Matching and Market Design: Overview

Fuhito Kojima\(^1\)

February 3, 2009

---

\(^1\)Yale University. http://sites.google.com/site/fuhitokojimaeconomics/. I am grateful to Al Roth and Utku Unver for allowing me to consult with their slides when I was preparing the lectures. Any error is my own responsibility.
What is Matching and Market Design?

- Most traditional economics focuses on analyzing economy as it is.
What is Matching and Market Design?

- Most traditional economics focuses on analyzing economy as it is.
- Economists have recently developed and applied the economics of “matching and market design” to design institutions.
Most traditional economics focuses on analyzing economy as it is.

Economists have recently developed and applied the economics of “matching and market design” to design institutions.

I will talk about the subfield “two-sided matching” theory, and it’s applications to (1) labor markets such as NRMP for medical residents and residency programs, and (2) student placement mechanisms, such as those in NYC and Boston.

- Medical students in many countries work as residents (interns) at hospitals.

- Medical students in many countries work as residents (interns) at hospitals.
- In the U.S. more than 20,000 medical students and 4,000 hospitals are matched.

- Medical students in many countries work as residents (interns) at hospitals.
- In the U.S. more than 20,000 medical students and 4,000 hospitals are matched.
- Beginning around 1900, the American hospital-intern market was decentralized, and suffered from unraveling of appointment dates.

- Medical students in many countries work as residents (interns) at hospitals.
- In the U.S. more than 20,000 medical students and 4,000 hospitals are matched.
- Beginning around 1900, the American hospital-intern market was decentralized, and suffered from *unraveling* of appointment dates.
- Students and hospitals made contracts 2 years in advance of graduation.

- Medical students in many countries work as residents (interns) at hospitals.
- In the U.S. more than 20,000 medical students and 4,000 hospitals are matched.
- Beginning around 1900, the American hospital-intern market was decentralized, and suffered from *unraveling* of appointment dates.
- Students and hospitals made contracts 2 years in advance of graduation.
- There were a lot of mismatch because students’ quality and interests were unknown early in the study.

- Medical students in many countries work as residents (interns) at hospitals.
- In the U.S. more than 20,000 medical students and 4,000 hospitals are matched.
- Beginning around 1900, the American hospital-intern market was decentralized, and suffered from *unraveling* of appointment dates.
- Students and hospitals made contracts 2 years in advance of graduation.
- There were a lot of mismatch because students’ quality and interests were unknown early in the study.
- This caused inefficiency, and doctors and hospitals tried to change their system.
A centralized matching mechanism called NIMP (now called NRMP) was established in 1952.
A centralized matching mechanism called NIMP (now called NRMP) was established in 1952. Students submitted rank order lists over hospitals and hospitals submitted rank order lists over students. The NIMP used these lists to decide who works where.
A centralized matching mechanism called NIMP (now called NRMP) was established in 1952. Students submitted rank order lists over hospitals and hospitals submitted rank order lists over students. The NIMP used these lists to decide who works where.

- The clearinghouse prevented unraveling, and is in use now.
A centralized matching mechanism called NIMP (now called NRMP) was established in 1952. Students submitted rank order lists over hospitals and hospitals submitted rank order lists over students. The NIMP used these lists to decide who works where.

The clearinghouse prevented unraveling, and is in use now.

Questions: Why did NIMP algorithm stop unraveling? Is there any room for improvement of the mechanism?
School Choice

In many countries, children were automatically sent to a school in their neighborhoods.
School Choice

- In many countries, children were automatically sent to a school in their neighborhoods.
- Recently, more and more cities in the United States and in other countries employ school choice programs: school authorities take into account preferences of children and their parents.
In many countries, children were automatically sent to a school in their neighborhoods.

Recently, more and more cities in the United States and in other countries employ school choice programs: school authorities take into account preferences of children and their parents.

Typical goals of school authorities are: (1) efficient placement, (2) fairness of outcomes, (3) easy for participants to understand and use, etc.
Abdulkadiroglu and Sonmez (2003) showed that placement mechanisms used in many cities such as Boston are flawed. (1) the mechanism is manipulable, i.e., students may benefit by report false preferences, and (2) the result may be neither fair nor efficient. They proposed new mechanisms to improve upon existing placement mechanisms.
Abdulkadiroglu and Sonmez (2003) showed that placement mechanisms used in many cities such as Boston are flawed. (1) the mechanism is manipulable, i.e., students may benefit by report false preferences, and (2) the result may be neither fair nor efficient. They proposed new mechanisms to improve upon existing placement mechanisms.

Based on this and other studies, Boston and New York City changed their student placement mechanisms.
Abdulkadiroglu and Sonmez (2003) showed that placement mechanisms used in many cities such as Boston are flawed. (1) the mechanism is manipulable, i.e., students may benefit by report false preferences, and (2) the result may be neither fair nor efficient. They proposed new mechanisms to improve upon existing placement mechanisms.

Based on this and other studies, Boston and New York City changed their student placement mechanisms.

