A Dynamic Framework of School Choice: Effects of Middle Schools on High School Choice

Dong Woo Hahm
University of Southern California

Minseon Park
University of Wisconsin-Madison

DSE 2022 Conference
August 19th, 2022

Introduction: Market Design Based School Choice

Introduction: Market Design Based School Choice

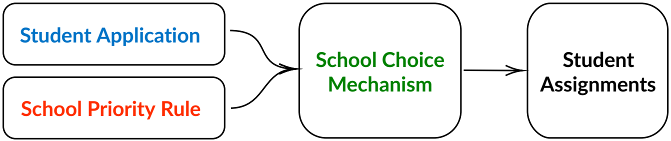
- **School choice**: a tool of assigning students to schools based on their choices

Introduction: Market Design Based School Choice

- **School choice**: a tool of assigning students to schools based on their choices
- Widely used to assign students to schools **at various levels of education**
  e.g. US (Boston, Chicago, **NYC**...), Australia, Chile, England, Hungary, Paris, Turkey

Introduction: Market Design Based School Choice

- **School choice**: a tool of assigning students to schools based on their choices
- Widely used to assign students to schools **at various levels of education**
  e.g. US (Boston, Chicago, **NYC**...), Australia, Chile, England, Hungary, Paris, Turkey



Motivation: Sequential Nature of School Choice

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools
   - Admissions reforms have targeted each level of education separately

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools
   - Admissions reforms have targeted each level of education separately
   - **e.g.** Racial segregation across high schools

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools
   - Admissions reforms have targeted each level of education separately
   - **e.g.** Racial segregation across high schools
     - Most existing policies **only** target high school admissions

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools
   - Admissions reforms have targeted each level of education separately
   - **e.g.** Racial segregation across high schools
     - Most existing policies **only** target high school admissions
     - But, segregation in high schools may develop earlier

Motivation: Sequential Nature of School Choice

- Students face school choices multiple times $\rightarrow$ these choices are closely related
- **Nevertheless**, this relationship has been neglected by both scholars and policymakers

1. No existing studies examine the sequential nature of school choices
   - Absence of **empirical evidence** and an **appropriate framework**

2. Policymakers failed to consider it when tackling segregation across schools
   - Admissions reforms have targeted each level of education separately
   - **e.g.** Racial segregation across high schools
     - Most existing policies **only** target high school admissions
     - But, segregation in high schools may develop earlier $\rightarrow$ any *high school-only* policy may fail to address this

Research Questions

Research Questions

1. **"Does a student's previous school choice affect the subsequent school choices?"**

Research Questions

1. **"Does a student's previous school choice affect the subsequent school choices?"**
2. **"How can one address racial segregation across schools using this relationship?"**

Research Questions

1. **“Does a student’s previous school choice affect the subsequent school choices?”**
2. **“How can one address racial segregation across schools using this relationship?”**

Objectives:

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Setting:

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Setting: NYC public middle and high school choices

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Setting: NYC public middle and high school choices

- Student-level application data: 2014-15 to middle schools, 2017-18 to high schools

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Setting: NYC public middle and high school choices

- Student-level application data: 2014-15 to middle schools, 2017-18 to high schools
- Student Proposing Deferred Acceptance (DA)

Research Questions

1. “Does a student’s previous school choice affect the subsequent school choices?”

2. “How can one address racial segregation across schools using this relationship?”

Objectives:

1. Develop a novel, evidence-based dynamic model of school choices
2. Provide a new insight into how to understand and address segregation

Setting: NYC public middle and high school choices

- Student-level application data: 2014-15 to middle schools, 2017-18 to high schools
- Student Proposing Deferred Acceptance (DA)
- Prespecified school priority rules: intrinsic priority groups + single tie-breaking

What We Do

What We Do

1. Reduced-form analysis
   - Empirical evidence of middle schools' effects on high school applications/assignments
     - Use randomness in middle school assignments generated by tie-breaking rules

What We Do

1. Reduced-form analysis
   - **Empirical evidence** of middle schools' effects on high school applications/assignments
     - Use randomness in middle school assignments generated by tie-breaking rules
2. Structural analysis of choices over two periods:
   - Develop and estimate a **novel dynamic model of school choice**

What We Do

1. Reduced-form analysis
   - Empirical evidence of middle schools' effects on high school applications/assignments
     - Use randomness in middle school assignments generated by tie-breaking rules

2. Structural analysis of choices over two periods:
   - Develop and estimate a novel dynamic model of school choice
   - Decompose middle schools' effects into two **channels**

$$\left.\begin{array}{l}\text{1. How } \textbf{students} \text{ rank } \textbf{high schools} \\ \text{2. How } \textbf{high schools} \text{ rank } \textbf{students}\end{array}\right\} \longrightarrow \text{high school assignments}$$

What We Do

1. Reduced-form analysis
   - **Empirical evidence** of middle schools' effects on high school applications/assignments
     - Use randomness in middle school assignments generated by tie-breaking rules

2. Structural analysis of choices over two periods:
   - Develop and estimate a **novel dynamic model of school choice**
   - Decompose middle schools' effects into two **channels**

     $\left.\begin{array}{l}\text{1. How } \textbf{students} \text{ rank } \textbf{high schools} \\ \text{2. How } \textbf{high schools} \text{ rank } \textbf{students}\end{array}\right\} \longrightarrow \text{high school assignments}$

3. Counterfactual analysis: concurrent admissions reforms in NYC
   - Effects of admission reforms on segregation, **when implemented at alternative educational stages**

Preview of Findings

Preview of Findings

1. Middle schools do change students' tastes on high schools
   - Middle schools of higher quality make students value the quality of high schools more
   - Middle schools with many students of the same race strengthen racial homophily

Preview of Findings

1. Middle schools do change students' tastes on high schools
   - Middle schools of higher quality make students value the quality of high schools more
   - Middle schools with many students of the same race strengthen racial homophily
2. Changes in tastes impact high school assignments by affecting how students rank high schools
   - The impact of middle schools on high school assignment mainly operates by affecting high school applications

Preview of Findings

1. Middle schools do change students' tastes on high schools
   - Middle schools of higher quality make students value the quality of high schools more
   - Middle schools with many students of the same race strengthen racial homophily
2. Changes in tastes impact high school assignments by affecting how students rank high schools
   - The impact of middle schools on high school assignment mainly operates by affecting high school applications
3. Middle school-only admissions reform can desegregate **not only** middle schools **but also** high schools

Preview of Findings

1. Middle schools do change students' tastes on high schools
   - Middle schools of higher quality make students value the quality of high schools more
   - Middle schools with many students of the same race strengthen racial homophily
2. Changes in tastes impact high school assignments by affecting how students rank high schools
   - The impact of middle schools on high school assignment mainly operates by affecting high school applications
3. Middle school-only admissions reform can desegregate **not only** middle schools **but also** high schools

Policy implication:

Dynamics of school choice can be used to design more effective policies

Related Literature

Effects of schools on future outcomes

- Academic performance (Jackson 2010; Pop-Eleches and Urquiola 2013; Abdulkadiroğlu, Angrist, and Pathak 2014; Deming, Hastings, Kane, and Staiger 2014; Dobbie and Fryer 2014), labor market outcomes (Card and Krueger 1992a,b; Clark and Bono 2016)
- **This paper**: first to evaluate the impacts on students' future academic choices in a K-12 context

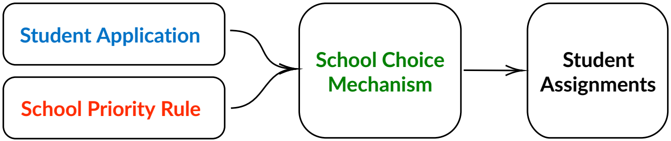
Quasi-experiments in student assignments

- Hoxby and Rockoff, 2004; Deming, Hastings, Kane, and Staiger, 2014; Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu, Angrist, and Pathak, 2014; Dobbie and Fryer, 2014; **Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017, 2021**

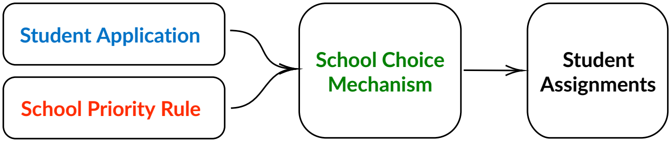
School choice

- Assignment mechanism (Abdulkadiroğlu, Che, and Yasuda 2015; Abdulkadiroğlu, Agarwal, and Pathak 2017; He 2017; Agarwal and Somaini 2018; Che and Tercieux 2019; Calsamiglia, Fu, and Güell 2020), information provision (Hastings and Weinstein 2008; Ajayi, Friedman, and Lucas 2017; Luflade 2017; Corcoran, Jennings, Cohodes, and Sattin-Bajaj 2018; Chen and He 2021a,b; Grenet, He, and Kübler 2021)
- **This paper**: first to incorporate a dynamic framework to study the relationship of school choices at different stages

Centralized Assignment in NYC

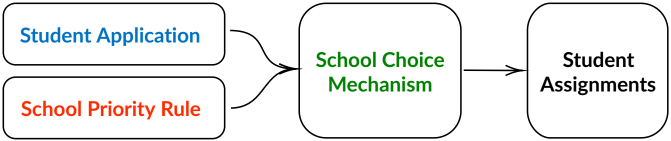


Centralized Assignment in NYC



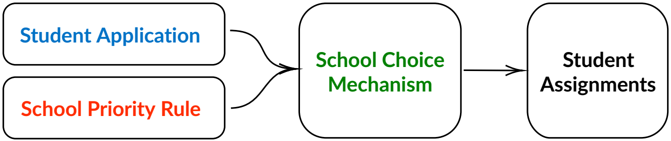
- Student applications: rank ordered lists (ROL) of schools

Centralized Assignment in NYC



- Student applications: rank ordered lists (ROL) of schools
- Prespecified school priorities:

Centralized Assignment in NYC



- Student applications: rank ordered lists (ROL) of schools

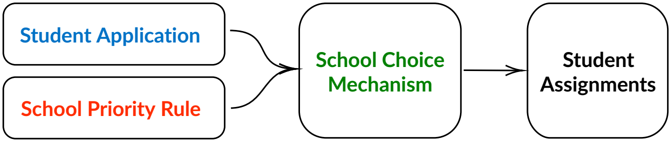
- Prespecified school priorities:

