Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation (Abdulkadiroğlu et al., 2017)

Mohit Karnani
Harvard

Empirical Methods in Education Reading Group - EMERG

March 11, 2025

Motivation

- Many school districts use matching mechanisms to assign students to schools.

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).
- Previous literature: use offers as instrumental variables (IV) for enrollment decisions.

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).
- Previous literature: use offers as instrumental variables (IV) for enrollment decisions.
- But how? These "random offers" still depend on parental preferences, priorities, etc.

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).
- Previous literature: use offers as instrumental variables (IV) for enrollment decisions.
- But how? These "random offers" still depend on parental preferences, priorities, etc.
- Counterthought: then just control for all these inputs of the matching system that determine the offers. Now offers are conditionally independent of potential outcomes.

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).
- Previous literature: use offers as instrumental variables (IV) for enrollment decisions.
- But how? These "random offers" still depend on parental preferences, priorities, etc.
- Counterthought: then just control for all these inputs of the matching system that determine the offers. Now offers are conditionally independent of potential outcomes.
- Problem: you can quickly run out of statistical power (e.g. it could be that no two students have the same "type").

Motivation

- Many school districts use matching mechanisms to assign students to schools.
- The most common mechanism is Deferred Acceptance (Gale and Shapley, 1962) and random lottery numbers are often used to break ties and determine offers.
- These tiebreakers can be used to overcome selection bias when estimating causal treatment effects, such as school value added (ATE of attending a given school).
- Previous literature: use offers as instrumental variables (IV) for enrollment decisions.
- But how? These "random offers" still depend on parental preferences, priorities, etc.
- Counterthought: then just control for all these inputs of the matching system that determine the offers. Now offers are conditionally independent of potential outcomes.
- Problem: you can quickly run out of statistical power (e.g. it could be that no two students have the same "type").
- This paper: we can leverage matching mechanisms to construct propensity scores and fully exploit the random tiebreakers, without relying on type-conditioning.

(Notation Slide)

- Applicants I indexed by $i = 1, \ldots, n$. Schools indexed by $s = 0, 1, \ldots, S$, where $s = 0$ is outside option. Capacity vector $\mathbf{q} = (q_0, q_1, \ldots, q_S)$, with $q_0 > n$.
- Applicants have strict preferences over schools $\succ_i$. $a \succ_i b$ means i prefers a over b.
- Applicants have priorities at each school $\rho_{is} \in \{1, \ldots, K, \infty\}$. Lower ρ_{is} means better priority, and $\rho_{is} = \infty$ means i is ineligible at s. Priority vector is $\boldsymbol{\rho}_i = (\rho_{i1}, \ldots, \rho_{iS})$.
- Applicant i's *rank* at school s is $\pi_{is} = \rho_{is} + r_i$, where $r_i \overset{iid}{\sim} U[0, 1]$ is a tiebreaker.
- Applicant *types* are the combination $\theta_i = (\succ_i, \boldsymbol{\rho}_i)$. The set of types is Θ.
- An *assignment* is a vector $\boldsymbol{\mu}$ mapping each student i to a school s, subject to $\mathbf{q}$.
- A (stochastic) *mechanism* φ maps an *economy* $(I, S, \mathbf{q}, \Theta)$ to a distribution of assignments described by a (scaled) bistochastic matrix with elements $p_{is} \in [0, 1]$, s.t. each row $\boldsymbol{p}_i = (p_{i0}, p_{i1}, \ldots, p_{i,S})$ adds to exactly 1, and each column s to at most q_s.

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Step 0 Draw an independently and identically distributed lottery number for each applicant.

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Step 0 Draw an independently and identically distributed lottery number for each applicant.

Step 1 Each applicant applies to their most preferred school. Each school ranks these applicants first by priority, then by random number within priority groups, and provisionally admits the highest-ranked applicants in this order up to its capacity. Other applicants are rejected.

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Step 0 Draw an independently and identically distributed lottery number for each applicant.

Step 1 Each applicant applies to their most preferred school. Each school ranks these applicants first by priority, then by random number within priority groups, and provisionally admits the highest-ranked applicants in this order up to its capacity. Other applicants are rejected.

Step k Each rejected applicant applies to their next most preferred school who hasn't rejected them. Each school ranks these new applicants together with applicants that it admitted provisionally in the previous round, first by priority and then by random number. From this pool, the school provisionally admits those it ranks highest up to capacity, rejecting the rest.

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Step 0 Draw an independently and identically distributed lottery number for each applicant.

