Lecture 1: The basic matching model

Thilo Klein

ZEW Mannheim

Introduction

Economics studies how goods and services are exchanged or distributed via a market.

Traditionally, markets determine allocation through prices, which are sufficient statistics to determine who gets what.

But in some cases prices may not be enough to characterize allocations (e.g., College admission, labor market, etc.)

$\Rightarrow$ It's also about choosing and being chosen.

Let's consider the extreme case where there's no price (i.e., no monetary transaction).

Matching markets are typical examples of markets where there is no monetary transactions (and thus no price):

- School assignment
- Medical match
- Allocation of dorms
- Assignment of cadets to branches
- Organ transplants
- Allocation of subsidized/public housing
- etc.

Warning

The next slides will present a model that has been formulated in the 1960's.

There is absolutely no intention to:

- promote an outdated view;
- hurt people's feelings;
- discriminate;
- claim that the model captures the situation it describes.

Marriage market

The marriage market model is the baseline model for most matching market models:

- There is a finite set of women: $W = \{w_1, w_2, \dots\}$.
- There is a finite set of men: $M = \{m_1, m_2, \dots\}$.

This model is known as a two-sided, one-to-one matching model. Other models are:

- One-sided matching (e.g., matching roommates);
- Many-to-one matching (matching students to colleges);
- Many-to-many matching (matching students and professors).

Preferences

- each man m has a preference ordering P_m over the women and the option of remaining single.
- each woman w has a preference ordering P_w over the men and the option of remaining single.

Example $P_m: \; w_1, w_3, w_6, m, w_2, w_4, \ldots$

- m's most preferred woman is w_1, then w_3 is the second woman, etc.
- m prefers to be matched to himself (stay single) than being matched to w_2 or w_4.
- w_2 and w_4 are unacceptable for m.

Preferences are assumed to be strict: nobody is indifferent between two different options.

Other ways to represent preferences:

- Read from left to right:

$$w_1 P_m w_3$$

  Man m (strictly) prefers w_1 to w_3.
- Read from top to bottom:

$$\begin{array}{c} P_m \\ \hline w_1 \\ w_3 \\ w_6 \end{array}$$

Matching

A matching is a mapping μ that says who is matched to whom:

- $\mu(m)$ is the partner of m under the matching μ
- $\mu(w)$ is the partner of w under the matching μ

Any matching μ must satisfy the following properties:

- For each man m, $\mu(m) \in W \cup \{m\}$
  A man m is matched to a woman or himself (but not to other men).
- For each woman w, $\mu(w) \in M \cup \{w\}$
  A woman w is matched to a man or herself (but not to other women).
- $\mu(m) = w \Leftrightarrow \mu(w) = m$.
  A man is matched to a women if, and only if, that woman is matched to that man.

Stability

There's no price in a matching problem, so we can't really talk about equilibrium. The "equivalent" concept is stability. It's a conjunction of two requirements: *individual rationality* and *absence of blocking pairs*.

Definition

A matching μ is individually rational if for each individual $v \in M \cup W$,

$$\mu(v) \; R_v \; v$$

R_v is the relation "preferred or indifferent to".

In words, a matching is individually rational if nobody would strictly prefer to remain single than staying with the partner prescribed by the matching.

Definition

A pair (m, w) block a matching μ if

- $\mu(m) \neq w$ *m and w are not matched together under* μ
- $w \; P_m \; \mu(m)$ *m prefers w to his match*
- $m \; P_w \; \mu(w)$ *w prefers m to her match*

Definition

A matching μ is stable if

- it is individually rational;
- there is no pair man-woman that blocks μ.

Example

P_{m_1}	P_{m_2}
w_1	$\boxed{w_1}$
$\boxed{w_2}$	w_2
m_1	m_2

P_{w_1}	P_{w_2}
m_1	$\boxed{m_1}$
$\boxed{m_2}$	m_2
w_1	w_2

Consider the matching $\mu(m_1) = w_2$ and $\mu(m_2) = w_1$.

That matching is not stable: m_1 and w_1 block μ:

- m_1 prefers w_1 to his match, $\mu(m_1) = w_2$
- w_1 prefers m_1 to her match, $\mu(w_1) = m_2$.

