Two-Sided Matching

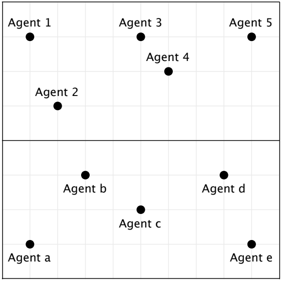
Joana Pais

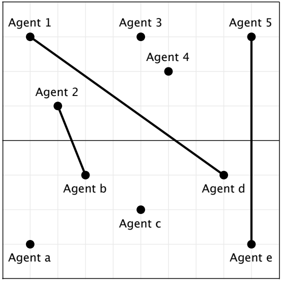
ISEG-UL + UECE/REM

PhD in Economics, 2024

Background

A two–sided **matching market** consists of two (finite and disjoint) sets of agents, where each agent has preferences over the other side of the market. Matching consists in assigning the members of the two sets to one another.





A Bit of History

- The seminal paper *College admissions and the stability of marriage* by David Gale and Lloyd Shapley (henceforth, GS) was published in 1962.
- In the 80s, Alvin Roth studied the matching of doctors to hospitals and showed that the algorithm that was being used since the 50s was equivalent to algorithm described in GS.
- Since then, game theorists have been using one–sided and two–sided matching to design institutions such as (1) student placement in schools, (2) labor markets, and (3) organ donation networks.
- In 2012, the Nobel prize in Economics was awarded to Alvin Roth and Lloyd Shapley “for the theory of stable allocations and the practice of market design.”

Recently

- Entry–level labor markets were the main application until the early 2000s.
- The NRMP matches 25 000 medical students to 4 000 hospitals are matched each year.
- Many cities in the U.S. (NY, Boston, Chicago, Denver, New Orleans) and in other countries (www.matching-in-practice.eu) employ school choice programs.
- In NY, 80 000 students are allocated to 700 schools each year.
- Each year, approximately 10 million high school seniors compete for 6 million seats at various universities in China.
- More than 75 000 patients are waiting for a kidney in the U.S. In 2005, 16 370 transplants were conducted and 4 200 patients died while waiting.
- Landing slots, assignment cadets to positions within the Army, refugee match, time banks,...

Taxonomy

- Matching can involve:
  - **One–sided matching**, i.e., the allocation or exchange of indivisible objects, such as dormitory rooms, transplant organs, courses, summer houses, etc.

    Agents have preferences over objects and may have priorities or claims over objects; the normative properties of the allocation are evaluated from agents' viewpoint.

Taxonomy

- Matching can involve:
  - **One–sided matching**, i.e., the allocation or exchange of indivisible objects, such as dormitory rooms, transplant organs, courses, summer houses, etc.

    Agents have preferences over objects and may have priorities or claims over objects; the normative properties of the allocation are evaluated from agents' viewpoint.
  - Or **two–sided matching**, in markets with two sides, such as firms and workers, students and schools, civil servants and positions, or men and women, that need to be matched with each other.

    Both sides have preferences and evaluation of the properties of the allocation may depend on preferences of both sides of the market.

Taxonomy

- Matching markets:
  - Can be **centralized** when a clearinghouse or matchmaker exists to perform the matching, e.g. some school choice programs, university admission, assignment of high-school teachers to schools.
  - Can be **decentralized** when no such clearinghouse exists, e.g. most job markets.
  - Can involve both centralized and decentralized matching.

The Marriage Market

A **marriage market** is a triple (F, W, R) with

- $F = \{f_1, ..., f_p\}$ is a set of men
- $W = \{w_1, ..., w_p\}$ is a set of women
- $R = (R_k)_{k \in F \cup W}$ is a preference profile where for any f and w

  R_f is an ordered list of preferences on $W \cup \{f\}$

  R_w is an ordered list of preferences on $F \cup \{w\}$.

- Let P_k be the strict preference relation derived from R_k

  iP_kj means iR_kj and not jR_ki.

- Let P_k be the strict preference relation derived from R_k

  iP_kj means iR_kj and not jR_ki.