Step 1 Each applicant applies to their most preferred school. Each school ranks these applicants first by priority, then by random number within priority groups, and provisionally admits the highest-ranked applicants in this order up to its capacity. Other applicants are rejected.

Step k Each rejected applicant applies to their next most preferred school who hasn't rejected them. Each school ranks these new applicants together with applicants that it admitted provisionally in the previous round, first by priority and then by random number. From this pool, the school provisionally admits those it ranks highest up to capacity, rejecting the rest.

End The algorithm terminates when there are no new applications (some applicants may remain unassigned).

Measuring the Effectiveness of Charter Schools in Denver

- Context: Denver Public Schools (DPS) uses a centralized matching system to assign students to both traditional and charter schools. We want to measure charter VA.
- They run the DA algorithm:

Step 0 Draw an independently and identically distributed lottery number for each applicant.

Step 1 Each applicant applies to their most preferred school. Each school ranks these applicants first by priority, then by random number within priority groups, and provisionally admits the highest-ranked applicants in this order up to its capacity. Other applicants are rejected.

Step k Each rejected applicant applies to their next most preferred school who hasn't rejected them. Each school ranks these new applicants together with applicants that it admitted provisionally in the previous round, first by priority and then by random number. From this pool, the school provisionally admits those it ranks highest up to capacity, rejecting the rest.

End The algorithm terminates when there are no new applications (some applicants may remain unassigned).

- The DA algorithm complies with the *Equal Treatment of Equals* property.

Equal Treatment of Equals (ETE)

- ETE in words: applicants with the same preferences and priorities at all schools have the same assignment probability at each school.

Equal Treatment of Equals (ETE)

- ETE in words: applicants with the same preferences and priorities at all schools have the same assignment probability at each school.
- ETE in math: $\theta_i = \theta_j \implies \boldsymbol{p}_i = \boldsymbol{p}_j$

Equal Treatment of Equals (ETE)

- ETE in words: applicants with the same preferences and priorities at all schools have the same assignment probability at each school.
- ETE in math: $\theta_i = \theta_j \implies \boldsymbol{p}_i = \boldsymbol{p}_j$
- Powerful idea: ETE implies conditional independence!

Equal Treatment of Equals (ETE)

- ETE in words: applicants with the same preferences and priorities at all schools have the same assignment probability at each school.
- ETE in math: $\theta_i = \theta_j \implies \boldsymbol{p}_i = \boldsymbol{p}_j$
- Powerful idea: ETE implies conditional independence!
- If $D_i(s)$ is an indicator for i being matched to s and W_i is a vector of characteristics, then ETE means that $D_i(s) \perp\!\!\!\perp W_i | \theta_i$.

Equal Treatment of Equals (ETE)

- ETE in words: applicants with the same preferences and priorities at all schools have the same assignment probability at each school.
- ETE in math: $\theta_i = \theta_j \implies \boldsymbol{p}_i = \boldsymbol{p}_j$
- Powerful idea: ETE implies conditional independence!
- If $D_i(s)$ is an indicator for i being matched to s and W_i is a vector of characteristics, then ETE means that $D_i(s) \perp\!\!\!\perp W_i | \theta_i$.
- Direct implication:

$$P[D_i(s) = 1 | W_i = w, \theta_i = \theta] = P[D_i(s) = 1 | \theta_i = \theta]$$

Intuition: Equal Treatment of Equals (ETE)

- Why does ETE imply conditional independence?

Intuition: Equal Treatment of Equals (ETE)

- Why does ETE imply conditional independence?
- It basically emulates a stratified RCT: applicants with the same type are randomly assigned to schools based on the same distributions, so the assignment into the treatment is independent of potential outcomes, conditional on the strata.

Intuition: Equal Treatment of Equals (ETE)

- Why does ETE imply conditional independence?
- It basically emulates a stratified RCT: applicants with the same type are randomly assigned to schools based on the same distributions, so the assignment into the treatment is independent of potential outcomes, conditional on the strata.
- So should we control for type? This paper: not really, we can do better!

Intuition: Equal Treatment of Equals (ETE)

- Why does ETE imply conditional independence?
- It basically emulates a stratified RCT: applicants with the same type are randomly assigned to schools based on the same distributions, so the assignment into the treatment is independent of potential outcomes, conditional on the strata.
- So should we control for type? This paper: not really, we can do better!
- Students with different types can still have the same probability of assignment to a given school s (the same propensity score).