The mechanism seems to work well. Example: NYC’s number of “unassigned” children.
Abdulkadiroglu and Sonmez (2003) showed that placement mechanisms used in many cities such as Boston are flawed. (1) the mechanism is manipulable, i.e., students may benefit by report false preferences, and (2) the result may be neither fair nor efficient. They proposed new mechanisms to improve upon existing placement mechanisms.

Based on this and other studies, Boston and New York City changed their student placement mechanisms.

The mechanism seems to work well. Example: NYC’s number of “unassigned” children.

Many studies are currently conducted to evaluate the current school choice mechanisms, and several mechanisms are proposed to improve the outcome.
A simple theory of matching

- Proposed by Gale and Shapley (1962). Suggested reading is Roth and Sotomayor (RS henceforth), Chapter 2.
A simple theory of matching

- Proposed by Gale and Shapley (1962). Suggested reading is Roth and Sotomayor (RS henceforth), Chapter 2.
- Finite sets $S$ of students and $C$ of colleges (we use student-college terminology just for convenience).
A simple theory of matching

- Proposed by Gale and Shapley (1962). Suggested reading is Roth and Sotomayor (RS henceforth), Chapter 2.
- Finite sets $S$ of students and $C$ of colleges (we use student-college terminology just for convenience).
- Each student can be matched to at most one college, and each college can admit at most one student (so the model is called “one-to-one matching”). Students have strict preferences over colleges and being unmatched (denoted by $\emptyset$) and colleges have strict preferences over students and being unmatched.
A simple theory of matching

- Proposed by Gale and Shapley (1962). Suggested reading is Roth and Sotomayor (RS henceforth), Chapter 2.
- Finite sets $S$ of students and $C$ of colleges (we use student-college terminology just for convenience).
- Each student can be matched to at most one college, and each college can admit at most one student (so the model is called “one-to-one matching”). Students have strict preferences over colleges and being unmatched (denoted by $\emptyset$) and colleges have strict preferences over students and being unmatched.
- $c \succ_s c'$ means “student $s$ strictly prefers college $c$ to college $c'$. $s \succ_c s'$ means “college $c$ strictly prefers student $s$ to student $s'$.
A simple theory of matching

- Proposed by Gale and Shapley (1962). Suggested reading is Roth and Sotomayor (RS henceforth), Chapter 2.
- Finite sets $S$ of students and $C$ of colleges (we use student-college terminology just for convenience).
- Each student can be matched to at most one college, and each college can admit at most one student (so the model is called “one-to-one matching”). Students have strict preferences over colleges and being unmatched (denoted by $\emptyset$) and colleges have strict preferences over students and being unmatched.
- $c \succ_s c'$ means “student $s$ strictly prefers college $c$ to college $c'$.
- $s \succ_c s'$ means “college $c$ strictly prefers student $s$ to student $s'$.
- If $i \succ_j \emptyset$ then we say $i$ is acceptable to $j$. 
The outcome of the matching market is a **matching**, which specifies which student attends which college.
The outcome of the matching market is a **matching**, which specifies which student attends which college.

Formally, \( \mu \) is a function from \( S \cup C \) to \( S \cup C \cup \{\emptyset\} \) such that

1. \( \mu(s) \in C \cup \{\emptyset\} \),
2. \( \mu(c) \in S \cup \{\emptyset\} \), and
3. \( \mu(s) = c \iff \mu(c) = s \), for every student \( s \in S \) and college \( c \in C \).
Roughly speaking, a matching is stable if there is no individual players or pairs of players who can profitably deviate from it. Formally,
Roughly speaking, a matching is stable if there is no individual players or pairs of players who can profitably deviate from it. Formally,

Matching $\mu$ is **blocked by an individual** $i$ if $\mu(i)$ is unacceptable to $i$, that is, $\emptyset \succ_i \mu(i)$.
Stability

- Roughly speaking, a matching is stable if there is no individual players or pairs of players who can profitably deviate from it. Formally,
- Matching $\mu$ is **blocked by an individual** $i$ if $\mu(i)$ is unacceptable to $i$, that is, $\emptyset \succ_i \mu(i)$.
- Matching $\mu$ is **blocked by a pair** $s$ and $c$ if each of them prefer each other to their partners under $\mu$, that is,

$$c \succ_s \mu(s), \text{ and } s \succ_c \mu(c).$$
Stability

- Roughly speaking, a matching is stable if there is no individual players or pairs of players who can profitably deviate from it. Formally,

- Matching $\mu$ is **blocked by an individual** $i$ if $\mu(i)$ is unacceptable to $i$, that is, $\emptyset \succ_i \mu(i)$.

- Matching $\mu$ is **blocked by a pair** $s$ and $c$ if each of them prefer each other to their partners under $\mu$, that is,

  \[ c \succ_s \mu(s), \text{ and } s \succ_c \mu(c). \]

- A matching is **stable** if it is not blocked by any individual or pair.
Roughly speaking, a matching is stable if there is no individual players or pairs of players who can profitably deviate from it. Formally,

- Matching $\mu$ is **blocked by an individual** $i$ if $\mu(i)$ is unacceptable to $i$, that is, $\emptyset \succ_i \mu(i)$.
- Matching $\mu$ is **blocked by a pair** $s$ and $c$ if each of them prefer each other to their partners under $\mu$, that is,

$$c \succ_s \mu(s), \text{ and } s \succ_c \mu(c).$$

A matching is **stable** if it is not blocked by any individual or pair.