  If agent $k \in F \cup W$ prefers to remain single rather than be matched to j, i.e., if kP_kj, then j is **unacceptable** to k.

- Let P_k be the strict preference relation derived from R_k

  iP_kj means iR_kj and not jR_ki.

  If agent $k \in F \cup W$ prefers to remain single rather than be matched to j, i.e., if kP_kj, then j is **unacceptable** to k.

  We assume preferences are strict.

Matchings and Stability

An outcome of the marriage market is a **matching**

$$\mu : F \cup W \to F \cup W$$

such that $\mu(f) = w$ iff $\mu(w) = f$ and for all f and w, $\mu(f) \in W \cup \{f\}$ and $\mu(w) \in F \cup \{w\}$.

Matchings and Stability

An outcome of the marriage market is a **matching**

$$\mu : F \cup W \to F \cup W$$

such that $\mu(f) = w$ iff $\mu(w) = f$ and for all f and w, $\mu(f) \in W \cup \{f\}$ and $\mu(w) \in F \cup \{w\}$.

A matching μ is **individually rational** iff there exists no $k \in F \cup W$ that finds $\mu(k)$ unacceptable.

Matchings and Stability

An outcome of the marriage market is a **matching**

$$\mu : F \cup W \to F \cup W$$

such that $\mu(f) = w$ iff $\mu(w) = f$ and for all f and w, $\mu(f) \in W \cup \{f\}$ and $\mu(w) \in F \cup \{w\}$.

A matching μ is **individually rational** iff there exists no $k \in F \cup W$ that finds $\mu(k)$ unacceptable.

A matching μ is **blocked by a pair** $(f, w) \in F \times W$ if they each prefer each other to their partners under μ, i.e.,

$$wP_f\mu(f) \text{ and } fP_w\mu(w).$$

Matchings and Stability

An outcome of the marriage market is a **matching**

$$\mu : F \cup W \rightarrow F \cup W$$

such that $\mu(f) = w$ iff $\mu(w) = f$ and for all f and w, $\mu(f) \in W \cup \{f\}$ and $\mu(w) \in F \cup \{w\}$.

A matching μ is **individually rational** iff there exists no $k \in F \cup W$ that finds $\mu(k)$ unacceptable.

A matching μ is **blocked by a pair** $(f, w) \in F \times W$ if they each prefer each other to their partners under μ, i.e.,

$$wP_f\mu(f) \text{ and } fP_w\mu(w).$$

A matching is **stable** if it is individually rational and if it is not blocked by any pair of agents.

Stability, Core, and Pareto Efficiency

Matching μ is in the **core** of a marriage market if there is no matching ν and coalition $T \subseteq F \cup W$ such that:

for any $k \in T, \nu(k) P_k \mu(k)$ and $\nu(k) \in T$.

Stability, Core, and Pareto Efficiency

Matching μ is in the **core** of a marriage market if there is no matching ν and coalition $T \subseteq F \cup W$ such that:

for any $k \in T, \nu(k) P_k \mu(k)$ and $\nu(k) \in T$.

Theorem

The set of stable matchings equals the core.

Theorem

A stable matching is Pareto efficient.

Mechanism

A mechanism is a systematic procedure which determines a matching for each marriage market. Let a matching market (F, W, R) be denoted by R.

For a mechanism ϕ, $\phi[R]$ is the matching assigned for market R (when agents reveal R).

A mechanism ϕ is stable if $\phi[R]$ is stable for any R.

Existence of a Stable Matching

The (Firm–Proposing) Deferred Acceptance Algorithm (DAA)

Step 1. Each firm proposes to its first choice (if any).

Existence of a Stable Matching

The (Firm–Proposing) Deferred Acceptance Algorithm (DAA)

Step 1. Each firm proposes to its first choice (if any).
Each worker rejects any unacceptable proposals and, if more than one acceptable proposal is received, tentatively holds the most preferred.

Existence of a Stable Matching

The (Firm–Proposing) Deferred Acceptance Algorithm (DAA)

Step 1. Each firm proposes to its first choice (if any).
Each worker rejects any unacceptable proposals and, if more than one acceptable proposal is received, tentatively holds the most preferred.