Intuition: Equal Treatment of Equals (ETE)

- Why does ETE imply conditional independence?
- It basically emulates a stratified RCT: applicants with the same type are randomly assigned to schools based on the same distributions, so the assignment into the treatment is independent of potential outcomes, conditional on the strata.
- So should we control for type? This paper: not really, we can do better!
- Students with different types can still have the same probability of assignment to a given school s (the same propensity score).
- As propensity scores solely determine the distribution from which the Bernoulli treatments $D_i(s)$ are drawn from, we only need to control for these propensity scores (Rosenbaum and Rubin, 1983). This helps because we can pool types!

Propensity Score Pooling

- The type space is very high dimensional, so conditioning on θ_i is not very practical.

Propensity Score Pooling

- The type space is very high dimensional, so conditioning on θ_i is not very practical.
- Rosenbaum and Rubin (1983) define the *propensity score*, which in this case is $p_s(\theta) = P[D_i(s) = 1|\theta_i = \theta]$. They show how conditioning on the propensity score is enough to guarantee conditional independence $D_i(s) \perp W_i|p_s(\theta_i)$.

Propensity Score Pooling

- The type space is very high dimensional, so conditioning on θ_i is not very practical.
- Rosenbaum and Rubin (1983) define the *propensity score*, which in this case is $p_s(\theta) = P[D_i(s) = 1|\theta_i = \theta]$. They show how conditioning on the propensity score is enough to guarantee conditional independence $D_i(s) \perp W_i|p_s(\theta_i)$.
- Therefore

$$P[D_i(s) = 1|W_i = w, p_s(\theta_i) = p] = P[D_i(s) = 1|p_s(\theta_i) = p] = p$$

Propensity Score Pooling

- The type space is very high dimensional, so conditioning on θ_i is not very practical.
- Rosenbaum and Rubin (1983) define the *propensity score*, which in this case is $p_s(\theta) = P[D_i(s) = 1|\theta_i = \theta]$. They show how conditioning on the propensity score is enough to guarantee conditional independence $D_i(s) \perp W_i|p_s(\theta_i)$.
- Therefore

$$P[D_i(s) = 1|W_i = w, p_s(\theta_i) = p] = P[D_i(s) = 1|p_s(\theta_i) = p] = p$$

- This is great! Now we can have applicants with different types that share the same propensity score and participate in the same "stratified experiment".

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!
- Great, but what about larger markets?

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!
- Great, but what about larger markets?
- It is hard to compute the distribution of assignments for moderately large markets in general (we need $n!$ permutations). So how do we calculate the propensity scores?

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!
- Great, but what about larger markets?
- It is hard to compute the distribution of assignments for moderately large markets in general (we need $n!$ permutations). So how do we calculate the propensity scores?
- One approach: simulate the matching mechanism millions of times!

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!
- Great, but what about larger markets?
- It is hard to compute the distribution of assignments for moderately large markets in general (we need $n!$ permutations). So how do we calculate the propensity scores?
- One approach: simulate the matching mechanism millions of times!
- This works, but you might still end up with too many different values for p, unless you smooth them (e.g. by rounding them).

Propensity Score Pooling (continued)

- Example 1 in the paper shows how 4 different types of applicants share the same propensity score $p_{ia} = 0.25$.
- Thus, if our treatment of interest is $D_i(a)$, instead of conditioning on 5 different types (not feasible with 5 observations), now we can just condition on 2 propensity scores!
- Great, but what about larger markets?
- It is hard to compute the distribution of assignments for moderately large markets in general (we need $n!$ permutations). So how do we calculate the propensity scores?
- One approach: simulate the matching mechanism millions of times!
- This works, but you might still end up with too many different values for p, unless you smooth them (e.g. by rounding them).
- (Also, it's hard to interpret the "source" of randomization.)

Example 1

EXAMPLE 1: Five applicants $\{1, 2, 3, 4, 5\}$ apply to three schools $\{a, b, c\}$, each with one seat. Applicant 5 has the highest priority at c and applicant 2 has the highest priority at b; otherwise the applicants have the same priority at all schools. We are interested in measuring the effect of an offer at school a. Applicant preferences are

$$1 : a \succ b,$$
$$2 : a \succ b,$$
$$3 : a,$$
$$4 : c \succ a,$$
$$5 : c.$$

Applicants 3 and 5 rank only one school.

All applicants have different types, so type-conditioning is not feasible, but applicants 1-4 have the same propensity score for admission into a. All 4 participate in the same stratified experiment!

