# Need vs. Merit: The Large Core of College Admissions Markets

Ran Shorrer, Penn State

Joint with Avinatan Hassidim (Bar-Ilan) and Assaf Romm (Hebrew University)

# Stable Matchings

- Important for the success of centralized marketplaces
  - Roth (2002)
- Prediction for behavior in decentralized markets
  - Banerjee et al. (2013)
- The collection of stable matchings (sometimes = "the core")
  - Not necessarily unique
  - Large literature motivated by multiplicity
- Recent studies suggest that this collection is typically small

# Small Core – Large Literature

- Same applicants assigned, same quotas filled
  - Rural hospital theorem (Roth 1986)
  - With contracts + substitutes (Hatfield Milgrom 2005)
- Small fraction of agents with different stable allocation
  - Evidence from NRMP (Roth Peranson 1999)
    - Simulation results with short lists
  - 1-1 with short preference lists (Immorlica Mahdian 2005)
  - q-1, for fixed q and responsive preferences (Kojima Pathak 2009)
  - q-1, for fixed q and substitutable preferences (Storms 2013)
  - With unequal number of men and women (Ashlagi et al. 2017, Pittel 2017)
  - Continuum of students (limit results) (Azevedo Leshno 2016)
- Small differences in payoffs
  - Preferences with common + independent components (Lee 2017)
  - Correlation in preferences (Holzman Samet 2014)

# Small Core – Implications

- Stability yields sharp predictions
- Pins down the welfare of the overwhelming majority of agents
  - Narrow margins for design of stable mechanisms (small impact on efficiency, equity, etc.)
- Truthful reporting under DA is safe
  - Strategy-proof for "proposers"
  - Equilibrium of complete information game with 1- $\epsilon$  fraction truthful
  - Truthful  $\epsilon$ -BNE as long as no superstar schools
    - Vanishing market power
  - SP-L (Azevedo Budish 2016)

#### אזור מנהל — בדיקת דירוג עדיפויות

יש לגרור את שם המגמה שברצונך לדרג לריבוע האפור. ניתן לשנות את סדר דירוג המגמות ע"י גרירת מגמה מעל/מתחת למגמה אחרת בתוך הריבוע האפור.



# The College Admissions Problem

- Each college, c, has  $q_c$  seats, and a smaller number of scholarships
  - Examples: Israeli Psychology Match, Hungary, Turkey, Australia, Russia, US
- Each college c has a ranking over all students,  $\gg_c$
- Colleges want to recruit the best students under capacity and budget constraints
  - Care lexicographically more about the composition of the cohort than about who gets financial aid
- DA stable and strategy-proof (Hatfield and Kominers 2017; HRS 2017)

# Highlights

- DA is stable, allocates financial aid based on merit
- The collection of stable matchings is LARGE
  - Theoretically (anti-Kojima Pathak result)
    - Can do other models
  - Empirically (>10% of college students in Hungary)
  - Meaningful tradeoffs in selection between stable matchings (can increase the size of incoming cohort in Hungary by >3%)
- Substantial scope for manipulation by colleges

# DA Corresponds to Merit-Based Financial Aid

- On the run of DA, when budget constraint is binding, lower ranked applicants will be rejected first
  - Assumption: applicants prefer to receive financial aid
  - True of both the student-proposing DA, and the contract-proposing DA, where each contract is treated as a separate program
- Used in Hungary, Turkey, Australia
- Is this a **bug** or a **feature**?
  - In all three examples, DA was chosen before funding was introduced

#### Warm-up

- *M* students, *N* colleges
- Colleges' capacity is 1, funding is available
- Any assignment is acceptable (IR)
- Agents like money, but care lexicographically more about the identity of their matched partner
- Claim.  $2\min\{M, N\}$  agents have multiple stable allocations
- **Proof.** Find your favorite stable allocation. It has min{*M*, *N*} matched agents on each side. Shifting money between them preserves stability (from lexicographic preferences).

#### Example

- One college, *c*, with two seats and one scholarship.
- Two applicants: r (rich) and p (poor).
- **r** prefers the funded seat to the unfunded seat  $r: r^f \succ_r r^u \succ_r \emptyset$
- p finds only the funded seat acceptable  $p: p^f \succ_p \emptyset \succ_p p^u$
- $r \gg_{c} p$ . Example c preferences:

$$c: \{r^f, p^u\} \succ_c \{r^u, p^f\} \succ_c \{r^u\} \succ_c \{r^f\} \succ_c \{p^u\} \succ_c \{p^f\}$$