  1. Priority group: applicants are classified into few groups

     e.g. 1st priority: Manhattan students

     2nd priority: information session attendees

Centralized Assignment in NYC



- Student applications: rank ordered lists (ROL) of schools

- Prespecified school priorities: lexicographic

  1. Priority group: applicants are classified into few groups

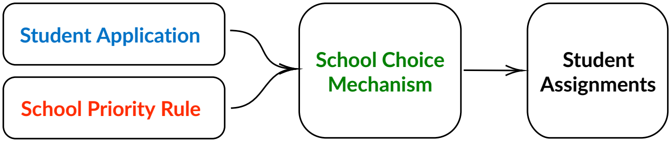
     e.g. 1st priority: Manhattan students

     2nd priority: information session attendees

  2. **Tie-breaking**: break ties within the same priority group

     e.g. Lotteries or test scores

Centralized Assignment in NYC



- Student applications: rank ordered lists (ROL) of schools

- Prespecified school priorities: lexicographic

  1. Priority group: applicants are classified into few groups

     e.g. 1st priority: Manhattan students

     2nd priority: information session attendees

  2. **Tie-breaking**: break ties within the same priority group

     e.g. Lotteries or test scores

- Mechanism: **Gale–Shapley Deferred Acceptance (DA)** DA

NYC DOE School Choice Data

NYC DOE School Choice Data

- Students
  - MS application: 2014-15
  - HS application: 2017-18 (corresponding to 2014-15 MSAP cohort)
  - Demographics, standardized test scores data, residence (census tract level) data
  - Unique student ID to link all data

NYC DOE School Choice Data

- Students
  - MS application: 2014-15
  - HS application: 2017-18 (corresponding to 2014-15 MSAP cohort)
  - Demographics, standardized test scores data, residence (census tract level) data
  - Unique student ID to link all data
- Schools
  - NYC DOE MS and HS directory

NYC DOE School Choice Data

- Students
  - MS application: 2014-15
  - HS application: 2017-18 (corresponding to 2014-15 MSAP cohort)
  - Demographics, standardized test scores data, residence (census tract level) data
  - Unique student ID to link all data

- Schools
  - NYC DOE MS and HS directory
  - Two definitions of groups of middle/high schools:
    1. High achievement: top 1/3 in terms of average test score
    2. High minority: top 1/3 in terms of % Black or Hispanic

    defined using demographics of current seniors Histogram

NYC DOE School Choice Data

- Students
  - MS application: 2014-15
  - HS application: 2017-18 (corresponding to 2014-15 MSAP cohort)
  - Demographics, standardized test scores data, residence (census tract level) data
  - Unique student ID to link all data

- Schools
  - NYC DOE MS and HS directory
  - Two definitions of groups of middle/high schools:
    1. High achievement: top 1/3 in terms of average test score
    2. High minority: top 1/3 in terms of % Black or Hispanic

    defined using demographics of current seniors Histogram

- 54,012 students applying to 472 middle schools, and then applying to 426 high schools 3 years later Attrition Summary: Stu

Summary Statistics: Schools

	(1)	(2)	(3)	(4)	(5)	(6)
	All		High Achievement		High Minority	
Variables	Mean	Std	Mean	Std	Mean	Std
Panel A: Middle School Characteristics						
Average Test Score (6th graders)	297.30	(20.52)	313.00	(17.72)	282.50	(11.82)
% Black or Hispanic	70.92	(30.51)	47.92	(30.82)	97.16	(2.74)
Panel B: High School Characteristics						
4yr Graduation Rate (%)	67.01	(17.34)	81.79	(10.81)	60.67	(16.19)
College Enrollment Rate (%)	58.38	(17.13)	73.96	(11.49)	52.23	(15.11)
Average Test Score (9th graders)	294.4	(17.49)	311.3	(14.36)	285.7	(13.37)
% Black or Hispanic	76.58	(23.40)	57.98	(26.10)	95.49	(04.20)

Note: Average Test Score is the average of the mean of statewide ELA and Math scores in a given school.

Summary Statistics: Schools

	(1)	(2)	(3)	(4)	(5)	(6)
	All		High Achievement		High Minority	
Variables	Mean	Std	Mean	Std	Mean	Std
Panel A: Middle School Characteristics						
Average Test Score (6th graders)	297.30	(20.52)	313.00	(17.72)	282.50	(11.82)
% Black or Hispanic	70.92	(30.51)	47.92	(30.82)	97.16	(2.74)
Panel B: High School Characteristics						
4yr Graduation Rate (%)	67.01	(17.34)	81.79	(10.81)	60.67	(16.19)
College Enrollment Rate (%)	58.38	(17.13)	73.96	(11.49)	52.23	(15.11)
Average Test Score (9th graders)	294.4	(17.49)	311.3	(14.36)	285.7	(13.37)
% Black or Hispanic	76.58	(23.40)	57.98	(26.10)	95.49	(04.20)
% From High Achievement MS	29.89	(26.68)	58.99	(24.80)	15.69	(17.40)
% From High Minority MS	32.96	(26.53)	15.62	(21.71)	52.78	(22.82)

Note: Average Test Score is the average of the mean of statewide ELA and Math scores in a given school.

Causal Effects of Middle School Attendance on High School Choice

- Provide evidence of middle schools' causal effects on high school choice and outcomes

Causal Effects of Middle School Attendance on High School Choice

- Provide evidence of middle schools' causal effects on high school choice and outcomes
- Specifically,

$$\text{Treatment effects of attending} \begin{cases} \text{high achievement MS} \\ \text{high minority MS} \end{cases} \text{on the} \begin{cases} \text{characteristics of ranked HS} \\ \text{characteristics of assigned HS} \end{cases}$$

Causal Effects of Middle School Attendance on High School Choice

- Provide evidence of middle schools' causal effects on high school choice and outcomes
- Specifically,

$$\text{Treatment effects of attending} \begin{cases} \text{high achievement MS} \\ \text{high minority MS} \end{cases} \text{on the} \begin{cases} \text{characteristics of ranked HS} \\ \text{characteristics of assigned HS} \end{cases}$$

- To overcome selection issue, use the quasi-experimental feature built in DA (Abdulkadiroğlu, Angrist, Narita and Pathak (2017, 2021, AANP)) AANP Identification

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. High school application: Yes

Main Findings

Q1. “Does a student’s previous school choice affect the subsequent school choices?”

1. High school application: Yes

- High achievement MS students **apply to** high schools w/ $\begin{cases} +1.8\text{pp}^{*} & \text{college enrollment} \\ +3.0\text{pp}^{**} & \text{\% of high performers} \end{cases}$

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. High school application: Yes

- High achievement MS students **apply to** high schools w/ $\begin{cases} +1.8\text{pp}^{*} & \text{college enrollment} \\ +3.0\text{pp}^{**} & \text{\% of high performers} \end{cases}$

2. High school assignment: Yes, and with larger magnitude

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. High school application: Yes

- High achievement MS students **apply to** high schools w/ $\begin{cases} +1.8\text{pp}^{*} & \text{college enrollment} \\ +3.0\text{pp}^{**} & \text{\% of high performers} \end{cases}$

2. High school assignment: Yes, and with larger magnitude

- High achievement MS students **are assigned to** high schools w/ $\begin{cases} +3.4\text{pp}^{*} & \text{college enrollment} \\ +5.3\text{pp}^{**} & \text{\% of high performers} \end{cases}$

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. High school application: Yes

- High achievement MS students **apply to** high schools w/ $\begin{cases} +1.8\text{pp}^{*} & \text{college enrollment} \\ +3.0\text{pp}^{**} & \text{\% of high performers} \end{cases}$

2. High school assignment: Yes, and with larger magnitude

- High achievement MS students **are assigned to** high schools w/ $\begin{cases} +3.4\text{pp}^{*} & \text{college enrollment} \\ +5.3\text{pp}^{**} & \text{\% of high performers} \end{cases}$
- Could be because treated students are also ranked higher by high schools $\rightarrow$ explored in the structural approach

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. High school application: Yes

- High achievement MS students **apply to** high schools w/ $\begin{cases} +1.8\text{pp}^{*} & \text{college enrollment} \\ +3.0\text{pp}^{**} & \text{\% of high performers} \end{cases}$

2. High school assignment: Yes, and with larger magnitude

- High achievement MS students **are assigned to** high schools w/ $\begin{cases} +3.4\text{pp}^{*} & \text{college enrollment} \\ +5.3\text{pp}^{**} & \text{\% of high performers} \end{cases}$
- Could be because treated students are also ranked higher by high schools $\rightarrow$ explored in the structural approach