Example 2

EXAMPLE 2: Four applicants $\{1, 2, 3, 4\}$ apply to three schools $\{a, b, c\}$, each with one seat. There are no school priorities and applicant preferences are

$$1 : c,$$

$$2 : c \succ b \succ a,$$

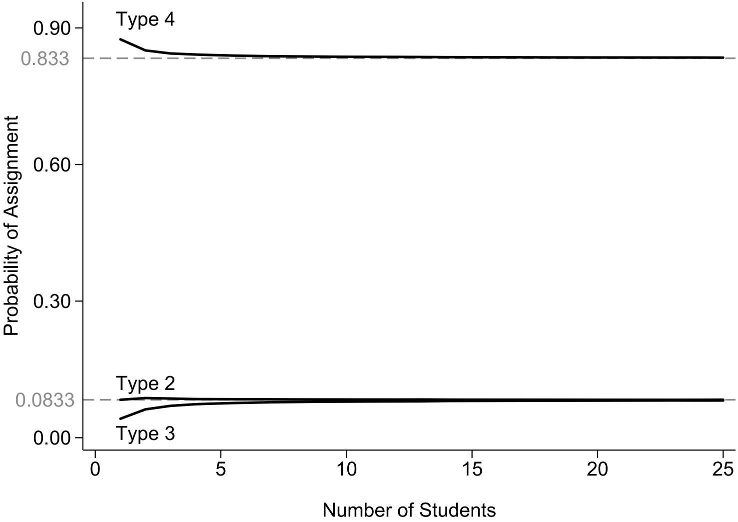
$$3 : b \succ a,$$

$$4 : a.$$

As in Example 1, each applicant is of a different type.

All applicants have different types, so type-conditioning is not feasible, but they also have different propensity scores ($p_{1a} = 0, p_{2a} = 1/12, p_{3a} = 1/24, p_{4a} = 7/8$), so there is no pooling in this small market...

Example 2 in the large



...but in a large market replicating Example 2 we can pool types 2 and 3!

A Large-Market Formula for the DA-score

- Let's assume that the market is *large*, i.e. there is a continuum of applicants $I = [0, 1]$ with a distribution of types and tiebreakers, and capacities $\mathbf{q}$ are proportions of I.

A Large-Market Formula for the DA-score

- Let's assume that the market is *large*, i.e. there is a continuum of applicants $I = [0, 1]$ with a distribution of types and tiebreakers, and capacities $\mathbf{q}$ are proportions of I.
- There is a cutoff characterization of DA, i.e. there are cutoffs $\mathbf{c} = (c_1, \ldots, c_S)$ such that DA assigns applicants to their most preferred school where they clear the cutoff.

A Large-Market Formula for the DA-score

- Let's assume that the market is *large*, i.e. there is a continuum of applicants $I = [0, 1]$ with a distribution of types and tiebreakers, and capacities $\mathbf{q}$ are proportions of I.
- There is a cutoff characterization of DA, i.e. there are cutoffs $\mathbf{c} = (c_1, \ldots, c_S)$ such that DA assigns applicants to their most preferred school where they clear the cutoff.
- With these cutoffs, we can define the *marginal priority* of each school ρ_s as the integer part of c_s. Applicants of type θ that clear this marginal priority ($\rho_{is} < \rho_s$) are *always seated* at s in the DA procedure, conditional on being rejected at the schools they prefer over s. We define these preferred schools as $B_{\theta s} = \{s' \in S | s' \succ_\theta s\}$.

A Large-Market Formula for the DA-score

- Let's assume that the market is *large*, i.e. there is a continuum of applicants $I = [0, 1]$ with a distribution of types and tiebreakers, and capacities $\mathbf{q}$ are proportions of I.
- There is a cutoff characterization of DA, i.e. there are cutoffs $\mathbf{c} = (c_1, \ldots, c_S)$ such that DA assigns applicants to their most preferred school where they clear the cutoff.
- With these cutoffs, we can define the *marginal priority* of each school ρ_s as the integer part of c_s. Applicants of type θ that clear this marginal priority $(\rho_{is} < \rho_s)$ are *always seated* at s in the DA procedure, conditional on being rejected at the schools they prefer over s. We define these preferred schools as $B_{\theta s} = \{s' \in S | s' \succ_\theta s\}$.
- Applicants that have a worse priority than ρ_s $(\rho_{is} > \rho_s)$ are *never seated* at s.