(for some theorists: the set of all stable matchings is equivalent to the **core**, and a stable matching is **Pareto efficient**.)
Stable matchings always exist

Theorem (Gale and Shapley 1962; RS Theorem 2.8)
There exists a stable matching in any one-to-one matching market.
Stable matchings always exist

Theorem (Gale and Shapley 1962; RS Theorem 2.8)
There exists a stable matching in any one-to-one matching market.

- Gale and Shapley propose the (student-proposing) **deferred acceptance algorithm**:
Stable matchings always exist

Theorem (Gale and Shapley 1962; RS Theorem 2.8)

There exists a stable matching in any one-to-one matching market.

- Gale and Shapley propose the (student-proposing) **deferred acceptance algorithm**:
- Given preferences of students and colleges, conduct the following algorithm:

  **Step 1**: (a) Each student “applies” to her first choice college.
               (b) Each college tentatively holds the most preferred applicant
                   (if s/he is acceptable) and rejects all other students.

  **Step \( t \geq 2 \)**: (a) Each student rejected in Step \((t - 1)\) applies to her next highest choice.
                        (b) Each college considers both new applicants and the student (if any) held at Step \((t-1)\), tentatively holds the most preferred acceptable student from the combined set of students, and rejects all other students.

- Terminate when no more applications are made. Termination happens in finite time.
Example of DA algorithm

Let \( S = \{s_1, s_2, s_3\} \), \( C = \{c_1, c_2\} \), and their preferences given by

\[
\succ_{s_1} : c_1, c_2, \\
\succ_{s_2} : c_1, \\
\succ_{s_3} : c_2, c_1, \\
\succ_{c_1} : s_3, s_2, s_1, \\
\succ_{c_2} : s_1, s_3.
\]
Example of DA algorithm

- Let \( S = \{s_1, s_2, s_3\} \), \( C = \{c_1, c_2\} \), and their preferences given by

\[
\succ_{s_1} : c_1, c_2, \\
\succ_{s_2} : c_1, \\
\succ_{s_3} : c_2, c_1, \\
\succ_{c_1} : s_3, s_2, s_1, \\
\succ_{c_2} : s_1, s_3.
\]

- Follow steps of the DA algorithm (I recommend each of you to do it with a piece of paper).
Example of DA algorithm

- Let $S = \{s_1, s_2, s_3\}$, $C = \{c_1, c_2\}$, and their preferences given by

$$
\succ_{s_1} : c_1, c_2, \\
\succ_{s_2} : c_1, \\
\succ_{s_3} : c_2, c_1, \\
\succ_{c_1} : s_3, s_2, s_1, \\
\succ_{c_2} : s_1, s_3.
$$

- Follow steps of the DA algorithm (I recommend each of you to do it with a piece of paper).
- The resulting matching $\mu = \{(s_1, c_2), (s_2, \emptyset), (s_3, c_1)\}$ is stable (verify it!).
Proof of Theorem (A stable matching always exists)

The proof is very simple.
Proof of Theorem (A stable matching always exists)

The proof is very simple.

1. The resulting matching \( \mu \) of DA is not blocked by an individual because at each step of the algorithm, no student applies to an unacceptable college and no college holds application of an unacceptable student.
Proof of Theorem (A stable matching always exists)

The proof is very simple.

1. The resulting matching $\mu$ of DA is not blocked by an individual because at each step of the algorithm, no student applies to an unacceptable college and no college holds application of an unacceptable student.

2. $\mu$ is not blocked by any pair because: Suppose $c \succ_s \mu(s)$ for some $s$ and $c$. This means that $s$ applied to $c$ and was rejected by $c$ at some step of DA. Since $c$’s tentative match only improves as the algorithm proceeds, the match $\mu(c)$ at the end of DA is still better for $c$ than $s$. So $c$ is not interested in blocking $\mu$ with $s$. 
Mechanisms in real markets

1. Stability is theoretically appealing, but does it matter in real life?
Mechanisms in real markets

1. Stability is theoretically appealing, but does it matter in real life?
2. Roth (1984) showed that the NIMP algorithm is equivalent to a (hospital-proposing) DA algorithm, so NIMP produces a stable matching.
Mechanisms in real markets

1. Stability is theoretically appealing, but does it matter in real life?
2. Roth (1984) showed that the NIMP algorithm is equivalent to a (hospital-proposing) DA algorithm, so NIMP produces a stable matching.
3. Roth (1991) studied British medical match, where different regions use different matching mechanisms. He found that stable mechanisms are successfully used (and is still in use) but most unstable mechanisms were abandoned after a short period of time.
Mechanisms in real markets

