1, r_2) \succ_c \mu(c)$ and if

either

- $\{$ $a_1 \notin \mu(r_1)$ and $a_1 \succ_{r_1} a_i$ for some $a_i \in \mu(r_1)$ or a_1 is acceptable to r_1 and $|\mu(r_1)| < q_1$ $\}$ *and* either $a_2 \in \mu(r_2)$ or $\{a_2 \notin \mu(r_2)$, $a_2 \succ_{r_2} a_j$ for some $a_j \in \mu(r_1)$ or a_2 is acceptable to r_2 and $|\mu(r_2)| < q_2$ $\}$

or

- $\{a_2 \notin \mu(r_2)$, $a_2 \succ_{r_2} a_j$ for some $a_j \in \mu(r_2)$ or a_2 is acceptable to r_2 and $|\mu(r_2)| < q_2\}$ and $a_1 \in \mu(r_1)$.

A matching is stable if it is not blocked by any individual agent or by a pair of agents consisting of an individual and a residency program, or by a couple together with one or two residency programs

Theorem 5.11 (Roth '84): In the college admissions model with couples, the set of stable matchings may be empty.

Proof: by example.

Furthermore, the following example shows that even when the set of stable matchings is non-empty, it may no longer have the nice properties we've come to expect.

Matching with couples (Example of Aldershof and Carducci, '96)

4 hospitals $\{h1, \ldots h4\}$ each with one position;

2 couples $\{s1, s2\}$ and $\{s3, s4\}$

Preferences:

h_1	h_2	h_3	h_4	$\{s_1, s_2\}$	$\{s_3, s_4\}$
s_4	s_2	s_2	s_2	h_3, h_2	h_2, h_1
s_3	s_3	s_4	s_3	h_2, h_3	h_2, h_3
	s_1	s_1		h_2, h_4	h_1, h_3
				h_3, h_4	h_4, h_1
				u, h_3	h_4, h_3
				u, h_2	
				u, h_4	
				h_3, u	
				h_2, u	

There are exactly two stable matchings:

$(h_1, s_4), (h_2, s_2), (h_3, s_1), (h_4, s_3)$ preferred by $h_1, h_2, h_4, \{s_1, s_2)$

$(h_1, s_4), (h_2, s_3), (h_3, s_{12}), (h_4, u)$ preferred by $h_3, \{s_2, s_3)$

So, even when stable matchings exist, there need not be an optimal stable matching for either side, and employment levels may vary.