A Large-Market Formula for the DA-score

- Let's assume that the market is *large*, i.e. there is a continuum of applicants $I = [0, 1]$ with a distribution of types and tiebreakers, and capacities $\mathbf{q}$ are proportions of I.
- There is a cutoff characterization of DA, i.e. there are cutoffs $\mathbf{c} = (c_1, \ldots, c_S)$ such that DA assigns applicants to their most preferred school where they clear the cutoff.
- With these cutoffs, we can define the *marginal priority* of each school ρ_s as the integer part of c_s. Applicants of type θ that clear this marginal priority ($\rho_{is} < \rho_s$) are *always seated* at s in the DA procedure, conditional on being rejected at the schools they prefer over s. We define these preferred schools as $B_{\theta s} = \{s' \in S | s' \succ_\theta s\}$.
- Applicants that have a worse priority than ρ_s ($\rho_{is} > \rho_s$) are *never seated* at s.
- Applicants that have marginal priority ($\rho_{is} = \rho_s$) are *conditionally seated*, i.e. they only get a spot at s if they clear the *lottery cutoff* at s, i.e. if the realization of their random tiebreaker r_i is lower than the decimal part of the cutoff $\tau_s = c_s - \rho_s$.

Example: Always, Never and Conditionally Seated

- This partitions the set of applicants who rank s into 3 disjoint subsets: $\Theta_s^n, \Theta_s^a, \Theta_s^c$:

(n) $\Theta_s^n = \{\theta \in \Theta_s \mid \rho_{\theta s} > \rho_s\}$ (never seated)

(a) $\Theta_s^a = \{\theta \in \Theta_s \mid \rho_{\theta s} < \rho_s\}$ (always seated)

(c) $\Theta_s^c = \{\theta \in \Theta_s \mid \rho_{\theta s} = \rho_s\}$ (conditionally seated)

Rank: $\pi_{is} = \rho_{is} + r_i$	Priority: ρ_{is}	Lottery Number: r_i	Offer: $D_i(s)$
1.13	1	0.13	1
1.99	1	0.99	1
2.05	2	0.05	1
2.35	2	0.35	1
2.57	2	0.57	0
2.61	2	0.61	0
3.12	3	0.12	0
3.32	3	0.32	0

In this example, the cutoff is $c_s = 2.35$, the marginal priority (integer part of cutoff) is $\rho_s = 2$, and the lottery cutoff (decimal part of cutoff) is $\tau_s = 0.35$. The first two are always seated, the last two are never seated, and the ones in the middle are conditionally seated.

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\mathrm{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta \tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta \tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max \{\tau_{\tilde{s}} \mid \rho_{\theta \tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s}\} & \text{if } \rho_{\theta \tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\text{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta\tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max\left\{\tau_{\tilde{s}} \mid \rho_{\theta\tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s}\right\} & \text{if } \rho_{\theta\tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

- This tells us how to truncate r_i when applying to s, conditional on other rejections:

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\mathrm{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta \tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta \tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max \left\{ \tau_{\tilde{s}} \mid \rho_{\theta \tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s} \right\} & \text{if } \rho_{\theta \tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

- This tells us how to truncate r_i when applying to s, conditional on other rejections:
  - 0 if priority is worse-than-marginal at all preferred schools (no truncation)

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\text{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta \tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta \tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max\left\{\tau_{\tilde{s}} \mid \rho_{\theta \tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s}\right\} & \text{if } \rho_{\theta \tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

- This tells us how to truncate r_i when applying to s, conditional on other rejections:
  - 0 if priority is worse-than-marginal at all preferred schools (no truncation)
  - 1 if preferred schools include an option where seated with certainty (full truncation)

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\text{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta\tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max\left\{\tau_{\tilde{s}} \mid \rho_{\theta\tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s}\right\} & \text{if } \rho_{\theta\tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

- This tells us how to truncate r_i when applying to s, conditional on other rejections:
  - 0 if priority is worse-than-marginal at all preferred schools (no truncation)
  - 1 if preferred schools include an option where seated with certainty (full truncation)
  - Somewhere in between if marginal or worse somewhere they prefer; rejections at "highly selective" schools where applicant is marginal are not informative, we only care about the most "lenient" cutoff.