$$\begin{array}{c} r: r^{f} \succ_{r} r^{u} \succ_{r} \emptyset \\ p: p^{f} \succ_{p} \emptyset \succ_{p} p^{u} \end{array}$$
$$c: \left\{ r^{f}, p^{u} \right\} \succ_{c} \left\{ r^{u}, p^{f} \right\} \succ_{c} \left\{ r^{u} \right\} \succ_{c} \left\{ r^{f} \right\} \succ_{c} \left\{ p^{u} \right\} \succ_{c} \left\{ p^{f} \right\} \end{array}$$

$$r: r^{f} \succ_{r} r^{u} \succ_{r} \emptyset$$
$$p: p^{f} \succ_{p} \emptyset \succ_{p} p^{u}$$
$$c: \{r^{f}, p^{u}\} \succ_{c} \{r^{u}, p^{f}\} \succ_{c} \{r^{u}\} \succ_{c} \{r^{f}\} \succ_{c} \{p^{u}\} \succ_{c} \{p^{f}\}$$

• Stable outcome 1 (result of student proposing DA)

$$r: r^{f} \succ_{r} r^{u} \succ_{r} \emptyset$$
$$p: p^{f} \succ_{p} \emptyset \succ_{p} p^{u}$$
$$c: \{r^{f}, p^{u}\} \succ_{c} \{r^{u}, p^{f}\} \succ_{c} \{r^{u}\} \succ_{c} \{r^{f}\} \succ_{c} \{p^{u}\} \succ_{c} \{p^{f}\}$$

- Stable outcome 1 (result of student proposing DA)
- Stable outcome 2
- Budget constraint  $\rightarrow$  loss of <u>substitutes</u>  $\rightarrow$ 
  - loss of lattice structure: no student optimal stable matching
  - No rural hospital theorem: different (number of) students matched

#### Large Markets

- Natural generalization of Kojima Pathak (2009)
  - Today, much less general
- n colleges, 2n students
- Each college has two seats and one scholarship
  - Arbitrary complete ranking over students
- Applicants draw uniformly i.i.d k colleges
- Draw uniformly an *acceptable permutation* over **2k** contracts
  - "Acceptable permutation" = each funded contract ranked over the corresponding unfunded contract

#### Generalizations

- Could have different "popularities"
- Some applicant may not draw all unfunded contracts
  - "poor"
- Some applicants don't like funding
  - Or make "obvious mistakes"
- Programs may have larger and different quotas, multiple levels of aid
- Market may be unbalanced

#### Main Theorem

The expected fraction of colleges that

- 1. can successfully manipulate DA, and
- 2. are assigned a different number of students in different stable allocations
- is bounded below by a positive constant, independent of *n*. Proof

With a bit more work, different number of students in (any) college

• Easy if "poor" students only interested in funded seats (as is common in Hungary)

# Comparison with Kojima Pathak (2009)

Kojima and Pathak's argument:

- 1. A school can successfully manipulate DA by dropping some students from its ROL if it has multiple stable assignments
- Run student proposing DA. Let a schools drop some students, and continue running the algorithm from this point. Schools have vanishing market power: rejection chains not likely to cycle back (likely absorbed by another school)

# Comparison with Kojima Pathak (2009)

• **ROL** – colleges' choice functions are more complex

- cannot be summarized by ROL and one quota
- DA contracts rejected under DA may be part of other stable allocations that the college prefers
  - recall the "poor" student from the example
- Vanishing market power A rejection chain starting at a funded contract has a good chance to end up in the unfunded contract with the same college. No need for "new" offers; Freed-up funds may be used to recruit previously rejected price sensitive students
- Manipulability and unique-stable are logically independent

# Manipulation

- Natural manipulation for colleges under DA
- Declare the "rich" unacceptable with funding
- More generally, applicants with no good outside options
  - Overlap group case, business school financial aid

	DB		Μ	Р
	$\operatorname{mean}$	$\operatorname{sd}$	$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

	DB		Μ	Р
	$\operatorname{mean}$	$\operatorname{sd}$	$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,0	)56
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

	DB		Μ	Р
	$\operatorname{mean}$	$\operatorname{sd}$	$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

	DB		Μ	Р
	mean sd		$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

	DB		Μ	Р
	$\mathrm{mean}$ sd		$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

	DB		Μ	Р
	mean sd		$\operatorname{mean}$	$\operatorname{sd}$
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate $(\%)$	7.95	4.668	6.77	4.162
Gross annual per capita income $(1000 \text{ USD})$	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Table 3: Characteristics of applicants in MP and DB