Example

P_{m_1}	P_{m_2}
$\boxed{w_1}$	w_1
w_2	$\boxed{w_2}$
m_1	m_2

P_{w_1}	P_{w_2}
$\boxed{m_1}$	m_1
m_2	$\boxed{m_2}$
w_1	w_2

Consider the matching $\mu'(m_1) = w_1$ and $\mu'(m_2) = w_2$. That matching is stable:

- m_2 would like to block with w_1 but she prefers $\mu'(w_1) = m_1$ to him.
- w_2 would like to block with m_1 but he prefers $\mu'(m_1) = w_1$ to her.
- w_1 and m_1 do not want to block: they are matched to their most preferred partner.

Theorem (David Gale & Lloyd Shapley, 1962)

For any preferences, there always exists at least one stable matching.

D. Gale and L. Shapley (1963) "College admissions and the stability of marriage," *American Mathematical Monthly*, vol. 69, pp. 9–15.

Deferred Acceptance algorithm (informal)

It is like "modern" dating:

- Men propose women in the same order as in their preferences:
  - Each man will start proposing his most preferred woman
  - If rejected, a man will propose his second most preferred woman
  - If rejected, a man will propose his third most preferred woman
  - etc.
- Each woman always keeps the best man (according to her preferences) among the man proposing her (if any), and rejects the others.
- The algorithm stops when there's no more rejection.

P_{m_1}	P_{m_2}
w_1	w_1
w_2	w_2
m_1	m_2

P_{w_1}	P_{w_2}
m_1	m_1
m_2	m_2
w_1	w_2

- **Step 1**: Men m_1 and m_2 both make an offer to w_1.
  - Woman w_1 picks her best choice, m_1, and rejects m_2.
  - Woman w_2 does nothing.
- **Step 2**: Man still courts m_1, m_2 now makes an offer to w_2.
  - Women w_1 is still with m_1.
  - Woman w_2 accepts m_2's proposal.
- There is no more rejection, so the algorithm stops.
  Final matching: $\mu(m_1) = w_1$ and $\mu(m_2) = w_2$.

Which side makes offers?

In the Deferred Acceptance algorithm it is important that:

- One side makes the offers, the other accept/rejects the proposal.
- No obligations to have men proposing, it can be women proposing.

Deferred Acceptance

Step 1

Each man proposes to his most preferred, acceptable woman

(if a man finds all women unacceptable he remains single).

Each woman who received at least one offer

- temporarily holds the offer from the most preferred man among those who made an offer to her and are acceptable.
- rejects the other offer(s).

Step $\mathbf{k}, \mathbf{k} \geq \mathbf{2}$

Each man whose offer has been rejected in the previous step proposes to his most preferred woman among the acceptable women he has not yet proposed.

(if there is no such woman he remains single).

Each woman who received at least one offer in this step

- temporarily holds the offer from the most preferred man among
  - those who made an offer to her in this step and are acceptable.
  - the man she held from the previous step (if any).
- rejects the other offer(s).

End: The algorithm stops when no man has an offer that is rejected.

Final matching:

- Each woman is matched to the man whose offer she was holding when the algorithm stopped (if any).

  That's why (final) acceptance was **deferred**

- Each man is matched to the woman he was temporarily matched when the algorithm stoped (if any).

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
w_2	w_1	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
w_2	w_1	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
m_2, m_3

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
w_2	$\cancel{w_1}$	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, m_3

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
w_2	$\cancel{w_1}$	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

$$\frac{w_1}{\cancel{m_2},\ m_3}$$

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, m_3

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	w_1
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, m_3

m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

$$\frac{w_1}{\cancel{m_2},\ \cancel{m_3}}$$

m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	w_2	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, $\cancel{m_3}$
m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	$\cancel{w_2}$	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, $\cancel{m_3}$
m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	$\cancel{w_2}$	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, $\cancel{m_3}$
m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	$\cancel{w_2}$	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, $\cancel{m_3}$
m_1

Deferred Acceptance

P_{m_1}	P_{m_2}	P_{m_3}
$\cancel{w_2}$	$\cancel{w_1}$	$\cancel{w_1}$
w_1	$\cancel{w_2}$	w_2
w_3	w_3	w_3

P_{w_1}	P_{w_2}	P_{w_3}
m_1	m_3	m_1
m_3	m_2	m_3
m_2	m_1	m_2

w_1
$\cancel{m_2}$, $\cancel{m_3}$
m_1
m_1

Proof of Gale and Shapley's theorem

The outcome of the Deferred Acceptance algorithm is necessarily individually rational:

- No man makes an offer to a woman he finds unacceptable;
- No woman accepts an offer from an unacceptable man.