A Large-Market Formula for the DA-score (continued)

- With all these ingredients, we can define the *Most Informative Disqualification* (MID):

$$\text{MID}_{\theta s} = \begin{cases} 0 & \text{if } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \\ 1 & \text{if } \rho_{\theta\tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max\left\{\tau_{\tilde{s}} \mid \rho_{\theta\tilde{s}} = \rho_{\tilde{s}}, \tilde{s} \in B_{\theta s}\right\} & \text{if } \rho_{\theta\tilde{s}} \geq \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s} \end{cases}$$

- This tells us how to truncate r_i when applying to s, conditional on other rejections:
  - 0 if priority is worse-than-marginal at all preferred schools (no truncation)
  - 1 if preferred schools include an option where seated with certainty (full truncation)
  - Somewhere in between if marginal or worse somewhere they prefer; rejections at "highly selective" schools where applicant is marginal are not informative, we only care about the most "lenient" cutoff.
- Simple interpretation: the MID is the probability of being seated at some $\tilde{s} \in B_{\theta s}$ you prefer over s.

A Large-Market Formula for the DA-score (continued)

- Provided we know $\text{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \text{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \text{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \text{MID}_{\theta s}}{1 - \text{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

A Large-Market Formula for the DA-score (continued)

- Provided we know $\mathrm{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \mathrm{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \mathrm{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \mathrm{MID}_{\theta s}}{1 - \mathrm{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

- Three cases again:

A Large-Market Formula for the DA-score (continued)

- Provided we know $\text{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \text{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \text{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \text{MID}_{\theta s}}{1 - \text{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

- Three cases again:
  - The first one is easy: never seated applicants are never seated.

A Large-Market Formula for the DA-score (continued)

- Provided we know $\text{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \text{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \text{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \text{MID}_{\theta s}}{1 - \text{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

- Three cases again:
  - The first one is easy: never seated applicants are never seated.
  - The second is not too hard: always seated applicants are seated *only if they don't get something better*. This case implies that even if a school is undersubscribed, there can be randomization!

A Large-Market Formula for the DA-score (continued)

- Provided we know $\mathrm{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \mathrm{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \mathrm{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \mathrm{MID}_{\theta s}}{1 - \mathrm{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

- Three cases again:
  - The first one is easy: never seated applicants are never seated.
  - The second is not too hard: always seated applicants are seated *only if they don't get something better*. This case implies that even if a school is undersubscribed, there can be randomization!
  - The third one has 2 parts: the first one is the probability of not getting something better, and the second one is the probability of being admitted into school s, conditional on being able to *point at it* during the DA procedure.

A Large-Market Formula for the DA-score (continued)

- Provided we know $\text{MID}_{\theta s}$, we can easily compute the DA propensity score as

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - \text{MID}_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - \text{MID}_{\theta s}) \times \max\left\{0, \frac{\tau_s - \text{MID}_{\theta s}}{1 - \text{MID}_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c \end{cases}$$

- Three cases again:
  - The first one is easy: never seated applicants are never seated.
  - The second is not too hard: always seated applicants are seated *only if they don't get something better*. This case implies that even if a school is undersubscribed, there can be randomization!
  - The third one has 2 parts: the first one is the probability of not getting something better, and the second one is the probability of being admitted into school s, conditional on being able to *point at it* during the DA procedure.
- Now we just take the empirical analogue of this formula and we're done!

Byproduct: MID helps us learn about the source of randomization

TABLE III

DA SCORE ANATOMY AT STRIVE PREP SCHOOLS (2013 APPLICANTS)[a]

				Θ_s^n	Θ_s^c		Θ_s^a		
	Eligible			$0 \le \text{MID} \le 1$	$\text{MID} \ge \tau_s$	$\text{MID} < \tau_s$	$\text{MID} = 1$	$0 < \text{MID} < 1$	$\text{MID} = 0$
	Applicants	Capacity	Offers	$p_s(\theta) = 0$	$p_s(\theta) = 0$	$0 < p_s(\theta) < 1$	$p_s(\theta) = 0$	$0 < p_s(\theta) < 1$	$p_s(\theta) = 1$
Campus	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GVR	324	147	112	0	0	0	159	116	49
Lake	274	147	126	0	0	0	132	26	116
Highland	244	147	112	0	0	0	121	21	102
Montbello	188	147	37	0	0	0	128	31	29
Federal	574	138	138	78	284	171	3	1	37
Westwood	494	141	141	53	181	238	4	0	18

[a]This table shows how formula scores are determined for STRIVE applicants in grade 6 (all 6th-grade seats at these schools are assigned in a single bucket; ineligible applicants are omitted) for applicants applying for seats in the fall 2013 school year. Column 3 records offers made to these applicants. Columns 4, 5, and 7 show the number of applicants in partitions with a score of zero. Columns 6 and 8 show the number of applicants subject to random assignment. Column 9 shows the number of applicants with certain offers.