Step k. Any firm who was rejected at step $k-1$ makes a new proposal to its most preferred acceptable worker who has not rejected it (if any).

Existence of a Stable Matching

The (Firm–Proposing) Deferred Acceptance Algorithm (DAA)

Step 1. Each firm proposes to its first choice (if any).
Each worker rejects any unacceptable proposals and, if more than one acceptable proposal is received, tentatively holds the most preferred.

Step k. Any firm who was rejected at step $k-1$ makes a new proposal to its most preferred acceptable worker who has not rejected it (if any).
Each worker tentatively holds his most preferred acceptable offer to date and rejects the rest.

Existence of a Stable Matching

The (Firm–Proposing) Deferred Acceptance Algorithm (DAA)

Step 1. Each firm proposes to its first choice (if any).
Each worker rejects any unacceptable proposals and, if more than one acceptable proposal is received, tentatively holds the most preferred.

Step k. Any firm who was rejected at step $k-1$ makes a new proposal to its most preferred acceptable worker who has not rejected it (if any).
Each worker tentatively holds his most preferred acceptable offer to date and rejects the rest.

The algorithm terminates after a step where no rejections occur.
Each worker is matched to the firm whose proposal he holds (if any).

Example

Let $F = f_1, f_2, f_3$, $W = w_1, w_2$, and their preferences given by

$$P_{f_1} : w_1, w_2$$
$$P_{f_2} : w_1$$
$$P_{f_3} : w_2, w_1$$
$$P_{w_1} : f_3, f_2, f_1$$
$$P_{w_2} : f_1, f_3.$$

Example

Let $F = f_1, f_2, f_3$, $W = w_1, w_2$, and their preferences given by

$$P_{f_1} : w_1, w_2$$
$$P_{f_2} : w_1$$
$$P_{f_3} : w_2, w_1$$
$$P_{w_1} : f_3, f_2, f_1$$
$$P_{w_2} : f_1, f_3.$$

The resulting matching $\mu = \{(f_1, w_2), (f_2, w_2), (f_3, w_1)\}$ is stable.

Theorem (GS)

A stable matching exists for every marriage market.

Theorem (GS)

A stable matching exists for every marriage market.

Proof.

Observation: Workers get (weakly) better off as the process goes on, and firms get (weakly) worse off as the process goes on.

Theorem (GS)

A stable matching exists for every marriage market.

Proof.

Observation: Workers get (weakly) better off as the process goes on, and firms get (weakly) worse off as the process goes on.

- The DAA always stops
- The DAA produces a matching
- The matching it produces, μ, is always stable with respect to the strict preferences (i.e., after any tie–breaking)
- μ is stable with respect to the original preferences

□

Theorem (GS2)

When all firms and workers have strict preferences, there always exists a firm-optimal stable matching (that every firm likes at least as well as any other stable matching), and a worker-optimal stable matching. Furthermore, the matching μ^F produced by the deferred acceptance algorithm with firms proposing is the firm-optimal stable matching. The worker-optimal stable matching is the matching μ^W produced by the algorithm when workers propose.

Proof.

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

Proof.

Terminology: w is achievable for f if there is some stable matching μ such that $\mu(f) = w$. Proof by induction:
Inductive step: suppose that up to step k of the algorithm, no firm has been rejected by an achievable partner, and that at step k worker w rejects firm f (who is acceptable to w) and (therefore) holds on to some f'.
Then w is not achievable for f. Why? Suppose μ is stable such that $\mu(f) = w$. Then $\mu(f')$ is achievable for f'. Then μ cannot be stable: by the inductive step, (f', w) would be a blocking pair.
Therefore every man is matched with the best achievable partner under μ^F, the outcome of the firm-proposing deferred acceptance algorithm, meaning that μ^F is firm-optimal stable matching. □

Mechanisms

Let ϕ^F be the firm-optimal stable mechanism ($\phi^F[R]$ is the matching determined by the firm-proposing deferred acceptance algorithm for the preference profile R).

