

# Endogenous information acquisition in matching markets: China's college admission mechanisms

Yao Luo

Boston College

May 1, 2024

# Motivation

Theoretical results tell us:

- ◀ Full information? Costly information acquisition to discover preferences. (Corcoran et al. 2018)
- ◀ Market designers should pay attention to the acquisition and flow of information.
- ◀ Finding regret-free stable matching  $\Leftrightarrow$  finding market-clearing cutoffs. (Azevedo and Loshno. 2016, Immorlica et al. 2020)
- ◀ Information deadlocks. Market-clearing cutoffs  $\rightarrow$  budget set  $\rightarrow$  preference formation  $\rightarrow$  determine cutoffs

# Motivation

## Real mechanism implementations:

- ▶ Achieve approximately regret-free stable outcomes by providing external historical data + perturbed capacities: Australia.
- ▶ 2023 Australia: 62,846, China: 12,910,000.
- ▶ Parallel admission (PA) - Direct serial dictatorship (DirSD) + length restriction of the rank-ordered list (ROL)
- ▶ Inner Mongolia dynamic admission (IM) - Sequential serial dictatorship (SeqSD) + sequential moves by groups + time constraints
- ▶ In 2025 Inner Mongolia will give up IM and use PA. IM began in 2007.

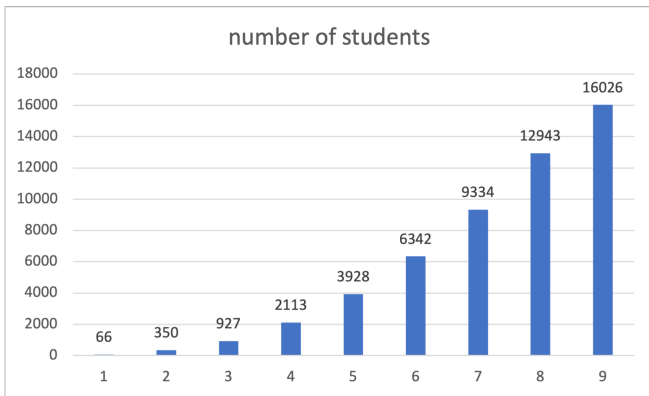
# Motivation

Why switch from IM to PA?

- ◀ Both PA and IM provides historical cutoff scores, but IM offers additional information regarding matching outcomes. Counterintuitive.
- ◀ Theoretically and experimentally, SeqSD leads to higher student welfare than DirSD. (Hakimov et al. 2023)

# Motivation

Real life implementation of SeqSD: groups + time constraints



# Research question

Does PA result in higher student welfare than IM? If so why?

- ◀ The time constraints in IM make the price discovery process too costly?
- ◀ Information communication in IM is not effective?
- ◀ The additional information is too noisy?
- ◀ Maybe IM is actually better than PA and the policy change is purely of political intention.
- ◀ Are there better ways to communicate information to students in IM? Change group size?

# Contribution

- ◀ Provide empirical and experimental comparisons of real-life implementation of DirSD (PA) and SeqSD (IM) mechanisms.
- ◀ Shed light on the importance of information flows in market design and validate costly endogenous information acquisition.
- ◀ Gong and Liang (2023) shows experimentally IM mechanism achieves similar stability as DA mechanism and similar efficiency as Bostom mechanism. Incomplete information. Low correlation of preferences.
- ◀ Chen and Kesten (2019) shows experimentally DA mechanism is better than PA in terms of stability, but the setup assumes complete information.

# Empirical strategy



$$y_i = \alpha_0 + \alpha_1 X_i + \beta Y_{2025} + \varepsilon_i$$

- ▶  $y_i$ : college prestige index (determined by cutoff scores of year 2022 and 2023).  
 $X_i$ : students' gender, ethnicity, rank by exam scores being normalized to be within (0,1).
- ▶ Implicitly assumes higher-ranked students prefer more prestigious colleges. Only care about big names without considering majors.



# Empirical strategy

- ◀ Data: students' gender, ethnicity, exam score, rank and admission result.
- ◀ Spearman's rank correlation coefficient.

$$\rho = 1 - \frac{6 \sum_i^N d_i^2}{n(n^2 - 1)}$$

# Experiment design

## General setup:

- ◀ Exam scores are randomly and independently drawn from the IM empirical distribution between 0 and 100.
- ◀ 30 students competing for 15 seats in 10 colleges. Admission rate is 50%.
- ◀ Students are told that any ranking of the university is equally possible. Preferences are randomly drawn from the space of rankings.
- ◀ Students only know their own exam scores and ranks.

# Experiment design

## General setup:

- ◀ University's quotas and historical cutoffs are common knowledge.
- ◀ Historical cutoffs are gotten by running DirSD.
- ◀ Students need to pay search costs to know their own preferences.
- ◀ Preferences are private knowledge.
- ◀ Students receive more rewards for being assigned to more preferred university.

# Experiment design

## Environments:

- ◀ Dimension 1: The degree of correlation of preferences among students.  
One tier, two tiers and three tiers.
- ◀ Dimension 2: The cost of information acquisition.  
Low and high costs.

# Experiment design

## Predictions:

- ◀ Hypothesis 1: Lower-ranked students gain more from IM compared to PA.
- ◀ Hypothesis 2: IM has a higher probability of being unmatched.
- ◀ Hypothesis 3: Students are more likely to oversearch in IM.

# Experiment design

## Predictions:

- ◀ Hypothesis 4: Welfare: on average IM  $\succ$  PA, but not hold for all students.
- ◀ Hypothesis 5: Smaller group size produces better matching outcomes.
- ◀ Hypothesis 6: Increasing the ROL in PA improves the matching outcomes.

# Empirical distribution