1. Stability is theoretically appealing, but does it matter in real life?
2. Roth (1984) showed that the NIMP algorithm is equivalent to a (hospital-proposing) DA algorithm, so NIMP produces a stable matching.
3. Roth (1991) studied British medical match, where different regions use different matching mechanisms. He found that stable mechanisms are successfully used (and is still in use) but most unstable mechanisms were abandoned after a short period of time.
4. In school choice, stability means “no justified envy”: no student is placed in a less preferred school to another school where a student with lower priority is assigned. NYC and Boston recently adopted DA in order to, among other things, to eliminate such unfair assignment.
### Mechanisms in real markets

<table>
<thead>
<tr>
<th>Market</th>
<th>Stable</th>
<th>Still in use</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRMP</td>
<td>yes</td>
<td>yes (new design 98-)</td>
</tr>
<tr>
<td>Edinburgh ('69)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Cardiff</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Birmingham</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Edinburgh ('67)</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Newcastle</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Sheffield</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Cambridge</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>London Hospital</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Medical Specialties</td>
<td>yes</td>
<td>yes (1/30 no)</td>
</tr>
<tr>
<td>Canadian Lawyers</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Dental Residencies</td>
<td>yes</td>
<td>yes (2/7 no)</td>
</tr>
<tr>
<td>Reform rabbis</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>NYC highschool</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
Student/College-optimal stable matchings

Theorem (Gale and Shapley 1962; RS Theorem 2.12)

There exists a student-optimal stable matching, that is, a stable matching that every student weakly prefers to any stable matching.
Theorem (Gale and Shapley 1962; RS Theorem 2.12)
There exists a student-optimal stable matching, that is, a stable matching that every student weakly prefers to any stable matching. The result of the student-proposing DA algorithm is the student-optimal stable matching. Similarly, college-proposing DA algorithm results in the college-optimal stable matching.
Student/College-optimal stable matchings

Theorem (Gale and Shapley 1962; RS Theorem 2.12)

There exists a student-optimal stable matching, that is, a stable matching that every student weakly prefers to any stable matching. The result of the student-proposing DA algorithm is the student-optimal stable matching. Similarly, college-proposing DA algorithm results in the college-optimal stable matching.

Moreover, the student-optimal stable matching is college-pessimal, that is, every college weakly dispreferences it to any stable matching, and vice versa (Theorem 2.13 of RS, but try to prove yourself as this is an easy exercise!)
The Theorem says that different stable matchings may benefit different market participants. In particular, each version of DA favors one side of the market at the expense of the other side.
The Theorem says that different stable matchings may benefit different market participants. In particular, each version of DA favors one side of the market at the expense of the other side.

This point was part of policy debate in NRMP in the 1990s. Recall that previous NIMP algorithm was hospital-proposing. Some medical students argued that the system favors hospitals at the expense of students and called for reconsideration of the mechanism.
The Theorem says that different stable matchings may benefit different market participants. In particular, each version of DA favors one side of the market at the expense of the other side.

This point was part of policy debate in NRMP in the 1990s. Recall that previous NIMP algorithm was hospital-proposing. Some medical students argued that the system favors hospitals at the expense of students and called for reconsideration of the mechanism.

We will come back to this point in a future lecture and discuss how important this is in the context of NRMP medical match.
Proof of Theorem

Terminology: \( c \) is **achievable** for \( s \) if there is some stable matching \( \mu \) such that \( \mu(s) = c \). It suffices to show that no student is rejected by an achievable college in any step of DA.
Terminology: $c$ is **achievable** for $s$ if there is some stable matching $\mu$ such that $\mu(s) = c$. It suffices to show that no student is rejected by an achievable college in any step of DA.

For contradiction, suppose a student is rejected by an achievable college. Consider the first step in which a student, say $s$, is rejected by an achievable college, say $c$ (let $\mu$ be a stable matching where $\mu(s) = c$.) This means that some other student $s'$ applied to $c$ and replaced the seat at $c$ at this step. Since this is the first step of DA where a student is rejected by an achievable college, we have $c \succ_{s'} \mu(s')$. Also we have $s' \succ_c s$ since $s'$ displaces $s$ at $c$ in DA. This means that pair $(s', c)$ blocks $\mu$, contradicting stability of $\mu$. 
The "Rural Hospital Theorem" (RS Theorem 2.22)

Rural Hospital Theorem (RS Theorem 2.22)
The set of students and colleges that are unmatched is the same for all stable matchings.
The “Rural Hospital Theorem” (RS Theorem 2.22)

Rural Hospital Theorem (RS Theorem 2.22)

The set of students and colleges that are unmatched is the same for all stable matchings.

- One motivation is the allocation of residents in rural hospitals. Hospitals in rural areas cannot fill positions for residents, and some people argue that the matching mechanisms should be changed so that more doctors end up in rural hospitals.
The “Rural Hospital Theorem” (RS Theorem 2.22)

Rural Hospital Theorem (RS Theorem 2.22)

The set of students and colleges that are unmatched is the same for all stable matchings.

- One motivation is the allocation of residents in rural hospitals. Hospitals in rural areas cannot fill positions for residents, and some people argue that the matching mechanisms should be changed so that more doctors end up in rural hospitals.
- But the theorem says that it is impossible as long as stable matching is implemented.
The “Rural Hospital Theorem” (RS Theorem 2.22)

Rural Hospital Theorem (RS Theorem 2.22)

The set of students and colleges that are unmatched is the same for all stable matchings.