Full Table | Heterogeneity | Mediation | High Minority | Sensitivity

Overview of Model Details

1st Period: Middle School Choice

2nd Period: High School Choice

Overview of Model

Details

1st Period: Middle School Choice

2nd Period: High School Choice

ROL

Priorities

→ Assignment

Overview of Model

Details

1st Period: Middle School Choice

2nd Period: High School Choice

Utilities

MS attended
HS Char.
Stu Char.
Distance
Unobs. Tastes

→ ROL

Priorities

→ Assignment

Overview of Model

Details

1st Period: Middle School Choice

Utilities

MS Char.
Stu Char.
Distance
Unobs. Tastes

ROL

Assignment

Priorities

2nd Period: High School Choice

Utilities

MS attended
HS Char.
Stu Char.
Distance
Unobs. Tastes

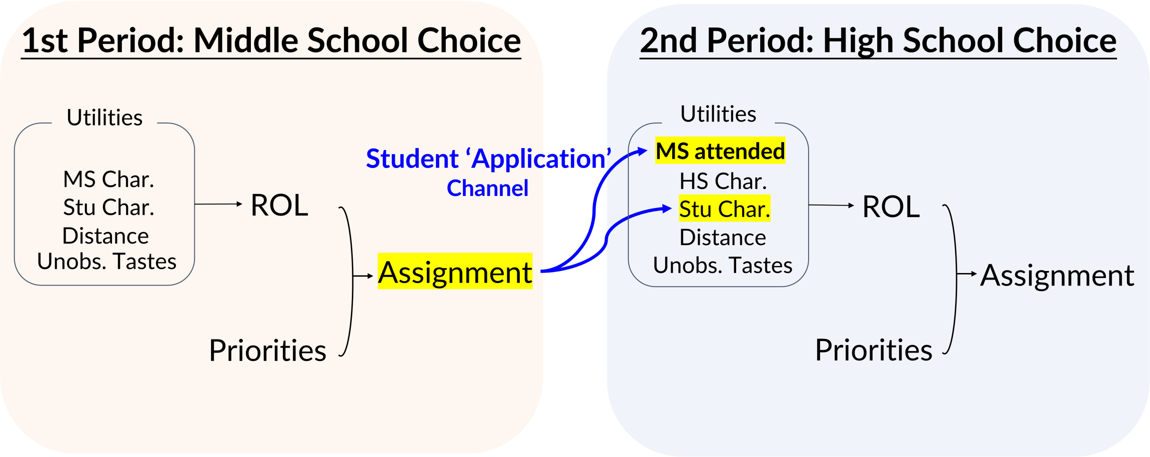
ROL

Assignment

Priorities

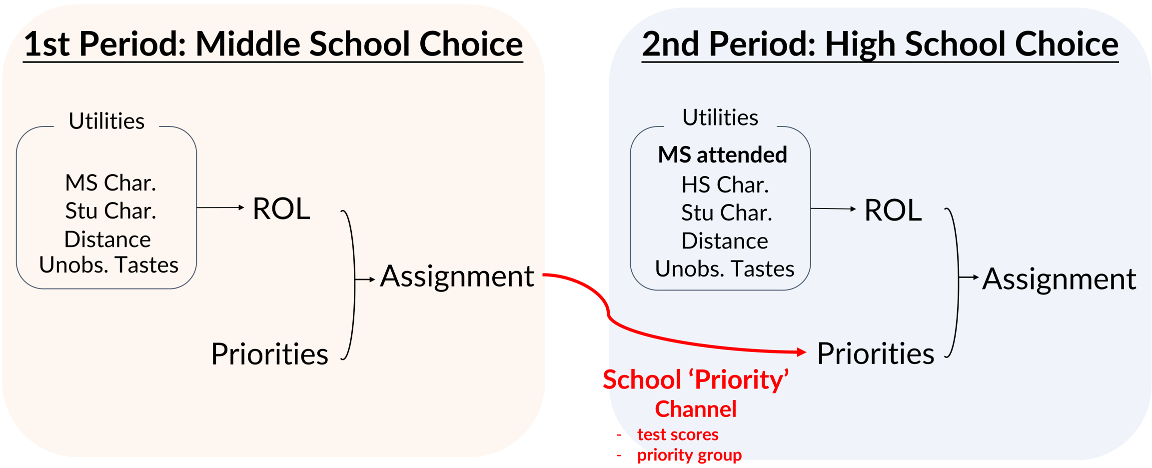
Overview of Model

Details

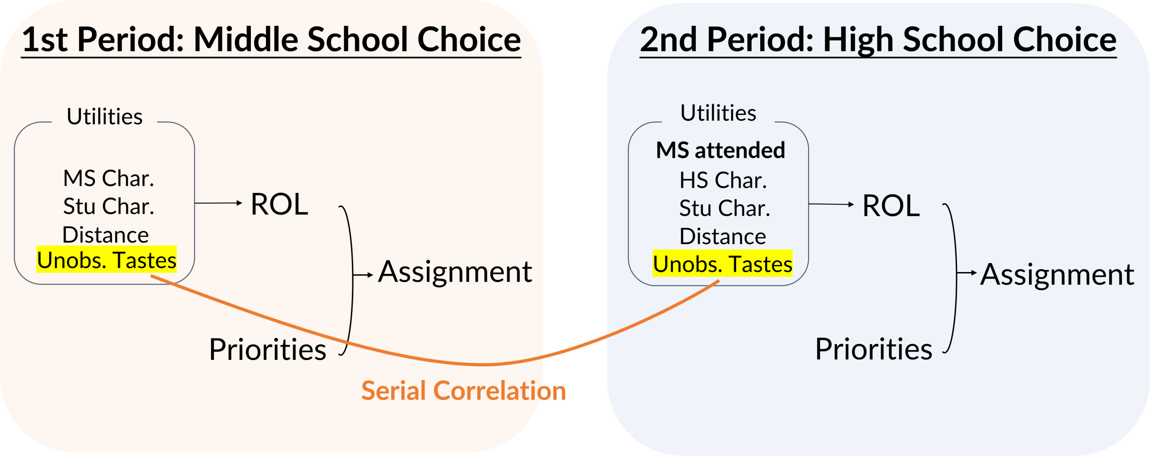


Overview of Model

Details

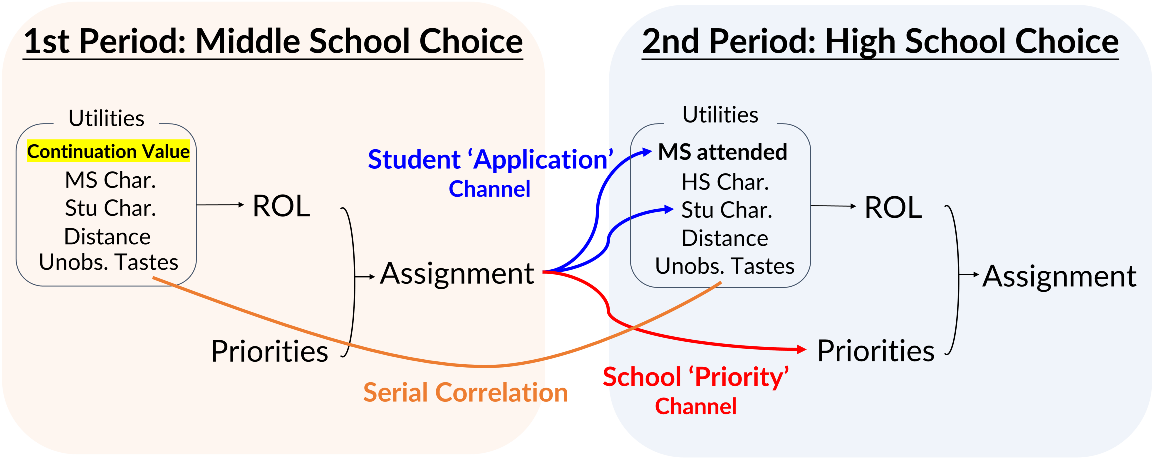


Overview of Model Details



Overview of Model

Details



Theoretical Framework: Second Period

Each forward looking player (student) i participates in school choices across two periods

Theoretical Framework: Second Period

Each forward looking player (student) i participates in school choices across two periods

In the **second period** (high school application),

Theoretical Framework: Second Period

Each forward looking player (student) i participates in school choices across two periods

In the **second period** (high school application),

- Student $i \in \mathcal{I}$ has utility V_{ij} from attending high school $j \in \mathcal{J}$:

$$V_{ij} = v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) + \eta_{ij}$$

Theoretical Framework: Second Period

Each forward looking player (student) i participates in school choices across two periods

In the **second period** (high school application),

- Student $i \in \mathcal{I}$ has utility V_{ij} from attending high school $j \in \mathcal{J}$:

$$V_{ij} = v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) + \eta_{ij}$$

  - $\tilde{X}_j$: HS char (% high performer, % White, 1(STEM))
    Z_i^H: student char (ethnicity, FRL, ELL, 8th grade test score)
    $m(i)$: i's attended middle school; $\tilde{d}_{ij}$: distance
    γ_i^H: unobserved taste on $\tilde{X}_j$; η_{ij}: idiosyncratic preference shock iid EVT1

Theoretical Framework: Second Period

Each forward looking player (student) i participates in school choices across two periods

In the **second period** (high school application),

- Student $i \in \mathcal{I}$ has utility V_{ij} from attending high school $j \in \mathcal{J}$:

$$V_{ij} = v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) + \eta_{ij}$$

  - $\tilde{X}_j$: HS char (% high performer, % White, 1(STEM))
    Z_i^H: student char (ethnicity, FRL, ELL, 8th grade test score)
    $m(i)$: i's attended middle school; $\tilde{d}_{ij}$: distance
    γ_i^H: unobserved taste on $\tilde{X}_j$; η_{ij}: idiosyncratic preference shock iid EVT1
  - Z_i^H may depend on $m(i)$ **e.g.** end of MS test scores VA

First Period

In the **first period** (middle school application),

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

  - X_m: MS char (% high performer, % White, 1(STEM))
    Z_i^M: student char (ethnicity, FRL, ELL, 5th grade test score); d_{im}: distance;
    γ_i^M: unobserved taste on X_m; ϵ_{im}: idiosyncratic preference shock iid EVT1

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times (\text{Continuation value of attending } m)$$

  - X_m: MS char (% high performer, % White, 1(STEM))
    Z_i^M: student char (ethnicity, FRL, ELL, 5th grade test score); d_{im}: distance;
    γ_i^M: unobserved taste on X_m; ϵ_{im}: idiosyncratic preference shock iid EVT1

Behavioral assumption:
In each period, students' ROLs are such that the resulting assignment outcomes are ex-post stable

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

  - X_m: MS char (% high performer, % White, 1(STEM))
    Z_i^M: student char (ethnicity, FRL, ELL, 5th grade test score); d_{im}: distance;
    γ_i^M: unobserved taste on X_m; ϵ_{im}: idiosyncratic preference shock iid EVT1

Behavioral assumption:
In each period, students' ROLs are such that the resulting assignment outcomes are ex-post stable

- DA produces cutoffs in terms of ex-post priority scores (=intrinsic priority group + tie-breaking lottery)

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

  - X_m: MS char (% high performer, % White, 1(STEM))
    Z_i^M: student char (ethnicity, FRL, ELL, 5th grade test score); d_{im}: distance;
    γ_i^M: unobserved taste on X_m; ϵ_{im}: idiosyncratic preference shock iid EVT1

Behavioral assumption:
In each period, students' ROLs are such that the resulting assignment outcomes are ex-post stable

- DA produces cutoffs in terms of ex-post priority scores (=intrinsic priority group + tie-breaking lottery)
- Ex-post stability: each student is assigned to the favorite, ex-post feasible school