Let ϕ^Wbe worker-optimal stable mechanism ($\phi^W[R]$ is the matching determined by worker-proposing deferred acceptance algorithm for the preference profile R).

Let $\mu >_F \mu^0$ denote that all firns like μ at least as well as μ^0, with at least one firm having strict preference.

Then $>_F$ is a partial order on the set of matchings, representing the common preferences of the firms.

Similarly, define $>_W$ as the common preference of the workers.

Theorem (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: if μ and μ^0 are stable matchings, then all firms like μ at least as well as μ^0 if and only if all workers like μ^0 at least as well as μ. That is, $\mu >_W \mu^0$ if and only if $\mu^0 >_F \mu$.

Proof.

Immediate from definition of stability. □

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

For any two matchings μ and μ^0, and for all firms and workers, define $\mu \bigvee^F \mu^0$ (join) as the function that assigns each firm her more preferred of the two matches, and each worker her less preferred:

- $\mu \bigvee^F \mu^0(f) = \mu(f)$ if $\mu(f) P_f \mu^0(f)$ and $\mu \bigvee^F \mu^0(f) = \mu^0(f)$ otherwise.
- $\mu \bigvee^F \mu^0(w) = \mu(w)$ if $\mu^0(w) P_w \mu(w)$ and $\mu \bigvee^F \mu^0(w) = \mu^0(w)$ otherwise.

Define $\mu \bigwedge^M \mu^0$ (meet) analogously, by reversing the preferences.

Lattice Theorem (Conway): When all preferences are strict, if μ and μ^0 are stable matchings, then the functions $\mu \bigvee^F \mu^0$ and $\mu \bigwedge^M \mu^0$ are both matchings and they are both stable.

Theorem

In a market (F, W, R) with strict preferences, the set of people who are matched is the same for all stable matchings.

Theorem (Weak Pareto optimality for firms)

There is no individually rational matching μ (stable or not) such that $\mu(f) P_f \mu^F(f)$ for all $f \in F$.

Theorem (Weak Pareto optimality for firms)

There is no individually rational matching μ (stable or not) such that $\mu(f) P_f \mu^F(f)$ for all $f \in F$.

Example (μ^F may not be strongly Pareto optimal for firms)

$F = \{f_1, f_2, f_3\}$, $W = \{w_1, w_2, w_3\}$

$P_{f_1} : w_2, w_1, w_3$, $P_{f_2} : w_1, w_2, w_3$, $P_{f_3} : w_1, w_2, w_3$,

$P_{w_1} : f_1, f_2, f_3$, $P_{w_2} : f_3, f_1, f_2$, $P_{w_3} : f_1, f_2, f_3$.

Theorem (Weak Pareto optimality for firms)

There is no individually rational matching μ (stable or not) such that $\mu(f) P_f \mu^F(f)$ for all $f \in F$.

Example (μ^F may not be strongly Pareto optimal for firms)

$F = \{f_1, f_2, f_3\}$, $W = \{w_1, w_2, w_3\}$
$P_{f_1} : w_2, w_1, w_3$, $P_{f_2} : w_1, w_2, w_3$, $P_{f_3} : w_1, w_2, w_3$,
$P_{w_1} : f_1, f_2, f_3$, $P_{w_2} : f_3, f_1, f_2$, $P_{w_3} : f_1, f_2, f_3$.
So: $\mu^F = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\} = \mu^W$.

But note that $\mu >^F \mu^F$ for:

$$\mu = \{(f_1, w_2), (f_2, w_3), (f_3, w_1)\}$$

A mechanism ϕ is strategy-proof if for any R, for any $k \in F \cup W$, for any R'_k:

$$\phi[R](k) R_k \phi[R'_k, R_{-k}](k).$$

Theorem (Impossibility Theorem (Roth))

There is no stable and strategy-proof mechanism.

Theorem (Impossibility Theorem (Roth))

There is no stable and strategy-proof mechanism.

Remark on proof: for an impossibility theorem, one example for which no stable matching mechanism induces a truthful revelation a dominant strategy is sufficient.

Proof.