- One motivation is the allocation of residents in rural hospitals. Hospitals in rural areas cannot fill positions for residents, and some people argue that the matching mechanisms should be changed so that more doctors end up in rural hospitals.
- But the theorem says that it is impossible as long as stable matching is implemented.
- Also, if some students are matched in some stable matching and not in others, the latter may be unfair to him/her. The theorem says that there is not need to worry.
The “Rural Hospital Theorem” (RS Theorem 2.22)

Rural Hospital Theorem (RS Theorem 2.22)
The set of students and colleges that are unmatched is the same for all stable matchings.

- One motivation is the allocation of residents in rural hospitals. Hospitals in rural areas cannot fill positions for residents, and some people argue that the matching mechanisms should be changed so that more doctors end up in rural hospitals.
- But the theorem says that it is impossible as long as stable matching is implemented.
- Also, if some students are matched in some stable matching and not in others, the latter may be unfair to him/her. The theorem says that there is not need to worry.
- In some markets, not all assumptions hold exactly, so the theorem does not hold exactly. Then it is important to know if the theorem holds approximately. I will come back to this topic in the context of NRMP in 1990.
Proof of Rural Hospital Theorem

Let $\mu^S$ be the student-optimal stable matching and $\mu$ be an arbitrary stable matching.
Proof of Rural Hospital Theorem

- Let $\mu^S$ be the student-optimal stable matching and $\mu$ be an arbitrary stable matching.
- Since $\mu^S$ is student-optimal, all the students that are matched in $\mu$ are matched in $\mu^S$. 
Proof of Rural Hospital Theorem

- Let $\mu^S$ be the student-optimal stable matching and $\mu$ be an arbitrary stable matching.
- Since $\mu^S$ is student-optimal, all the students that are matched in $\mu$ are matched in $\mu^S$.
- Since $\mu^S$ is college-pessimal, all the colleges that are matched in $\mu^S$ are matched in $\mu$. 
Proof of Rural Hospital Theorem

Let $\mu^S$ be the student-optimal stable matching and $\mu$ be an arbitrary stable matching.

Since $\mu^S$ is student-optimal, all the students that are matched in $\mu$ are matched in $\mu^S$.

Since $\mu^S$ is college-pessimal, all the colleges that are matched in $\mu^S$ are matched in $\mu$.

But the number of matched students and colleges are the same in any matching. This means that the same set of students and colleges are matched in $\mu^S$ and $\mu$. 

We have learned properties of stable matching, given information about preferences of market participants.
Strategic behavior (RS Chapter 4)

- We have learned properties of stable matching, given information about preferences of market participants.
- But in reality, such information is private, so the clearinghouse should ask participants.
Strategic behavior (RS Chapter 4)

- We have learned properties of stable matching, given information about preferences of market participants.
- But in reality, such information is private, so the clearinghouse should ask participants.
- Do people have incentive to tell the truth?
Strategic behavior: terminology

- A mechanism is a rule that produces a matching for any reported preference.
Strategic behavior: terminology

- A **mechanism** is a rule that produces a matching for any reported preference.
- DA is an example of a mechanism.
Strategic behavior: terminology

- A **mechanism** is a rule that produces a matching for any reported preference.
- DA is an example of a mechanism.
- A mechanism is **strategy-proof** if telling the true preferences is a **dominant strategy** (that is, a best action no matter what others do) for everyone.
DA is not strategy-proof

- Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and

  $\succ_{s_1} : c_1, c_2,$
  $\succ_{s_2} : c_2, c_1,$
  $\succ_{c_1} : s_2, s_1$
  $\succ_{c_2} : s_1, s_2.$
DA is not strategy-proof

Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and

$$
\succ_s s_1 : c_1, c_2,
\succ_s s_2 : c_2, c_1,
\succ_c c_1 : s_2, s_1
\succ_c c_2 : s_1, s_2.
$$

When everyone reports true preferences, DA produces
DA is not strategy-proof

- Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and 
\[
\succ_{s_1} : c_1, c_2,
\succ_{s_2} : c_2, c_1,
\succ_{c_1} : s_2, s_1
\succ_{c_2} : s_1, s_2.
\]
- When everyone reports true preferences, DA produces 
$\mu = \{(s_1, c_1), (s_2, c_2)\}$. 

When $c_1$ reports $\succ':_{c_1} : s_2$, then DA produces 
$\mu' = \{(s_1, c_2), (s_2, c_1)\}$, which $c_1$ prefers to $\mu(c_1) = s_1$. 
So DA is not strategy-proof.
DA is not strategy-proof

- Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and

$$
\succ_{s_1} : c_1, c_2, \\
\succ_{s_2} : c_2, c_1, \\
\succ_{c_1} : s_2, s_1, \\
\succ_{c_2} : s_1, s_2.
$$

- When everyone reports true preferences, DA produces

$$
\mu = \{(s_1, c_1), (s_2, c_2)\}.
$$

- When $c_1$ reports $\succ_{c_1}' : s_2$, then DA produces

$$
\mu' = \{(s_1, c_2), (s_2, c_1)\},
$$

which $c_1$ prefers to $\mu(c_1) = s_1$. 