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times (\text{Continuation value of attending } m)$$

  - X_m: MS char (% high performer, % White, 1(STEM))
    Z_i^M: student char (ethnicity, FRL, ELL, 5th grade test score); d_{im}: distance;
    γ_i^M: unobserved taste on X_m; ϵ_{im}: idiosyncratic preference shock iid EVT1

Behavioral assumption:
In each period, students' ROLs are such that the resulting assignment outcomes are ex-post stable

- DA produces cutoffs in terms of ex-post priority scores (=intrinsic priority group + tie-breaking lottery)
- Ex-post stability: each student is assigned to the favorite, ex-post feasible school
- Agnostic about exact strategies, consistent with truth-telling or more robust assumptions

Game

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

Forward-looking behavior:

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

Forward-looking behavior:

- $\omega \in \Omega$: source of uncertainty that determines ex-post scores and cutoffs e.g. lottery tie-breakers

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

Forward-looking behavior:

- $\omega \in \Omega$: source of uncertainty that determines ex-post scores and cutoffs e.g. lottery tie-breakers
  - $O_i(Z_i^H, m; \omega)$: i's ex-post feasible set of high schools when she attends m, given ω

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times \text{(Continuation value of attending } m\text{)}$$

Forward-looking behavior:

- $\omega \in \Omega$: source of uncertainty that determines ex-post scores and cutoffs e.g. lottery tie-breakers
  - $O_i(Z_i^H, m; \omega)$: i's ex-post feasible set of high schools when she attends m, given ω
- Ex-post stability implies:

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times (\text{Continuation value of attending } m)$$

Forward-looking behavior:

- $\omega \in \Omega$: source of uncertainty that determines ex-post scores and cutoffs e.g. lottery tie-breakers
  - $O_i(Z_i^H, m; \omega)$: i's ex-post feasible set of high schools when she attends m, given ω
- Ex-post stability implies:

$$(i\text{'s second period payoff when attending } m \text{ given } \omega) = \max_{j \in O_i(Z_i^H, m; \omega)} V_{ij}$$

First Period

In the **first period** (middle school application),

- Student $i \in \mathcal{I}$ has utility U_{im} from attending middle school $m \in \mathcal{M}$:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \delta \times E_{\gamma_i^H, \omega, \eta_i, Z_i^H}\left[\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij} \middle| Z_i^M, \gamma_i^M, \boldsymbol{\epsilon}_i, m\right]$$

Forward-looking behavior:

Available Info

- $\omega \in \Omega$: source of uncertainty that determines ex-post scores and cutoffs e.g. lottery tie-breakers
  - $O_i(Z_i^H, m; \omega)$: i's ex-post feasible set of high schools when she attends m, given ω
- Ex-post stability implies:

$$(i\text{'s second period payoff when attending } m \text{ given } \omega) = \max_{j \in O_i(Z_i^H, m; \omega)} V_{ij}$$

Parameterization: Preferences

A random coefficient model.

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = X_m' \beta_i^M - \lambda^M d_{im}$$
$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij}$$

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = X_m' \beta_i^M - \lambda^M d_{im}$$

$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij}$$

- Heterogeneous taste parameters:

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = X_m' \beta_i^M - \lambda^M d_{im}$$
$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij}$$

- Heterogeneous taste parameters:

$$\beta_{i,l}^M = \beta_{0,l}^M + Z_i^{M\prime} \beta_{Z,l}^M + \gamma_{i,l}^M$$

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = X_m' \beta_i^M - \lambda^M d_{im}$$

$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij}$$

- Heterogeneous taste parameters:

$$\beta_{i,l}^M = \beta_{0,l}^M + Z_i^{M\prime} \beta_{Z,l}^M + \gamma_{i,l}^M$$

$$\beta_{i,l}^H = \beta_{0,l}^H + \boxed{Z_i^{H\prime}(m(i))}\, \beta_{Z,l}^H + \gamma_{i,l}^H + \underbrace{\boxed{\sum_{\tau=1}^{T} \rho_{\tau,l} 1\left\{\tau(m(i)) = \tau\right\}}}_{\text{Middle school type effect}}$$

where $\tau(m(i))$ is the type of i's middle school $m(i)$

Parameterization: Preferences

A random coefficient model.

- Flow payoffs:

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = X_m' \beta_i^M - \lambda^M d_{im}$$

$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij}$$

- Heterogeneous taste parameters:

$$\beta_{i,l}^M = \beta_{0,l}^M + Z_i^{M\prime} \beta_{Z,l}^M + \gamma_{i,l}^M$$

$$\beta_{i,l}^H = \beta_{0,l}^H + \boxed{Z_i^{H\prime}(m(i))}\, \beta_{Z,l}^H + \gamma_{i,l}^H + \underbrace{\boxed{\sum_{\tau=1}^{T} \rho_{\tau,l} 1\left\{\tau(m(i)) = \tau\right\}}}_{\text{Middle school type effect}}$$

where $\tau(m(i))$ is the type of i's middle school $m(i)$

- Unobservable tastes:

$$\gamma_i^H = \boldsymbol{R_0} \gamma_i^M + \xi_i$$

Assumptions CV

Estimation

Estimation

- Identifying assumption: ex-post stability

Estimation

- Identifying assumption: ex-post stability
  - A student's assigned school is the most preferred one among the feasible schools
    $\Rightarrow$ conditional mixed logit model

Estimation

- Identifying assumption: ex-post stability
  - A student's assigned school is the most preferred one among the feasible schools
    $\Rightarrow$ conditional mixed logit model
- Estimation: Simulated Maximum Likelihood Estimator (SMLE) Likelihood Identification

Estimation

- Identifying assumption: ex-post stability
  - A student's assigned school is the most preferred one among the feasible schools $\Rightarrow$ conditional mixed logit model
- Estimation: Simulated Maximum Likelihood Estimator (SMLE) Likelihood Identification
- Focus on students and residents of Staten Island for tractability
  - Can be considered as an independent school district Staten Island

Estimation

- Identifying assumption: ex-post stability
  - A student's assigned school is the most preferred one among the feasible schools $\Rightarrow$ conditional mixed logit model
- Estimation: Simulated Maximum Likelihood Estimator (SMLE) Likelihood Identification
- Focus on students and residents of Staten Island for tractability
  - Can be considered as an independent school district Staten Island
  - 2,626 students applying to 20 middle schools and 47 high schools

Main Findings

Q1. “Does a student’s previous school choice affect the subsequent school choices?”

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics

Main Findings

Q1. “Does a student’s previous school choice affect the subsequent school choices?”

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance
      High achievement MS makes students willing to travel $+$ 0.11 miles for 10pp $\uparrow$ of % high performers

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance
      High achievement MS makes students willing to travel $+$ 0.11 miles for 10pp $\uparrow$ of % high performers
   b. Tastes on racial composition

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
    a. Tastes on academic performance

       High achievement MS makes students willing to travel + 0.11 miles for 10pp ↑ of % high performers

    b. Tastes on racial composition

       High minority MS makes students willing to travel — 0.26 miles for 10pp ↑ of % White

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance
      High achievement MS makes students willing to travel + 0.11 miles for 10pp ↑ of % high performers

   b. Tastes on racial composition
      High minority MS makes students willing to travel − 0.26 miles for 10pp ↑ of % White

→ Segregation has a **reinforcing** effect

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance
      High achievement MS makes students willing to travel + 0.11 miles for 10pp ↑ of % high performers
   b. Tastes on racial composition
      High minority MS makes students willing to travel − 0.26 miles for 10pp ↑ of % White

   → Segregation has a **reinforcing** effect

2. Application channel is quantitatively more important than priority channel

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance
      High achievement MS makes students willing to travel + 0.11 miles for 10pp ↑ of % high performers

   b. Tastes on racial composition
      High minority MS makes students willing to travel − 0.26 miles for 10pp ↑ of % White

   → Segregation has a **reinforcing** effect

2. Application channel is quantitatively more important than priority channel
   - 2/3 of middle schools' effects operate through the student application channel

Main Findings

Q1. "Does a student's previous school choice affect the subsequent school choices?"

1. Middle schools affect how students value different high school characteristics
   a. Tastes on academic performance

      High achievement MS makes students willing to travel + 0.11 miles for 10pp ↑ of % high performers

   b. Tastes on racial composition

      High minority MS makes students willing to travel − 0.26 miles for 10pp ↑ of % White

   → Segregation has a **reinforcing** effect

2. Application channel is quantitatively more important than priority channel
   - 2/3 of middle schools' effects operate through the student application channel

Full Estimates | By Race | Scatter Plot | Model Fit | Decomposition

Racial Segregation in NYC

- NYC public schools are racially segregated
  - 2017-18: About 69% of public school students are Black or Hispanic
  - **However**, at half of high schools, % of Black or Hispanic exceeds 90% Segregation

Racial Segregation in NYC

- NYC public schools are racially segregated
  - 2017-18: About 69% of public school students are Black or Hispanic
  - **However**, at half of high schools, % of Black or Hispanic exceeds 90% Segregation

- Ongoing policy reforms in NYC:
  1. Eliminate geography-based priority rules
  2. Eliminate selecting students based on test scores

Racial Segregation in NYC

- NYC public schools are racially segregated
  - 2017-18: About 69% of public school students are Black or Hispanic
  - **However**, at half of high schools, % of Black or Hispanic exceeds 90% Segregation

- Ongoing policy reforms in NYC:
  1. Eliminate geography-based priority rules
  2. Eliminate selecting students based on test scores

- We evaluate effects of **two actual policies combined**, **implemented at different timings**:
  1. **MS**: reform only middle school admissions
  2. **HS**: reform only high school admissions
  3. **MSHS**: reform both middle and high school admissions

Racial Gap in Characteristics of Co-assigned Peers

Racial Gap in Characteristics of Co-assigned Peers

1. **MS** can desegregate middle schools

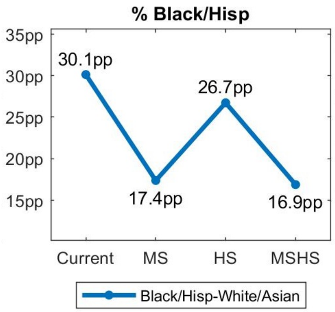
   Racial gap in the proportion of Black/Hispanic peers: $\downarrow$ 42% for MS