Conditional-on-score IV

- As in previous literature, our instrument is an indicator for being assigned to a charter school $D_i = \sum_{s \in Charters} D_i(s)$, and we want to measure the causal effect of enrolling in a charter school C_i on test scores Y_i.

Conditional-on-score IV

- As in previous literature, our instrument is an indicator for being assigned to a charter school $D_i = \sum_{s \in Charters} D_i(s)$, and we want to measure the causal effect of enrolling in a charter school C_i on test scores Y_i.
- The new thing here is that we have propensity scores $p_s(\theta)$, which we aggregate to any-charter propensity scores $p_D(\theta) = \sum_{s \in Charters} p_s(\theta)$, and we can use them to make D_i conditionally independent of any W_i (including potential outcomes).

Conditional-on-score IV

- As in previous literature, our instrument is an indicator for being assigned to a charter school $D_i = \sum_{s \in Charters} D_i(s)$, and we want to measure the causal effect of enrolling in a charter school C_i on test scores Y_i.
- The new thing here is that we have propensity scores $p_s(\theta)$, which we aggregate to any-charter propensity scores $p_D(\theta) = \sum_{s \in Charters} p_s(\theta)$, and we can use them to make D_i conditionally independent of any W_i (including potential outcomes).
- Our conditional-on-score Wald-IV estimand is

$$\frac{\mathbb{E}[Y_i|D_i = 1, p_D(\theta_i) = x] - \mathbb{E}[Y_i|D_i = 0, p_D(\theta_i) = x]}{\mathbb{E}[C_i|D_i = 1, p_D(\theta_i) = x] - \mathbb{E}[C_i|D_i = 1, p_D(\theta_i) = x]} = \mathbb{E}[Y_{1i} - Y_{0i}|p_D(\theta_i) = x, C_{1i} - C_{0i} > 0].$$

Conditional-on-score IV

- As in previous literature, our instrument is an indicator for being assigned to a charter school $D_i = \sum_{s \in Charters} D_i(s)$, and we want to measure the causal effect of enrolling in a charter school C_i on test scores Y_i.
- The new thing here is that we have propensity scores $p_s(\theta)$, which we aggregate to any-charter propensity scores $p_D(\theta) = \sum_{s \in Charters} p_s(\theta)$, and we can use them to make D_i conditionally independent of any W_i (including potential outcomes).
- Our conditional-on-score Wald-IV estimand is

$$\frac{\mathbb{E}[Y_i|D_i = 1, p_D(\theta_i) = x] - \mathbb{E}[Y_i|D_i = 0, p_D(\theta_i) = x]}{\mathbb{E}[C_i|D_i = 1, p_D(\theta_i) = x] - \mathbb{E}[C_i|D_i = 1, p_D(\theta_i) = x]} = \mathbb{E}[Y_{1i} - Y_{0i}|p_D(\theta_i) = x, C_{1i} - C_{0i} > 0].$$

- Instead of reporting LATEs for every level x of the propensity score, they aggregate everything in a 2SLS framework:

$$C_i = \sum_x \gamma(x) d_i(x) + \delta D_i + X_i' \lambda + \nu_i$$

$$Y_i = \sum_x \alpha(x) d_i(x) + \beta C_i + X_i' \mu + \varepsilon_i$$

Empirical Application

First, conditioning on the propensity score indeed balances the offer and no-offer groups:

TABLE V

STATISTICAL TESTS FOR BALANCE[a]