So DA is not strategy-proof.
DA is not strategy-proof

- Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and
  
  $\succsim_{s_1} : c_1, c_2,$
  $\succsim_{s_2} : c_2, c_1,$
  $\succsim_{c_1} : s_2, s_1$
  $\succsim_{c_2} : s_1, s_2.$

- When everyone reports true preferences, DA produces
  $\mu = \{(s_1, c_1), (s_2, c_2)\}.$

- When $c_1$ reports $\succsim_{c_1} : s_2$, then DA produces
  $\mu' = \{(s_1, c_2), (s_2, c_1)\},$ which $c_1$ prefers to $\mu(c_1) = s_1.$
DA is not strategy-proof

- Let $S = \{s_1, s_2\}$, $C = \{c_1, c_2\}$ and
  
  \[
  \succ_{s_1} : c_1, c_2,
  \succ_{s_2} : c_2, c_1,
  \succ_{c_1} : s_2, s_1
  \succ_{c_2} : s_1, s_2.
  \]

- When everyone reports true preferences, DA produces
  \[
  \mu = \{(s_1, c_1), (s_2, c_2)\}.
  \]

- When $c_1$ reports $\succ_{c_1}' : s_2$, then DA produces
  \[
  \mu' = \{(s_1, c_2), (s_2, c_1)\},
  \] which $c_1$ prefers to $\mu(c_1) = s_1$.

- So DA is not strategy-proof.
Impossibility Theorem

- DA is not strategy-proof, so people may have incentives to manipulate the mechanism.
Impossibility Theorem

- DA is not strategy-proof, so people may have incentives to manipulate the mechanism.
- Unfortunately, we cannot overcome the difficulty by finding another mechanism.
Impossibility Theorem

- DA is not strategy-proof, so people may have incentives to manipulate the mechanism.
- Unfortunately, we cannot overcome the difficulty by finding another mechanism.

Impossibility Theorem (Roth 1982; RS Theorem 4.4)

There is no stable mechanism that is strategy-proof.
Impossibility Theorem

- DA is not strategy-proof, so people may have incentives to manipulate the mechanism.
- Unfortunately, we cannot overcome the difficulty by finding another mechanism.

Impossibility Theorem (Roth 1982; RS Theorem 4.4)

There is no stable mechanism that is strategy-proof.

- Proof is a modification of the last slide (proof is available in RS page 88, but it’s a good idea to try to prove it yourself.)
Impossibility Theorem

- DA is not strategy-proof, so people may have incentives to manipulate the mechanism.
- Unfortunately, we cannot overcome the difficulty by finding another mechanism.

Impossibility Theorem (Roth 1982; RS Theorem 4.4)

There is no stable mechanism that is strategy-proof.

- Proof is a modification of the last slide (proof is available in RS page 88, but it’s a good idea to try to prove it yourself.)
- As before, it is still important to study whether manipulation is likely under stable mechanisms in applications. This will be the subject in a future lecture.
DA is strategy-proof for one side

Theorem (Dubins and Freedman 1981, Roth 1982; RS Thm 4.7)

*The student-proposing DA is strategy-proof for students. That is, telling the truth is a dominant strategy for every student.*
Theorem (Dubins and Freedman 1981, Roth 1982; RS Thm 4.7)

The student-proposing DA is strategy-proof for students. That is, telling the truth is a dominant strategy for every student.

- Actually it is group strategy-proof for students. That is, even a group of students can tell a lie together and make every member of the group strictly better off. See Hatfield and Kojima (forthcoming, Games and Economic Behavior) for the most general result.
DA is strategy-proof for one side

Theorem (Dubins and Freedman 1981, Roth 1982; RS Thm 4.7)

The student-proposing DA is strategy-proof for students. That is, telling the truth is a dominant strategy for every student.

- Actually it is **group strategy-proof** for students. That is, even a group of students can tell a lie together and make every member of the group strictly better off. See Hatfield and Kojima (forthcoming, Games and Economic Behavior) for the most general result.
- Proof is skipped (Intuition: students are not punished by applying to preferred colleges (this is in a contrast with the “Boston mechanism”).
Many-to one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
Many-to one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except
Advance the theory to account for colleges with multiple positions.

Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
Many-to one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
- Matching $\mu$ is a correspondence from $S \cup C$ to $S \cup C \cup \{\emptyset\}$ such that
Many-to one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
- Matching $\mu$ is a correspondence from $S \cup C$ to $S \cup C \cup \{\emptyset\}$ such that
  1. $\mu(s) \in C \cup \{\emptyset\}$. 

Stable matchings always exist in many-to-one matching
Other properties
Married Couples
Many-to-one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
- Matching $\mu$ is a correspondence from $S \cup C$ to $S \cup C \cup \{\emptyset\}$ such that
  1. $\mu(s) \in C \cup \{\emptyset\}$,
  2. $\mu(c) \subseteq S$ (each college is matched to a group of students), and
Many-to one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
- Matching $\mu$ is a correspondence from $S \cup C$ to $S \cup C \cup \{\emptyset\}$ such that
  1. $\mu(s) \in C \cup \{\emptyset\}$,
  2. $\mu(c) \subseteq S$ (each college is matched to a group of students), and
  3. $\mu(s) = c \iff s \in \mu(c)$, for every student $s \in S$ and college $c \in C$. 

