Revealed Preference Analysis of School Choice Models

by Agarwal and Somaini

Annual Review of Economics, 2020

School Assignment

- Many countries/districts assign students to schools using an assignment mechanism.
- Students/families submit preferences and are "matched" to schools.
- Examples:
  - high school (NY, Boston, Oakland, many other countries);
  - college admission in non-US settings;
  - medical school in the US.

Why do economists care?

- Understand demand for education,
- private incentives for human capital investment,
- equity in opportunity and outcomes,
- allocate human capital investment efficiently.

Challenges

- No prices
- Strategic behavior of applicants

Model: Preferences

Students: $i \in \mathcal{I}$

Schools/programs: $j \in \mathcal{J} = 0, 1 \ldots J$

$$\nu_{ij} = \nu(x_j, z_i, \xi_j, \gamma_i, \varepsilon_{ij}) - d_{ij}$$

with $\nu_{ij} = 0$

Estimation

$$(\gamma_i, \varepsilon_{ij}) \perp d_{ij} \mid z_i, \{x_j, \xi_j\}_{j=1}^{J}$$

Implicit restrictions:

1. Students' utility only depends on their own assignment.
2. No price -- comparing utility within student based on utility for distance, but comparing across students is difficult.
3. No costs of acquiring info, preferences are well-formed.

Mechanisms

Maps student rankings and priorities (exam scores) to a school assignment.

This can include a tie-breaker.

- Two people with the exact same exam score.
- Priorities are intervals of exam scores (Harvard admits SAT>1400)

Researcher knows the mechanism and the priority.

Student Proposing Deferred Acceptance

1. Each student applies to highest ranked school.
    - applications for the highest priority students are *tenatively held* (up to capacity).
2. Students who were rejected apply to highest ranked school that did not reject them.
3. Repeat this process.

DA is **strategy proof** and **stable**.

Strategy Proof

Truth telling is weakly dominant strategy.

Stable

No i and j such that
(i) i strictly prefers j over the school to which they are assigned;
(ii) if school j does not have spare capacity, then i has a higher score than another student assigned to j.

Immediate Acceptance (Boston Mechanism)

1. Each student applies to highest ranked school. Highest priority students are *assigned*.
2. Rejected students apply to next ranked school and assignment is made.
3. Repeat.

IA prioritizes based on the list. Truth telling not weakly dominant.

If you don't have a high test score, don't rank Harvard above State U.

Revealed Preference Analysis

1. Use Stability condition.

2. Use truth-telling property.

3. Other assumptions when students are strategic.

Stability

- Researcher and student both know eligibility score (e_{ij})
- Scores to not depend on lists or lotteries
- cutoffs for admissions are predictable by students and schools

If i is assigned to j then $\nu_{ij} > \nu_{ij'}$ for all j' where that student meets the eligibility.

To learn about preferences outside of choice set, we need to assume that preferences are conditionally independent of eligibility given the observables.

Fack et al (2019)

Paris high school uses a DA.

But students often look like they make "mistakes."

Truth-telling not quite right, but it turns out mistakes are relatively benign (e.g. not ranking).

Using stability to estimate is robust to these benign mistakes.

Truth-telling

Ranking $j > j'$ implies $\nu_{ij} > \nu_{ij'}$.

It is straightforward to write down the likelihood of a ranking (exploded logit).

If a j is not ranked, it must have $\nu_{ij} < 0$.

We can also easily accommodate for only eligible schools.

Strategic Behavior

Can we estimate preferences if all we assume is that students behave according to Bayesian Nash Equilibrium?

Yes

BNE Framework I

Students know mech. and correctly conjecture the distribution of reports by other agents.

Student belief:

$$\mathbf{L}_{R_i} = \int \Phi(R_i, R_{-i}) \prod_{i' \neq i} \sigma_{R_{i'}}(\nu_{i'}) f_V(\nu_{i'}) d\nu_{-i}$$

where Φ is the assignment and $\sigma(\nu)$ is the prob that agent with list utility ν submits list R.

By varying their R, and agent can implicitly choose different

BNE Framework II

Expected utility is $\nu_i \cdot L_{R_i}$.

Agent i will report R_i only if

$$\nu_i \cdot L_{R_i} > \nu_i \cdot L_R$$

for all $R \in \mathcal{R}$.

Estimation: Stability and Truth Telling

Stability

$$P(i \to j) = \frac{exp(\delta_j x_j \gamma z_i - d_{ij})}{1 + \sum_k 1\{\text{k is eligible}\} exp(\delta_j x_j \gamma z_i - d_{ij})}$$

Truth-telling

$$P(i \to j) = \prod_{k=1}^{K_i} \frac{exp(\delta_j x_j \gamma z_i - d_{ij})}{1 + \sum_k 1\{k \neq R_{ik'} \text{ for } k' < k\} exp(\delta_j x_j \gamma z_i - d_{ij})}$$

Probit Alternative

Probit can be used, although the choice probability expression are not tractable.

The solution is to employ a Gibbs sampler to sample from the posterior of the choice probability.

- Agarwal and Somaini (2018) is an example with strategic rankings.
- With strategic rankings, there are beliefs about being accepted that multiply the choice probabilities, so it gets more complicated.

Incomplete Models and Moment Inequalities

$$\mathbb{P}(\nu_i \in C^*_{R_i} \mid x_j, z_i, \xi_j; F_V) \leq \mathbb{P}(R_i \mid x_j, z_i, \xi_j) \leq \mathbb{P}(\nu_i \in C_{R_i} \mid x_j, z_i, \xi_j; F_V)$$

R is rationalizable for ν_i under the model if R belongs to the set of permissable reports.

C_R is the *set* of utilities ν_i such that R is rationalizable.

C^*_R is when R is the only ratinalizable report.

Next Class

LaVerde (2020)