# Table 1: Characteristics of applicants who submitted ROLs with funded contracts only

Dependent variable	Funded contracts only						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NABC-based SES index	$-0.062^{***}$ (0.0017)	$-0.069^{***}$ (0.0017)					
11th-grade GPA $(1-5)$		$\begin{array}{c} 0.077^{***} \\ (0.0023) \end{array}$		$0.094^{***}$ (0.0012)		$0.094^{***}$ (0.0012)	$0.093^{***}$ (0.0012)
Income (1000 USD)			-0.038*** (0.0006)	$-0.039^{***}$ (0.0006)			
Unemployment (%)					$0.008^{***}$ (0.0002)	$0.008^{***}$ (0.0002)	
Observations	78064	78064	284701	284701	284701	284701	284701
$R^2$	0.017	0.032	0.016	0.038	0.007	0.028	0.021

\* p < 0.1,\*\* p < 0.05,\*\*\* p < 0.01

#### Conclusions

- The college admissions markets typically have large cores
  - Loss of substitutes complicates the situation
- DA allocates funding based on merit
  - Other stable allocations, more "need-based"
- Meaningful tradeoffs for market designers
  - E.g., incentives vs. quantity/efficiency

# Main theorem: Core is large, DA manipulable

- The expected fraction of:
- 1. Students with multiple stable allocations
- 2. Colleges with different size of (stable) incoming cohort
- 3. Colleges that can manipulate the student-proposing DA

Is bounded below by  $\Delta > 0$  where  $\Delta$  does not depend on n.

- Let E(r, p, h, c) denote the event that:
- h prefers r to p (i.e.,  $r \gg_h p$ ).

• 
$$r: h^{funded} >_r h^{unfunded} >_r \dots$$

• 
$$p: h^{funded} >_p c^{funded} >_p \dots$$

• *h* only acceptable for *r*, *p*. *c* only acceptable for *p*.

- Let E(r, p, h, c) denote the event that:
- h prefers r to p (i.e.,  $r \gg_h p$ ).

• 
$$r: \mathbf{h}^{funded} \succ_r \mathbf{h}^{unfunded} \succ_r \dots$$

• 
$$p: h^{funded} >_p c^{funded} >_p \dots$$

• *h* only acceptable for *r*, *p*. *c* only acceptable for *p*.

- Let E(r, p, h, c) denote the event that:
- h prefers r to p (i.e.,  $r \gg_h p$ ).

•  $r: h^{funded} >_r h^{unfunded} >_r \dots$ 

• 
$$p: h^{funded} >_p c^{funded} >_p \dots$$

• *h* only acceptable for *r*, *p*. *c* only acceptable for *p*.

Back

- Let E(r, p, h, c) denote the event that: given h, ways to choose  $r, p, c \approx 4n^3$
- *h* prefers *r* to *p* (i.e.,  $r \gg_h p$ ).  $\times \frac{1}{2}$  choices of r,p,c

• 
$$r: h^{funded} >_r h^{unfunded} >_r \dots$$
  $Pr = \frac{1}{n} \times \frac{1}{k}$   
•  $p: h^{funded} >_p c^{funded} >_p \dots$   $Pr \approx \frac{1}{n^2}$ 

• h only acceptable for r, p. c only acceptable for p. All events are disjoint

$$Pr \approx \left(1 - \frac{k}{n}\right)^{4n} \approx e^{-4k}$$

$$r: r^{f} \succ_{r} r^{u} \succ_{r} \emptyset$$

$$p: p^{f} \succ_{p} \emptyset \succ_{p} p^{u}$$

$$c: \{r^{f}, r^{u}\} \succ_{c} \{r^{f}, p^{u}\} \succ_{c} \{r^{u}, p^{f}\} \succ_{c} \{r^{u}\} \succ_{c} \{r^{f}\} \succ_{c} \{p^{u}\} \succ_{c} \{p^{f}\}$$

Substitutable completion (Hatfield Kominers 17')

- Stable outcome 1 (result of "student proposing DA")
- Stable outcome 2
- Budget constraint  $\rightarrow p^f \notin Ch_c(\{p^f, r^f\}), p^f \in Ch_c(\{p^f, r^f, r^u\}) \rightarrow \text{loss of}$  (unilateral) <u>substitutes</u> (Hatfield Kojima 10')  $\rightarrow$ 
  - loss of lattice structure: no student optimal stable matching
  - No rural hospital theorem: different (number of) students matched

# Empirical Evidence – Hungary

- Thousands of programs, ~100,000 applicants, ~60,000 assigned
- State funded positions are historical norm. Currently ~40,000
- Average ROL has ~4 contracts with ~3 programs
  - 60% rank funded seats only
  - Other ROLs: ~50% have all funded over all unfunded
  - $\rightarrow$  funding plays a more important role for many applicants
- But some students rank the funded seat directly above the unfunded seat in the same program
  - And others list infeasible options between the two
  - $\rightarrow$  the program has **market power** over them