Stability of a matching is defined similarly.
Many-to-one matching (RS Chapter 5)

- Advance the theory to account for colleges with multiple positions.
- Everything is the same as before (Finite sets $S$ of students and $C$ of colleges etc) except each college $c$ has $q_c$ positions to fill.
- Matching $\mu$ is a correspondence from $S \cup C$ to $S \cup C \cup \{\emptyset\}$ such that
  1. $\mu(s) \in C \cup \{\emptyset\}$,
  2. $\mu(c) \subseteq S$ (each college is matched to a group of students), and
  3. $\mu(s) = c \iff s \in \mu(c)$, for every student $s \in S$ and college $c \in C$.

- Stability of a matching is defined similarly.
Stable matchings always exist in many-to-one matching

Theorem (Gale and Shapley 1962; RS Lemma 5.6)

There exists a stable matching in any one-to-one matching market.
Stable matchings always exist in many-to-one matching

Theorem (Gale and Shapley 1962; RS Lemma 5.6)
There exists a stable matching in any one-to-one matching market.

- One easy proof: think of a college $c$ as $q_c$ different colleges with one position each. Then, the theorem for one-to-one matching applies.
Overview

Simple Theory: One-to-one matching
Theory: Many-to-one matching

Summary

Stable matchings always exist in many-to-one matching

Other properties

Married Couples

Stable matchings always exist in many-to-one matching

Theorem (Gale and Shapley 1962; RS Lemma 5.6)

There exists a stable matching in any one-to-one matching market.

- One easy proof: think of a college $c$ as $q_c$ different colleges with one position each. Then, the theorem for one-to-one matching applies.
- Or we could directly generalize the (student-proposing) DA:

**Step 1**

(a) Each student “applies” to her first choice college.
(b) Each college tentatively holds the most preferred applicants up to its quota (if s/he is acceptable) and rejects all other students.

**Step $t \geq 2$**

(a) Each student rejected in Step $(t - 1)$ applies to her next highest choice.
(b) Each college considers both new applicants and the student (if any) held at Step $(t-1)$, tentatively holds the most preferred acceptable students up to its quota from the combined set of students, and rejects all other students.

- Terminate when no more applications are made. Termination happens in finite time.
Stable matchings always exist in many-to-one matching

Theorem (Gale and Shapley 1962; RS Lemma 5.6)

There exists a stable matching in any one-to-one matching market.

- One easy proof: think of a college \( c \) as \( q_c \) different colleges with one position each. Then, the theorem for one-to-one matching applies.
- Or we could directly generalize the (student-proposing) DA:

  **Step 1**: (a) Each student “applies” to her first choice college.
  (b) Each college tentatively holds the most preferred applicants **up to its quota** (if s/he is acceptable) and rejects all other students.

  **Step \( t \geq 2 \)**: (a) Each student rejected in Step \((t-1)\) applies to her next highest choice.
  (b) Each college considers both new applicants and the student (if any) held at Step \( (t-1) \), tentatively holds the most preferred acceptable students **up to its quota** from the combined set of students, and rejects all other students.

- Terminate when no more applications are made. Termination happens in finite time.

- Proof that DA results in a stable matching is essentially the same (good exercise!)
Many properties carry over to many-to-one matching
Many properties carry over to many-to-one matching

Because we can think of each college $c$ as $q_c$ different colleges with one position, many theories of one-to-one matching carry over to many-to-one matching (so one-to-one matching theory was useful after all!). Examples:
Many properties carry over to many-to-one matching

- Because we can think of each college $c$ as $q_c$ different colleges with one position, many theories of one-to-one matching carry over to many-to-one matching (so one-to-one matching theory was useful after all!). Examples:
  1. Student/college-proposing DA result in the student/college-optimal stable matchings.
Many properties carry over to many-to-one matching

Because we can think of each college $c$ as $q_c$ different colleges with one position, many theories of one-to-one matching carry over to many-to-one matching (so one-to-one matching theory was useful after all!). Examples:

1. Student/college-proposing DA result in the student/college-optimal stable matchings.
2. Rural hospital theorem: all colleges fill the same number of positions across stable matchings. Any student unmatched in any one stable matching is unmatched in all stable matching.
Some properties fail in many-to-one matching

- Not all properties carry over to many-to-one matching, especially strategic properties.
Some properties fail in many-to-one matching

- Not all properties carry over to many-to-one matching, especially strategic properties.
  