$\Rightarrow$ If the outcome of the algorithm is not stable then there exists a blocking pair.

Suppose that $(\text{Alice}, \text{Bob})$ is a blocking pair:

- Alice prefers Bob to her match.
- Bob prefers Alice to his match.

  $\Rightarrow$ In the algorithm, Bob proposed to Alice before proposing to his final match.

If Bob made further proposals (after proposing Alice), he must have been rejected by Alice.

$\Rightarrow$ Alice got an offer from a man X she prefers to Bob.

$\Rightarrow$ Alice preferes her final match to Bob.

(Her final match is man X or a more preferred man.)

$\Rightarrow$ Alice and Bob cannot be a blocking pair, a contradiction.

Optimality

Denote by μ_M be the matching obtained with the DA algorithm with men proposing.

Proposition

Each man prefers μ_M to any other stable matching. Each woman prefers any stable matching to μ_M.

μ_M is called the man-optimal matching.

Since the model is symmetric between men and woman:

Proposition

Each woman prefers μ_W, the woman-optimal matching, to any other stable matching, and each man prefers any stable matching to μ_W.

- Matching obtain when running DA with men proposing:

  Man-optimal matching.

- Matching obtain when running DA with women proposing:

  Woman-optimal matching.

Proof: μ_M = best stable matching for men

Definition

Man m and woman w are achievable if there exists a stable matching μ such that $\mu(m) = w$.

To prove the result it is enough to show this:

Claim

Under DA (men proposing) no man can be rejected by an achievable woman.

> $\Rightarrow$ All women preferred to the man-optimal mate are not achievable.
>
> $\Rightarrow$ The man-optimal mate is the most preferred achievable partner.

- Suppose there exist an achievable pair, but in DA with men proposing the woman rejects the man.
- Let m be the first man rejected by an achievable woman (i.e., didn't occur any any earlier step).

  Let w be that woman.
- w rejects m because she prefers another man m'

  (m achievable $\Rightarrow$ m acceptable for w)

$$m' \; P_w \; m$$

  $\Rightarrow m'$ propose to w before (or at the same step) m is rejected by w.

- All women m' prefers to w are not achievable for m'.

  (because m first man rejected by an achievable)

  $\Rightarrow$ m' prefers (weakly) w to any other achievable woman.

- Take μ stable with $\mu(m) = w$.

  $\Rightarrow$ So m' matched to a woman less preferred than w (the women preferred to w are unachievable).

$$w \; P_{m'} \; \mu(m')$$

- So (m', w) block $\mu \quad \Rightarrow \quad \mu$ not stable, a contradiction.

Proof: men and women have opposite preferences (over stable matchings)

We prove the following:

Proposition

Let μ and μ' be two stable matchings. Suppose all men (weakly) prefer μ to μ'.
Then all women (weakly) prefer μ' to μ.

Suppose the proposition is not true:

There exists a woman *w* who also prefers μ to μ'.

Let $m = \mu(w)$ and $m' = \mu'(w)$.

So we have

$$m \; P_w \; m'$$

Since m and m' are two different men, man m is matched to a different woman under μ', i.e., $\mu(m) \neq \mu'(m)$.

So we have

$$w = \mu(m) \; P_m \; \mu'(m).$$

$\Rightarrow$ (m, w) block μ', so μ' is not stable, a contradiction.

Incentives with the Deferred Acceptance algorithm

So far, we assumed that the algorithm was using individuals' true preferences.

We consider the following mechanism:

1. Men and woman submit (simultaneously) their preferences;
2. A matching is constructed using the Deferred Acceptance algorithm and the submitted preferences.
3. The matching is announced.

Question: With this mechanism, do men and women have any incentive to submit their true preferences?

Theorem

A matching mechanism that uses the Deferred Acceptance algorithm is strategyproof for the proposing side

(i.e., it is a dominant strategy to submit one's true preferences).

Proof

The proof needs the following result:

Lemma (Blocking lemma)

Let μ be any matching and let $\widehat{M} \subset M$ be the set of men who prefer μ to μ_M. Then there is a pair (m, w) which blocks μ, where $m \notin \widehat{M}$.