Consider an example with 2 agents on each side, with true preferences R as follows:

$P_{f_1} : w_1, w_2$, $P_{f_1} : w_2, w_1$

$P_{w_1} : f_2, f_1$, $P_{w_2} : f_1, f_2$

Proof.

Consider an example with 2 agents on each side, with true preferences R as follows:

$P_{f_1} : w_1, w_2$, $P_{f_1} : w_2, w_1$

$P_{w_1} : f_2, f_1$, $P_{w_2} : f_1, f_2$

In this example, what must an arbitrary stable mechanism do? I.e. what is the range of $\phi[R]$ if ϕ is a stable mechanism?

Proof.

Consider an example with 2 agents on each side, with true preferences R as follows:

$P_{f_1} : w_1, w_2$, $P_{f_1} : w_2, w_1$

$P_{w_1} : f_2, f_1$, $P_{w_2} : f_1, f_2$

In this example, what must an arbitrary stable mechanism do? I.e. what is the range of $\phi[R]$ if ϕ is a stable mechanism?

Answer: The set of stable matchings. In this market, $\mu_1 = \mu^F = \{(f_1, w_1), (f_2, w_2)\}$ and $\mu_2 = \mu^W = \{(f_1, w_2), (f_2, w_1)\}$ are the only stable matchings.

Proof.

Consider an example with 2 agents on each side, with true preferences R as follows:

$P_{f_1} : w_1, w_2,\ P_{f_1} : w_2, w_1$

$P_{w_1} : f_2, f_1,\ P_{w_2} : f_1, f_2$

In this example, what must an arbitrary stable mechanism do? I.e. what is the range of $\phi[R]$ if ϕ is a stable mechanism?

Answer: The set of stable matchings. In this market, $\mu_1 = \mu^F = \{(f_1, w_1), (f_2, w_2)\}$ and $\mu_2 = \mu^W = \{(f_1, w_2), (f_2, w_1)\}$ are the only stable matchings.

If $\phi[R]= \mu_1$ then w_1 would be better off by stating R'_{w_1} such that $P'_{w_1} = f_2$. In the market $[R'_{w_1}, R_{?w_1}]$ there is a single stable matching μ_2, so necessarily $\phi[R'_{w_1}, R_{?w_1}] = \mu_2$ and

$$\phi[R'_{w_1}, R_{?w_1}](w_1) P_{w_2} \phi[R](w_1)$$

If $\phi[R]= \mu_2$ then f_1 would be better off by stating R'_{f_1} such that $P'_{f_1} = w_1$....

Theorem

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 misstated preferences.)

Theorem (Dubins and Freedman, Roth)

The firm-optimal stable mechanism, ϕ^F makes it a dominant strategy for each firm to state her true preferences.

Consider a complete information game, in which all agents simultaneously report their preferences and a mechanism finds the outcome of the induced marriage market.

Consider a complete information game, in which all agents simultaneously report their preferences and a mechanism finds the outcome of the induced marriage market.

Pure strategy Nash-equilibria exist:

Theorem (Gale and Sotomayor)

When all preferences are strict, let μ be a stable matching for (F, W, R). Suppose each $w \in \mu(F)$ chooses the strategy of listing only $\mu(w)$ on her stated preference list of acceptable firms (and each firm states hers true preferences). This is a Nash-equilibrium in the game induced by the firm-optimal stable mechanism (and is the matching that results).

Furthermore, every Nash-equilibrium misrepresentation by the workers must nevertheless yield a matching that is stable with respect to the true preferences.

Furthermore, every Nash-equilibrium misrepresentation by the workers must nevertheless yield a matching that is stable with respect to the true preferences.

Theorem (Roth)

Suppose each firm $f \in F$ chooses her dominant strategy and states her true preferences R_f, and any worker $w \in W$ chooses any set of strategies (preference lists) R'_w so that R' is an equilibrium for the matching game induced by the firm-optimal stable mechanism. Then the corresponding firm-optimal stable matching for (F, W, R') is one of the stable matchings of (F, W, R), that is $\phi^M[R']$ is stable under R.