Racial Gap in Characteristics of Co-assigned Peers

1. MS can desegregate middle schools

Racial gap in the proportion of Black/Hispanic peers: ↓ 42% for MS

Middle Schools

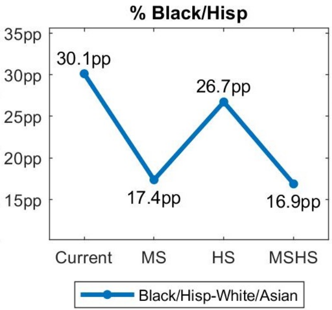


Racial Gap in Characteristics of Co-assigned Peers

1. **MS** can desegregate middle schools and high schools

Racial gap in the proportion of Black/Hispanic peers: ↓ 42% for MS, ↓ 13% for HS

Middle Schools

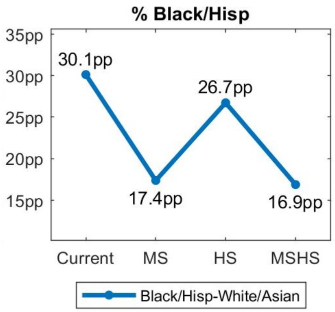


Racial Gap in Characteristics of Co-assigned Peers

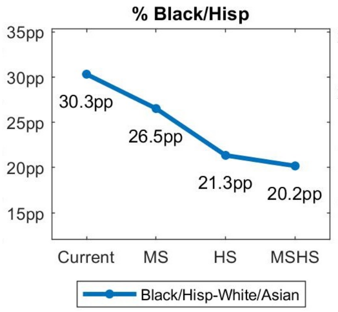
1. MS can desegregate middle schools and high schools

Racial gap in the proportion of Black/Hispanic peers: ↓ 42% for MS, ↓ 13% for HS

Middle Schools



High Schools

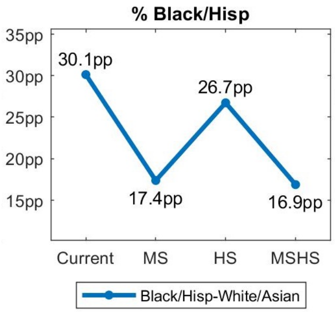


Racial Gap in Characteristics of Co-assigned Peers

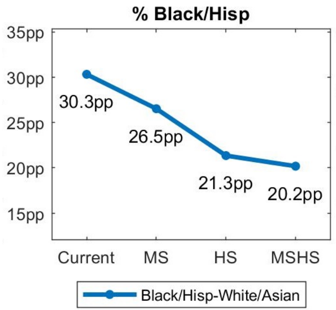
1. MS can desegregate middle schools and high schools

Racial gap in the proportion of Black/Hispanic peers: ↓ 42% for MS, ↓ 13% for HS

Middle Schools



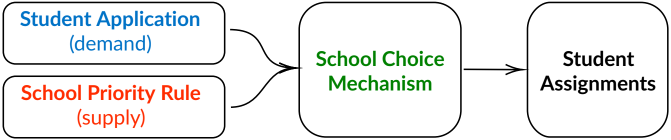
High Schools



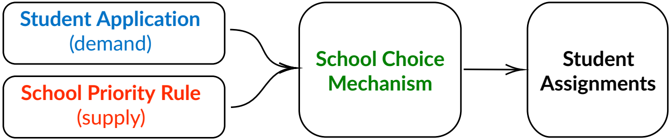
2. MSHS can improve on **HS** for desegregating high schools

Char Gap

Policy Implication

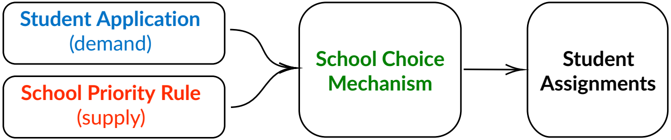


Policy Implication



- Student-school matching is an equilibrium outcome:
  - Determined by both demand (*student applications*) and supply (*school priorities*)

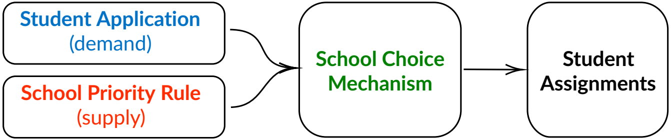
Policy Implication



- Student-school matching is an equilibrium outcome:
  - Determined by both demand (*student applications*) and supply (*school priorities*)

- Most desegregation policies focused on reforming the supply side i.e., how schools select students

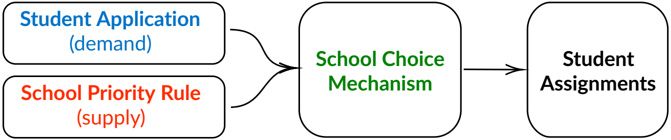
Policy Implication



- Student-school matching is an equilibrium outcome:
  - Determined by both demand (*student applications*) and supply (*school priorities*)

- Most desegregation policies focused on reforming the supply side i.e., how schools select students
  - Little attention to how we can influence the demand side

Policy Implication



- Student-school matching is an equilibrium outcome:
  - Determined by both demand (*student applications*) and supply (*school priorities*)

- Most desegregation policies focused on reforming the supply side i.e., how schools select students
  - Little attention to how we can influence the demand side

- We suggest that:
  - ***Early*** intervention on the supply side can alter ***subsequent*** demand side behaviors
  - Such dynamic relationship can be used to design more effective policies

Conclusion

- The first to examine the **dynamic relationship** of school choices
  1. Empirical evidence of middle schools' effects on high school applications/assignments
  2. A novel dynamic framework of school choice
     - Middle schools' effects are mainly by changing student applications to high schools
  3. A new perspective on how to understand/address segregation across public schools

Conclusion

- The first to examine the **dynamic relationship** of school choices
  1. Empirical evidence of middle schools' effects on high school applications/assignments
  2. A novel dynamic framework of school choice
     - Middle schools' effects are mainly by changing student applications to high schools
  3. A new perspective on how to understand/address segregation across public schools

- We open a new avenue of research in the school choice literature by
  1. Bringing the dynamic aspect of school choice to the front
  2. Providing a new framework that is applicable to many other topics in school choice

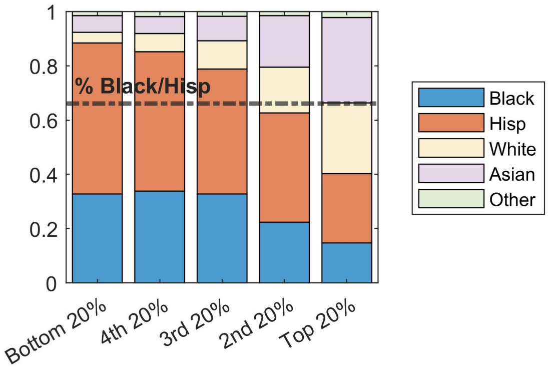
Thank you!

Comments or Suggestions?

dongwooh@usc.edu | mpark88@wisc.edu

Segregation in NYC High Schools CF

Black and Hispanic students are underrepresented in 'good' high schools



Note: Quintiles based on average performance at statewide exams.

Focusing on High School is Insufficient

e.g. NYC college enrollment 2016-17

	(1) Enrolled in College	(2) Enrolled in College	(3) Enrolled in College
Regents Score	0.141***	0.146***	0.139***
	(0.005)	(0.005)	(0.005)
HS Quality	0.028***		0.017***
	(0.007)		(0.006)
MS Quality		0.042***	0.034***
		(0.007)	(0.006)
N	51672	50942	50851
R2	0.160	0.162	0.163

Robust standard errors in parentheses. $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

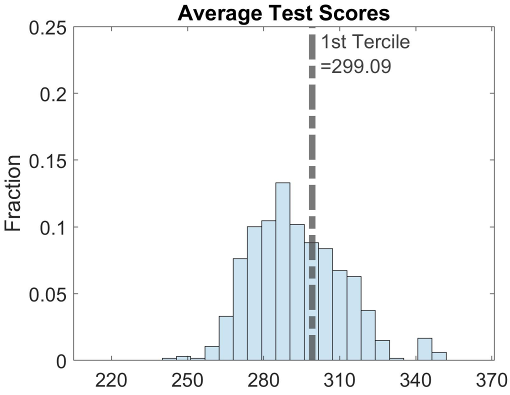
Middle school quality is more highly correlated with student's college outcome than high school quality!

Student Proposing Deferred Acceptance

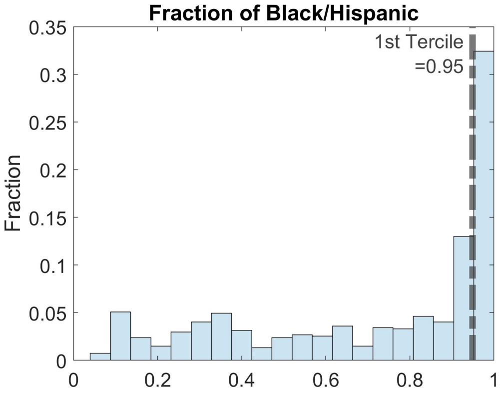
- **Step 1**
  Each student proposes to her first choice. Each school tentatively assigns seats to its proposers one at a time, following their priority order. The student is rejected if no seats are available at the time of consideration.
- **Step *k***
  Each student who was rejected in the previous step proposes to her next best choice. Each school considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the school's priority order. The student is rejected if no seats are available when she is considered.
- The algorithm terminates either when there are no new proposals or equally when all rejected students have exhausted their preference lists.