	Non-Offered Mean (1)	No Controls (2)	Simulated Score Controls: Rounded (Hundredths) (3)	Simulated Score Controls: Rounded (Thousandths) (4)	DA Score Controls: Frequency (Saturated) (5)	DA Score Controls: Formula (Saturated) (6)	Full Applicant Type Controls (7)
A. Application variables							
Number of schools ranked	4.438	−0.544***	−0.076	−0.084	−0.075	−0.104	−0.046
		(0.031)	(0.070)	(0.073)	(0.071)	(0.069)	(0.041)
Number of charter schools ranked	1.450	0.443***	−0.027	−0.029	0.036	−0.020	0.001
		(0.018)	(0.035)	(0.037)	(0.037)	(0.035)	(0.018)
First school ranked is charter	0.275	0.639***	−0.026	−0.022	−0.005	−0.003	0.000
		(0.007)	(0.017)	(0.015)	(0.015)	(0.014)	(0.000)
B. Baseline covariates							
Origin school is charter	0.087	0.118***	−0.027*	−0.029**	−0.028**	−0.038***	0.024
		(0.007)	(0.014)	(0.014)	(0.012)	(0.012)	(0.015)
Female	0.512	−0.017*	0.030	0.023	0.020	0.020	0.027
		(0.010)	(0.025)	(0.027)	(0.026)	(0.026)	(0.055)
Hispanic	0.597	0.102***	−0.011	−0.013	−0.014	−0.007	0.025
		(0.010)	(0.021)	(0.023)	(0.021)	(0.022)	(0.034)
Black	0.188	−0.052***	0.004	0.000	0.006	0.003	−0.020
		(0.007)	(0.019)	(0.020)	(0.019)	(0.019)	(0.028)
Subsidized lunch	0.782	0.052***	−0.007	−0.003	0.004	0.010	0.031
		(0.008)	(0.018)	(0.019)	(0.018)	(0.018)	(0.031)
Limited English proficient	0.305	0.089***	0.006	0.017	0.002	0.019	0.007
		(0.010)	(0.023)	(0.026)	(0.024)	(0.025)	(0.051)
Special education	0.093	−0.005	0.014	0.009	0.006	0.012	0.036
		(0.006)	(0.014)	(0.016)	(0.014)	(0.015)	(0.023)
N	5,674	9,879	2,714	2,291	2,445	2,404	464

Results: Charters are Good!

TABLE VI

CHARTER EFFECTS ESTIMATED USING ALTERNATIVE SCORE CONTROLS[a]

	Simulated Score Controls (Rounded to Hundredths)			DA Score Controls (With Covariates)		No Score Controls (With Covariates)	
	Semiparametric (1)	2SLS (No Covariates) (2)	2SLS (With Covariates) (3)	Frequency (Saturated) (4)	Formula (Saturated) (5)	2SLS (6)	OLS (7)
First stage	0.389***	0.415***	0.420***	0.443***	0.435***	0.561***	
	{0.053}	(0.024)	(0.024)	(0.024)	(0.024)	(0.016)	
Math	0.372***	0.351***	0.415***	0.417***	0.409***	0.231***	0.230***
	{0.133}	(0.108)	(0.052)	(0.050)	(0.051)	(0.030)	(0.010)
Reading	0.180	0.083	0.166***	0.174***	0.166***	0.066**	0.094***
	{0.162}	(0.108)	(0.053)	(0.050)	(0.052)	(0.029)	(0.010)
Writing	0.217	0.184*	0.274***	0.295***	0.315***	0.141***	0.171***
	{0.136}	(0.105)	(0.058)	(0.056)	(0.058)	(0.032)	(0.011)
N	2,229	2,308	2,308	2,099	2,058	2,947	8,528

Thoughts / Extensions / Future Research

- Followup papers: Abdulkadiroğlu et al. (2022); Angrist et al. (2024)
- Fun fact: at least 1/3 of countries around the world use some sort of centralized matching system at some education level (Neilson, 2024). Can you get access to this type of data in some other country/school district?
- Other matching mechanisms might have similar propensity score formulas. Can you derive the formula for Top Trading Cycles (TTC)?
- The propensity scores computed via the matching mechanism are useful to study many other topics: strategizing in non-strategy proof mechanisms (Agarwal and Somaini, 2018); RCTs with interference (Karnani, 2023); school integration (Angrist et al., 2022); VA of Catholic schools (Munevar, 2023). Can you think of new applications?

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2017). Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation. *Econometrica*, 85(5):1373–1432.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2022). Breaking ties: Regression discontinuity design meets market design. *Econometrica*, 90(1):117–151.

Agarwal, N. and Somaini, P. (2018). Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism. *Econometrica*, 86(2):391–444.

Angrist, J., Gray-Lobe, G., Idoux, C., and Pathak, P. A. (2022). Still Worth the Trip? School Busing Effects in Boston and New York. *SSRN Electronic Journal*, (July).

Angrist, J., Hull, P., Pathak, P. A., and Walters, C. (2024). Credible School Value-Added with Undersubscribed School Lotteries. *The Review of Economics and Statistics*, 106(1):1–19.

Gale, D. and Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15.

Karnani, M. (2023). RCTs with Interference in Matching Systems. *(MIT Dissertation)*.

Munevar, I. (2023). Catholic schools and student performance: Evidence from Chile's school-admission lottery. *Job Market Paper*.

Neilson, C. A. (2024). The Rise of Centralized Mechanisms in Education Markets Around the World: ccas-project.org.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.