The result we will prove is in fact stronger:

Theorem

Let P be any preference profile (for men and women). Then there is no set of men $\widehat{M}$ and preference profile $\widehat{P}'_{\widehat{M}}$ such that all men in $\widehat{M}$ prefer man-optimal matching with the preference profile $\widehat{P} = (P_{-M}, \widehat{P}'_{\widehat{M}}, P_W)$ to the true preference profile $P = (P_{-M}, P_{\widehat{M}}, P_W)$. (DA is group-strategyproof for the proposing side.)

Define $\widehat{\mu}_M =$ man-optimal with $\widehat{P}$.

Suppose each man in $\widehat{M}$ prefers $\widehat{\mu}_M$ to μ_M.

By the *Blocking Lemma*, there is a pair (m, w) such that:

- (m, w) blocks $\widehat{\mu}_M$ under the profile P; and
- $m \notin \widehat{M}$.

Hence, neither m nor w misrepresent their preferences.

$\Rightarrow (m, w)$ also block $\widehat{\mu}_M$ under the profile $\widehat{P}$.

$\Rightarrow \widehat{\mu}_M$ is not stable for the profile $\widehat{P}$, a contradiction.

But the two sides don't have the same incentives. . .

P_{m_1}	P_{m_2}
w_1	w_2
w_2	w_1

P_{w_1}	P_{w_2}
m_2	m_1
$\boxed{m_1}$	m_2

P'_{w_2}
m_1

The man-optimal matching with the true preferences.

$$\mu_M(m_1) = w_1 \qquad \text{and} \qquad \mu_M(m_2) = w_2\,.$$

Suppose w_1 is not truthful and submits instead $P'_{w_2} : m_1$ (only m_1 is declared acceptable). The man-optimal matching when w_2 lies:

$$\mu'_M(m_1) = w_2 \qquad \text{and} \qquad \mu'_M(m_2) = w_1\,.$$

$\Rightarrow$ w_2 prefers her match when lying (m_1) to her match when being truthful (m_2).

$\Rightarrow$ Deferred Acceptance is not strategyproof for both sides.

Is there another algorithm that would do the job?

Theorem

There is no matching mechanism that satisfies, for any matching problem, the following two properties at the same time:

(a) *The matching is stable with respect to the submitted preference lists*

(b) *The mechanism is strategyproof for all individuals.*

Proof

P_{m_1}	P_{m_2}
w_1	w_2
w_2	w_1

P_{w_1}	P_{w_2}
m_2	m_1
m_1	m_2

$$\mu_M(m_1) = w_1 \qquad \text{and} \qquad \mu_M(m_2) = w_2$$
$$\mu_W(m_1) = w_2 \qquad \text{and} \qquad \mu_W(m_2) = w_1$$

To satisfy *(a)* we must select for that case either μ_M or μ_W.

If the algorithm selects:

- $\mu_M \quad \Rightarrow \quad w_1$ and w_2 can be better off lying.
- $\mu_W \quad \Rightarrow \quad m_1$ and m_2 can be better off lying.

Take-away

- In a two-sided, one-to-one matching model each side has a strict preference over partners from the other side.
- If X prefers to remain single than being matched to Y, we say Y is unacceptable for X.
- A matching is a function that says, for each individual, who is matched to whom.
- A matching is individually rational if nobody is matched to an unacceptable partner.
- A pair (m, w) block a matching μ if they both prefer each other to their partner under μ.

- A matching is stable if it is
  - Individually rational
  - Not blocked by any pair.
- The Deferred Acceptance algorithm (DA) produces a stable matching:
  - The most preferred stable matching for the proposing side.
  - The least preferred stable matching for the receiving side.
- DA is strategyproof for the proposing side, but not for the receiving side.
- We cannot have, in general, strategyproofness for both sides and stability.
- DA with men proposing yields the man-optimal matching.

  DA with women proposing yields the woman-optimal matching.

- All men prefer the man-optimal matching to any other stable matching.

  All women prefer the woman-optimal matching to any other stable matching.
- All men prefer any stable matching to the woman-optimal matching.

  All women prefer any stable matching to the man-optimal matching.
- DA is strategyproof for the proposing side, but not for the receiving side.
- We cannot have, in general, strategyproofness for both sides and stability.