Middle School Groups

High Achievement Middle Schools

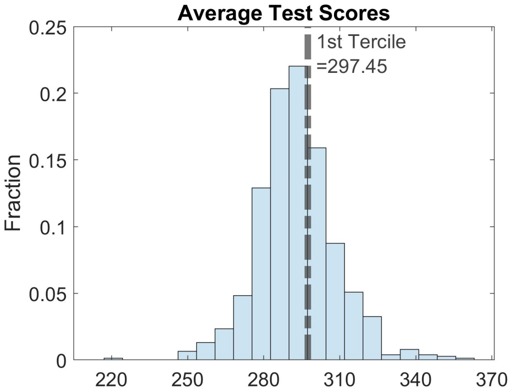


High Minority Middle Schools

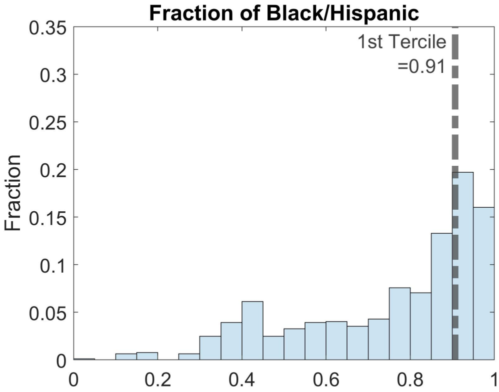


High School Groups

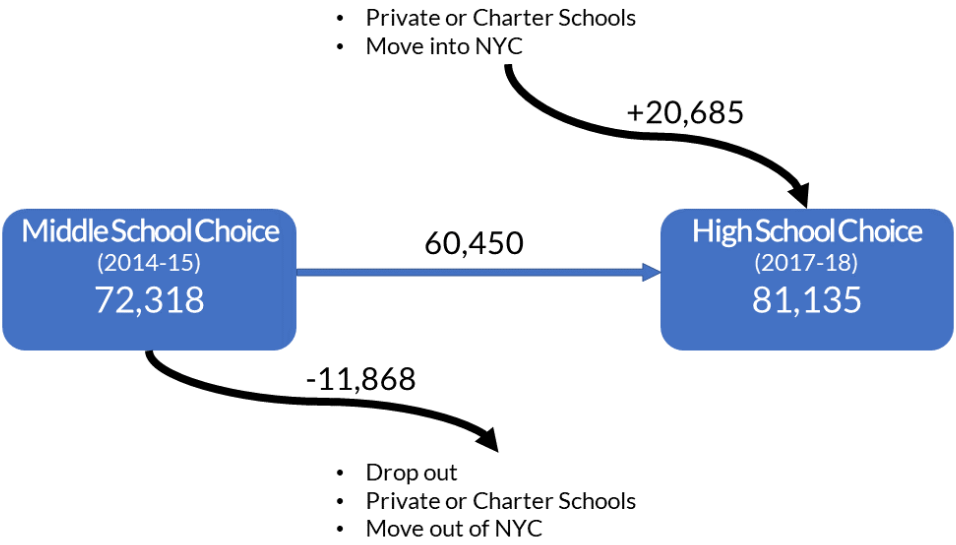
High Achievement High Schools



High Minority High Schools

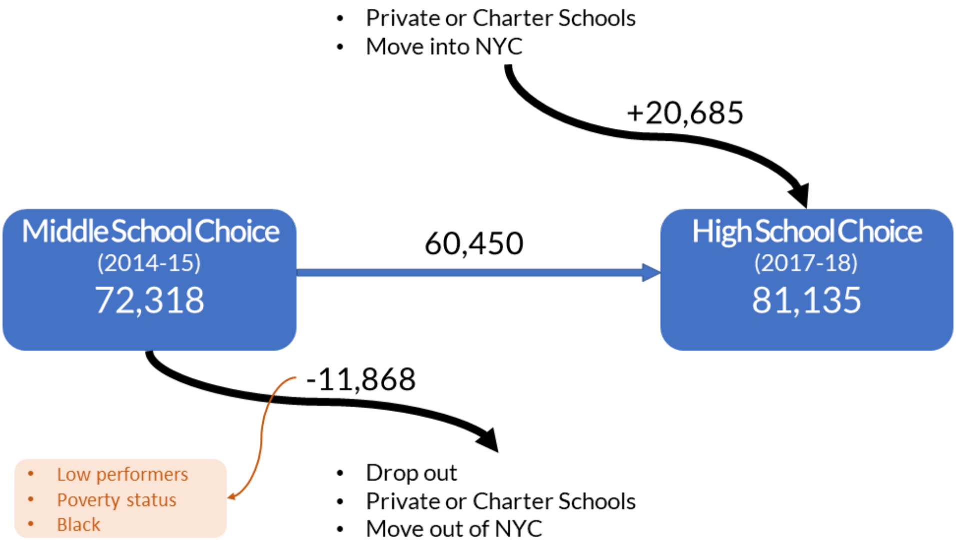


Attrition



Private or Charter Schools
Move into NYC
+20,685
Middle School Choice
(2014-15)
72,318
60,450
High School Choice
(2017-18)
81,135
-11,868
Drop out
Private or Charter Schools
Move out of NYC

Attrition



Private or Charter Schools
Move into NYC
+20,685
Middle School Choice (2014-15) 72,318
60,450
High School Choice (2017-18) 81,135
-11,868
Low performers
Poverty status
Black
Drop out
Private or Charter Schools
Move out of NYC

Summary Statistics: Students

Variables	N	Mean	Std
5th Grade Math Score	54,012	311.3	37.31
English Language Learner	54,012	0.12	0.32
Free or Reduced Lunch	54,012	0.73	0.45
Asian	54,012	0.19	0.39
Black	54,012	0.23	0.42
Hispanic	54,012	0.41	0.49
White	54,012	0.17	0.37

Note: The scale of 5th grade math score is from 125 to 402.

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities
  - Conditional on applications and priorities, assignment is random

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities
  - Conditional on applications and priorities, assignment is random
  - **But** full conditioning is impossible

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities
  - Conditional on applications and priorities, assignment is random
  - **But** full conditioning is impossible

- DA maps those (applications, priorities) into DA propensity score (=admission probability)

Quasi-experimental Variation in School Assignment

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities
  - Conditional on applications and priorities, assignment is random
  - **But** full conditioning is impossible

- DA maps those (applications, priorities) into DA propensity score (=admission probability)

- AANP formally shows that:

  At each school, conditional on DA propensity score, assignments of applicants are **random**

Quasi-experimental Variation in School Assignment ◂

Imagine that all schools use lotteries for tie-breaking

- Tie-breaker determines admissions of applicants with the same applications (=ROL) and priorities
  - Conditional on applications and priorities, assignment is random
  - **But** full conditioning is impossible

- DA maps those (applications, priorities) into DA propensity score (=admission probability)

- AANP formally shows that:

  At each school, conditional on DA propensity score, assignments of applicants are **random**
  $\Longrightarrow$ **'stratified randomized trial'**

AANP | Propensity Score

AANP 2017, 2021

- Abdulkadiroğlu, Angrist, Narita and Pathak (2017,2021):
  1. Formally prove **Conditional Independence**
     $\Rightarrow$ eliminate OVB by conditioning on DA propensity score
  2. Provide a compact way of calculating DA propensity scores in a general framework with non-random tie-breaking, by combining RD and large-market matching model

Calculating Propensity Score

Two cases:

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking
   - Even conditional on applications and priorities, assignments are no longer random

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking
   - Even conditional on applications and priorities, assignments are no longer random
   - **Solution**: non-parametric RD framework

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
    - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking
    - Even conditional on applications and priorities, assignments are no longer random
    - **Solution**: non-parametric RD framework
        - In the small neighborhood around cutoffs, applicants have constant risk of clearing cutoffs of 1/2

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking
   - Even conditional on applications and priorities, assignments are no longer random
   - **Solution**: non-parametric RD framework
     - In the small neighborhood around cutoffs, applicants have constant risk of clearing cutoffs of $1/2 \Rightarrow$ *as good as random* assignment

Calculating Propensity Score

Two cases:

1. *Unscreened* programs: pure lottery tie-breaking
   - With lotteries, propensity score depends on only a few school-level cutoffs
2. *Screened* programs: non-random tie-breaking
   - Even conditional on applications and priorities, assignments are no longer random
   - **Solution**: non-parametric RD framework
     - In the small neighborhood around cutoffs, applicants have constant risk of clearing cutoffs of $1/2 \Rightarrow$ *as good as random* assignment
     - For those, **local conditional independence** holds

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\ \underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}\ ]$

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\ \underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}\ ]$
- Priority structure: $\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group} \in \mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker} \in [0,1]}$

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}]$
- Priority structure: $\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group} \in \mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker} \in [0,1]}$
  - Tie-breaking rule: $\text{TB}_{iA} = \text{TB}_{iB} \sim U[0,1]$, $\text{TB}_{iC} \sim F_i$. F_i is unknown

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}]$
- Priority structure: $\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group} \in \mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker} \in [0,1]}$
  - Tie-breaking rule: $\text{TB}_{iA} = \text{TB}_{iB} \sim U[0, 1]$, $\text{TB}_{iC} \sim F_i$. F_i is unknown
  - Priority group lexicographically dominates tie-breakers

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}]$
- Priority structure: $\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group} \in \mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker} \in [0,1]}$
  - Tie-breaking rule: $\text{TB}_{iA} = \text{TB}_{iB} \sim U[0,1]$, $\text{TB}_{iC} \sim F_i$. F_i is unknown
  - Priority group lexicographically dominates tie-breakers
  - Admitted to school j if $\text{score}_{ij} > \text{cutoff}_j$ **and at the same time** rejected from all schools ranked above j

Calculating Propensity Score: An Example

- Consider student i who submits $ROL_i = [\underbrace{A \succ B}_{unscreened} \succ \underbrace{C}_{screened}]$
- Priority structure: $\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group} \in \mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker} \in [0,1]}$
  - Tie-breaking rule: $\text{TB}_{iA} = \text{TB}_{iB} \sim U[0,1]$, $\text{TB}_{iC} \sim F_i$. F_i is unknown
  - Priority group lexicographically dominates tie-breakers
  - Admitted to school j if $\text{score}_{ij} > \text{cutoff}_j$ **and at the same time** rejected from all schools ranked above j

	A	B	C
PG_{ij}	1	1	2
Cutoff	2.2	1.4	2.6
Admission Probability	0	1×0.6	$1 \times 0.4 \times (1 - F_i(0.6))$
Local Admission Probability	0	1×0.6	$1 \times 0.4 \times \mathbf{0.5}$

Identification Strategy

Identification Strategy

- 2SLS using DA assignment as an IV for actual attendance:

Identification Strategy

- 2SLS using DA assignment as an IV for actual attendance:

$$Y_i = \alpha_0 + \boxed{\beta} C_i + \sum_x \alpha_2(x) d_i(x) + \eta_i$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + \nu_i$$

Identification Strategy

- 2SLS using DA assignment as an IV for actual attendance:

$$Y_i = \alpha_0 + \boxed{\beta} C_i + \sum_x \alpha_2(x) d_i(x) + \eta_i$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + \nu_i$$

  - Y_i: HS application/assignment, C_i: treatment (attending some group of MS)
  - D_i: DA assignment into treatment schools
  - $\{d_i(x)\}_x$: DA propensity score fixed effects

Identification Strategy ◄

- 2SLS using DA assignment as an IV for actual attendance:

$$Y_i = \alpha_0 + \boxed{\beta} C_i + \sum_x \alpha_2(x) d_i(x) + \eta_i$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + \nu_i$$

  - Y_i: HS application/assignment, C_i: treatment (attending some group of MS)
  - D_i: DA assignment into treatment schools
  - $\{d_i(x)\}_x$: DA propensity score fixed effects

- Exclusion restriction: DA assignments D_i have no effect on outcomes Y_i other than by affecting the actual attendance C_i after controlling for DA propensity scores Balance

Identification Strategy ◂

- 2SLS using DA assignment as an IV for actual attendance:

$$Y_i = \alpha_0 + \boxed{\beta} C_i + \sum_x \alpha_2(x) d_i(x) + \eta_i$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + \nu_i$$

  - Y_i: HS application/assignment, C_i: treatment (attending some group of MS)
  - D_i: DA assignment into treatment schools
  - $\{d_i(x)\}_x$: DA propensity score fixed effects

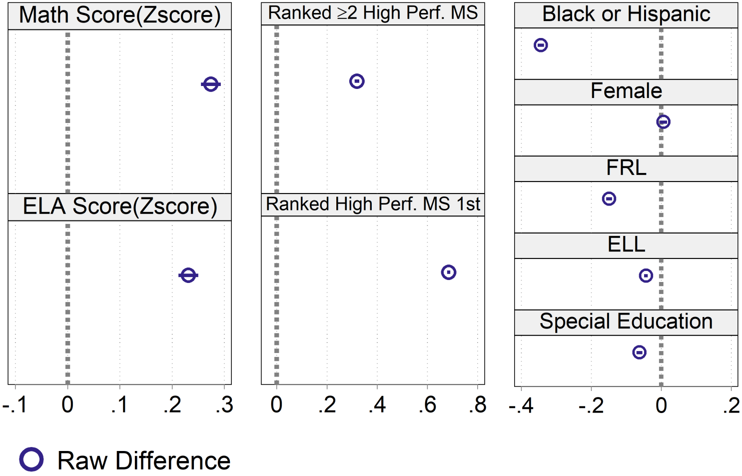
- Exclusion restriction: DA assignments D_i have no effect on outcomes Y_i other than by affecting the actual attendance C_i after controlling for DA propensity scores Balance
- Sample restriction: those with nondegenerate risk of being treated Marginal NDR-DR

Covariate Balance: Offered Students v.s. Non-offered Students

$W_i = \alpha_0 + \gamma D_i + e_i$

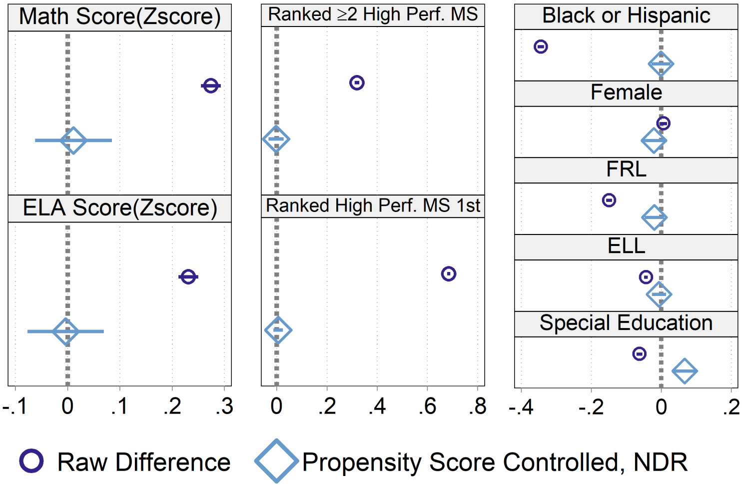
Covariate Balance: Offered Students v.s. Non-offered Students

$$W_i = \alpha_0 + \gamma D_i + e_i$$



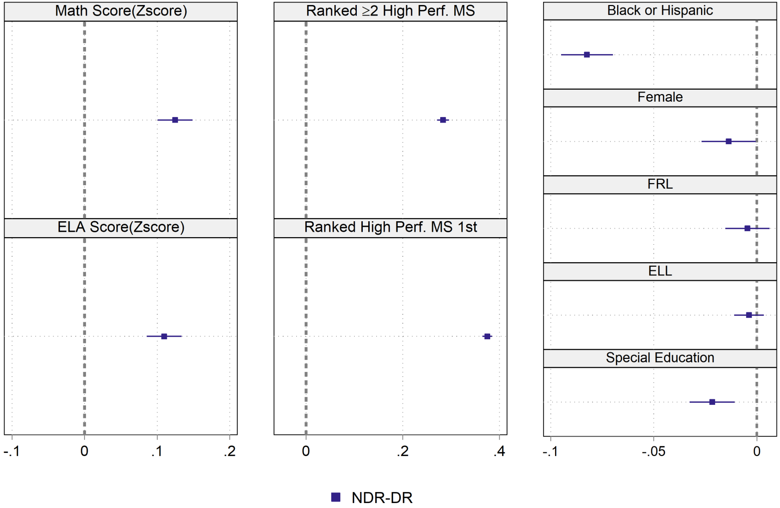
Covariate Balance: Offered Students v.s. Non-offered Students

$$W_i = \alpha_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + h(\mathcal{R}_i) + e_i$$



Nondegenerate v.s. Degenerate Risk Sample

Group mean difference of NDR and DR samples



Who Are Those Marginal Students?

Variables	All		Marginal to High-achievement MS		Marginal to High-minority MS	
	Mean	Std	Mean	Std	Mean	Std
5th Grade Math Score	311.3	37.3	313.5	35.3	292.0	32.6
English Language Learner	0.12	0.32	0.07	0.25	0.12	0.32
Free or Reduced Lunch	0.73	0.45	0.76	0.43	0.90	0.30
Asian	0.19	0.39	0.22	0.42	0.02	0.15
Black	0.23	0.42	0.17	0.38	0.34	0.47
Hispanic	0.41	0.49	0.43	0.50	0.62	0.49
White	0.17	0.37	0.16	0.36	0.01	0.10

Note: The scale of 5th grade math score is from 125 to 402.

Main Result

Dependent Variable	Avg of Top 5 Ranked		Assigned	
Model	OLS	2SLS	OLS	2SLS
Sample	All	NDR	All	NDR
Panel A: College Enrollment Rate (%p)				
From High Achievement MS	2.854***	1.755*	4.530***	3.414**
	(0.516)	(1.011)	(0.669)	(1.566)
N	44158	7060	41546	6679
R2	0.367	0.459	0.244	0.310
$\bar{y}$	71.217	72.197	65.653	67.204
Panel B: % High Performing Students (%p)				
From High Achievement MS	5.188***	2.986*	6.886***	5.292**
	(0.840)	(1.805)	(0.825)	(2.105)
N	44237	7062	42180	6751
R2	0.450	0.502	0.388	0.400
$\bar{y}$	39.731	40.934	33.058	34.978
First Stage F-stat		135.2		135.2
Student Obs. Char.	✓	✓	✓	✓
Local Linear Control	✓	✓	✓	✓

Note: Standard errors clustered at attended middle school level.

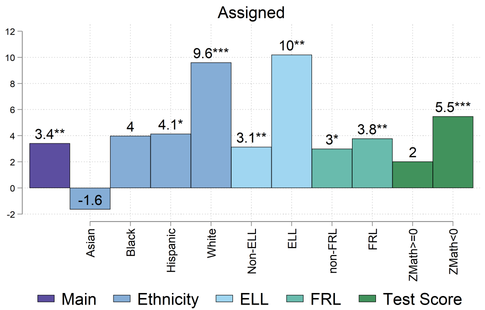
Main Result

Dependent Variable	Avg of Top 5 Ranked		Assigned	
Model	OLS	2SLS	OLS	2SLS
Sample	All	NDR	All	NDR
Panel C: % White (%p)				
From High Achievement MS	5.080***	0.311	5.755***	0.301
	(0.750)	(0.655)	(0.793)	(0.832)
N	44237	7062	42180	6751
R2	0.633	0.717	0.555	0.621
$\bar{y}$	18.627	20.334	15.097	16.761
Panel D: 1(STEM)				
From High Achievement MS	-0.053***	0.041	-0.057***	0.055
	(0.013)	(0.035)	(0.016)	(0.044)
N	44237	7062	42182	6751
R2	0.098	0.275	0.041	0.172
$\bar{y}$	0.324	0.318	0.314	0.322
First Stage F-stat		135.2		135.2
Student Obs. Char.	✓	✓	✓	✓
Local Linear Control	✓	✓	✓	✓

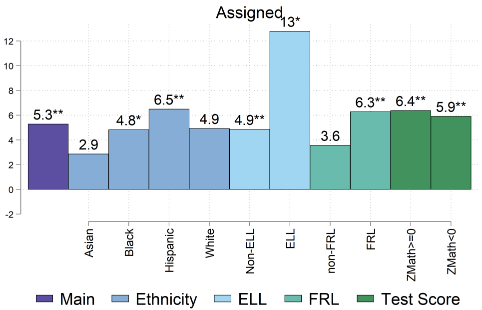
Note: Standard errors clustered at attended middle school level.