  1. No stable mechanism is strategy-proof for colleges (RS; Theorem 5.14). In particular, even college-proposing DA is not strategy-proof for colleges.
Some properties fail in many-to-one matching

- Not all properties carry over to many-to-one matching, especially strategic properties.
  1. No stable mechanism is strategy-proof for colleges (RS; Theorem 5.14). In particular, even college-proposing DA is not strategy-proof for colleges.
  2. On the contrary, student-proposing DA is still strategy-proof for students.
Some properties fail in many-to-one matching

- Not all properties carry over to many-to-one matching, especially strategic properties.
  1. No stable mechanism is strategy-proof for colleges (RS; Theorem 5.14). In particular, even college-proposing DA is not strategy-proof for colleges.
  2. On the contrary, student-proposing DA is still strategy-proof for students.
  3. Colleges may benefit by misreporting capacities. Sonmez (1997) shows that there is no stable mechanism that is immune to manipulation via capacities. Subsequent papers such as Konishi and Unver (2006) and Kojima (2007) identify conditions under which DA's are immune to manipulation via capacities.
Married Couples (RS section 5.4.3)

There are many married couples in medical match (1,000 out of 20,000 in NRMP, 1990s; 30-40 out of 3,000 in psychologist match, 2000s.), and they usually want to work in the same city.

DA fails to accommodate couples: it may assign the husband to Boston, the wife to LA, for example.

Participation of medical students in NIMP dropped in 1970s, especially among couples.
Married Couples (RS section 5.4.3)

There are many married couples in medical match (1,000 out of 20,000 in NRMP, 1990s; 30-40 out of 3,000 in psychologist match, 2000s.), and they usually want to work in the same city.
Married Couples (RS section 5.4.3)

- There are many married couples in medical match (1,000 out of 20,000 in NRMP, 1990s; 30-40 out of 3,000 in psychologist match, 2000s.), and they usually want to work in the same city.
- DA fails to accommodate couples: it may assign the husband to Boston, the wife to LA, for example.
Married Couples (RS section 5.4.3)

- There are many married couples in medical match (1,000 out of 20,000 in NRMP, 1990s; 30-40 out of 3,000 in psychologist match, 2000s.), and they usually want to work in the same city.
- DA fails to accommodate couples: it may assign the husband to Boston, the wife to LA, for example.
- Participation of medical students in NIMP dropped in 1970s, especially among couples.
There may be no stable matching with couples

- There are $C = \{c_1, c_2\}$ and one single student $s$ and one couple $(m, w)$

\[ s \succ c_1, c_2, \]
\[ (m, w) \succ (c_1, c_2), \]
\[ c_1 \succ m, s \]
\[ c_2 \succ s, w. \]
There may be no stable matching with couples

- There are $C = \{c_1, c_2\}$ and one single student $s$ and one couple $(m, w)$

  $\succ s : c_1, c_2,$

  $\succ (m, w) : (c_1, c_2),$  

  $\succ c_1 : m, s$  

  $\succ c_2 : s, w.$

- There is no stable matching (exercise).
There may be no stable matching with couples

There are $C = \{c_1, c_2\}$ and one single student $s$ and one couple $(m, w)$

$s \succ c_1, c_2,$

$(m, w) \succ (c_1, c_2),$  

$c_1 \succ m, s$

$c_2 \succ s, w.$

There is no stable matching (exercise).

So, the problem is “impossible to solve” in a sense. Then, what should we do?
Stability is important for matching in labor markets. Theoretically, 1) DA produces a stable matching if the market is simple (no couples etc). 2) Depending on which DA to use (student or college proposing), one side benefits at the expense of the other but the set of matched colleges and students do not change. 3) DA is not strategy-proof. 4) With couples, stable mechanisms may not work. Next we look at the real market and see if these theories can (or cannot) guide design of the market institution.
Stability is important for matching in labor markets.
Summary

- Stability is important for matching in labor markets.
- Theoretically,
Summary

- Stability is important for matching in labor markets.
- Theoretically,
  - DA produces a stable matching if the market is simple (no couples etc).

1. DA produces a stable matching if the market is simple (no couples etc).
Summary

- Stability is important for matching in labor markets.
- Theoretically,
  1. DA produces a stable matching if the market is simple (no couples etc).
  2. Depending on which DA to use (student or college proposing), one side benefits at the expense of the other but the set of matched colleges and students do not change.
Stability is important for matching in labor markets.

Theoretically,

1. DA produces a stable matching if the market is simple (no couples etc).
2. Depending on which DA to use (student or college proposing), one side benefits at the expense of the other but the set of matched colleges and students do not change.
3. DA is not strategy-proof.
Summary

- Stability is important for matching in labor markets.
- Theoretically,
  1. DA produces a stable matching if the market is simple (no couples etc).
  2. Depending on which DA to use (student or college proposing), one side benefits at the expense of the other but the set of matched colleges and students do not change.
  3. DA is not strategy-proof.
  4. With couples, stable mechanisms may not work.
Summary

- Stability is important for matching in labor markets.
- Theoretically,
  1. DA produces a stable matching if the market is simple (no couples etc).
  2. Depending on which DA to use (student or college proposing), one side benefits at the expense of the other but the set of matched colleges and students do not change.
  3. DA is not strategy-proof.
  4. With couples, stable mechanisms may not work.
- Next we look at the real market and see if these theories can (or cannot) guide design of the market institution.
Reading for next class:
