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 $\mu : M \cup W \rightarrow M \cup W$ 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;
  - $\nu(m) = w \implies \nu(w) = m$: follows easily from stability.
  - $\nu(w) = m \implies \nu(m) = w$: 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 $W(\mu')$.