2SLS: Subgroup

College Enrollment Rate



% High Performing Students



- In general, more effective for groups with smaller baseline (with the exception for White)

  e.g. ELL, FRL, lower baseline math score

2SLS: Mediation Analysis

Is the effect mainly due to change in test scores? $\Rightarrow$ additionally include end-of-MS test scores

	Top 5 2SLS	Assigned 2SLS
Panel A: College Enrollment Rate (%p)		
From High Perf. MS	1.751*	3.301**
	(0.967)	(1.542)
8th Grade ELA Score (σ)	1.314***	2.070***
	(0.205)	(0.328)
8th Grade Math Score (σ)	0.910***	1.416***
	(0.231)	(0.374)
N	7060	6679
Panel B: % High Performing Students (%p)		
From High Perf. MS	2.913*	5.185**
	(1.748)	(2.061)
8th Grade ELA Score (σ)	2.114***	3.023***
	(0.351)	(0.409)
8th Grade Math Score (σ)	1.258***	1.315**
	(0.397)	(0.522)
N	7062	6751

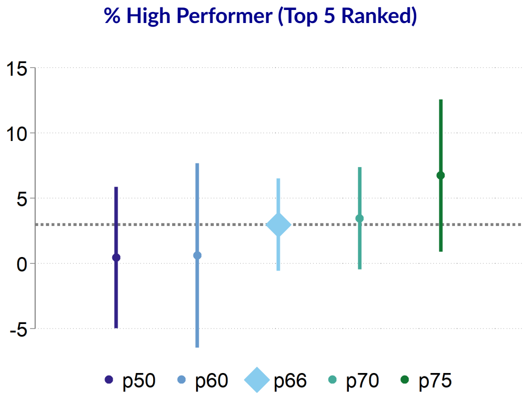
While coefficients on 8th grade test scores are significantly positive, LATEs are largely unchanged

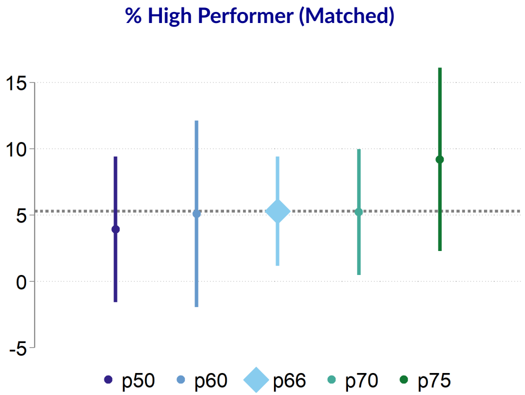
2SLS: Alternative Treatment: High Minority

	(1)	(2)	(3)	(4)
Dependent Variable	Top 5	Top 5	Assigned	Assigned
Model	OLS	2SLS	OLS	2SLS
Sample	All	NDR	All	NDR
Panel A: College Enrollment Rate (%p)				
From High Minority MS	-1.686***	0.248	-2.189***	-0.794
	(0.553)	(1.459)	(0.661)	(2.383)
N	46630	3307	43843	3091
R2	0.363	0.358	0.237	0.260
$\bar{y}$	71.371	66.679	65.829	60.183
Panel B: % High Performing Students (%p)				
From High Minority MS	-4.024***	3.188	-3.875***	3.957*
	(0.850)	(2.084)	(0.800)	(2.240)
N	46723	3317	44579	3163
R2	0.441	0.370	0.376	0.333
$\bar{y}$	39.839	28.252	33.146	21.158

2SLS: Sensitivity to Other Definitions of Treatments

Are the identified LATEs sensitive to the choice of threshold (top 1/3)?





Value-added Estimation ◂

- Let $y_{i,m}^{H}$: potential end-of-MS test score when i attends middle school m
- Estimate each m's production function based on selection on observables:

$$E\left[y_{i,m}^{H} \mid Z_i^M, m\right] = \alpha_m + Z_i^{M\prime}\beta_m$$

- OLS of $y_{i,m(i)}^{H}$ on school indicators interacted with Z_i^M where $m(i)$ is the actual middle school attendance in the data

Value-added Estimation

Mean and Standard Deviation of VA Coefficients Across Schools

	8th Grade Math		8th Grade ELA	
	Coefficient	SE	Coefficient	SE
Baseline Test Score	0.346	0.035	0.331	0.033
	(0.060)	(0.015)	(0.040)	(0.013)
Female	1.591	1.650	3.077	1.517
	(1.425)	(0.412)	(2.327)	(0.352)
Asian	6.002	3.993	6.029	3.402
	(4.892)	(2.108)	(4.617)	(1.547)
Black	-2.422	4.542	-2.502	4.642
	(6.194)	(2.527)	(3.826)	(3.216)
Hispanic	-2.309	2.708	-0.738	2.472
	(3.945)	(1.260)	(3.391)	(1.008)
English Language Learner	-2.862	5.691	1.239	6.066
	(7.230)	(2.669)	(6.273)	(3.045)
Student with Disability	-6.885	2.345	-5.571	2.212
	(3.192)	(0.690)	(2.122)	(0.663)
Free or Subsidized Lunch	-1.380	2.264	-1.501	2.013
	(2.124)	(1.190)	(1.974)	(0.863)

A Game of Incomplete Information

Che, Hahm and He (2022)

- Schools $\{c_1, ..., c_C\}$ with seats $S = (S_1, \cdots, S_C) \in \mathbb{N}^C$
- k students, each with an *ex-ante* type $\theta = (u, q) \in \Theta$ with distribution η
  - $u = (u_1, ..., u_C) \in [\underline{u}, \overline{u}]^C$: utility at each school
  - $q \in \mathcal{Q}$: "intrinsic" priorities at the schools **e.g.** priority groups in NYC
  - ex-post scores $s \in [0, 1]^C$: distribution $G_q(s)$
    - schools rank students by ex-post scores in admissions
    - **e.g.** lottery tie-breaking: s **=** q **+ lottery**
- Private information: student type $\theta = (u, q)$; Common knowledge: DA, seats S, distributions η & $G_q(s)$
- A game of incomplete information: strategy is a measureable function $\sigma_i : \Theta \rightarrow \Delta(\mathcal{R})$
  - $\mathcal{R}$: set of all possible ROLs

Robust Equilibrium and Its Stability

Definition

An infinite profile σ is a **robust equilibrium** if, for any $\epsilon > 0$, there exists $K \in \mathbb{N}$ such that for $k > K$, its k-truncation σ^k is an interim ϵ-Bayes Nash equilibrium.

- Namely, for i, σ_i gives student i of each type θ a payoff within ϵ of the highest possible (i.e., supremum) payoff she can get by using any strategy when all the others employ σ^k_{-i}

Theorem (Stability Theorem)

Any regular robust equilibrium is asymptotically stable.

- **Asymptotic stability**: as $k \to \infty$ (the economy becomes large)

  The fraction of students assigned their most-preferred feasible schools $\xrightarrow{p} 1$

  - given any realized state of the world (e.g., realization of the tie-breaking lottery)

What is Known to the Student? ◀

	Unobserved Taste on School Char.		Idiosyncratic Preference Shock		Program Characteristics	Student's own Characteristics	Uncertainty in High School Choice
	γ_i^M	γ_i^H	ϵ_{im}	η_{ij}	$X_m, \tilde{X}_j$	Z_i^M, Z_i^H	ω
1st Period (MSAP)	✓		✓		✓	✓	
2nd Period (HSAP)	✓	✓	✓	✓	✓	✓	

Distributional Assumptions

- Assumptions: let $\Psi_{1i} = (Z_i^M, \gamma_i^M, \epsilon_i, m)$

$$\eta_{ij} \perp \epsilon_{im} \Big| \gamma_i^M, \xi_i, \quad \forall i, j, m \tag{1}$$

$$\xi_i, \eta_i, \Psi_{1i} \text{ are mutually independent}, \quad \forall i \tag{2}$$

$$\omega \perp (\xi_i, \eta_{ij}) \Big| \Psi_{1i} \text{ and } \omega \perp \Psi_{1i}, \quad \forall i, j, m \tag{3}$$

- Idiosyncratic preferences:
  ϵ_{im} and η_{ij} both follow EVT1

- Unobservable tastes:
  $\gamma_i^M \perp \xi_i$, $\gamma_i^M \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\gamma)$, $\xi_i \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\xi)$

Continuation Value ◂

Based on the assumptions on the unobservables,

$$E_{\gamma_i^H,\omega,\boldsymbol{\eta}_i,Z_i^H}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|Z_i^M,\gamma_i^M,\boldsymbol{\epsilon}_i,m\right]$$
$$=\int_\omega\int_{\xi_i}\left(\mu+\log\left(\sum_{j\in O_i(Z_i^H,m;\omega)}\exp(v_{ij}(\xi_i))\right)\right)d\Phi(\xi_i|\Sigma_\xi)dH(\omega)$$

where

- $v_{ij}=V_{ij}-\eta_{ij}$
- $\Phi(\cdot|\Sigma)$: cdf of $\mathcal{N}(0,\Sigma)$
- $H(\cdot)$: cdf of ω

Source of Identification

1. Nonparametric identification of utilities:

$$(\gamma_i^M, \epsilon_{im}) \perp d_{im} \Big| X_m, Z_i^M$$

$$(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \Big| \tilde{X}_j, Z_i^H, m(i)$$

+ additive separability of distance (Agarwal and Somaini 2018)

Source of Identification

1. Nonparametric identification of utilities:

$$(\gamma_i^M, \epsilon_{im}) \perp d_{im} \Big| X_m, Z_i^M$$

$$(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \Big| \tilde{X}_j, Z_i^H, m(i)$$

   + additive separability of distance (Agarwal and Somaini 2018)

2. Popularity of schools with certain characteristics, for students with certain characteristics $\Rightarrow \beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H$

Source of Identification

1. Nonparametric identification of utilities:

$$(\gamma_i^M, \epsilon_{im}) \perp d_{im} \Big| X_m, Z_i^M$$

$$(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \Big| \tilde{X}_j, Z_i^H, m(i)$$

   + additive separability of distance (Agarwal and Somaini 2018)

2. Popularity of schools with certain characteristics, for students with certain characteristics $\Rightarrow \beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H$
3. Persistent pattern on preference for some school characteristics $\Rightarrow$ $\boldsymbol{R_0}$

Source of Identification

1. Nonparametric identification of utilities:

$$(\gamma_i^M, \epsilon_{im}) \perp d_{im} \Big| X_m, Z_i^M$$

$$(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \Big| \tilde{X}_j, Z_i^H, m(i)$$

+ additive separability of distance (Agarwal and Somaini 2018)

2. Popularity of schools with certain characteristics, for students with certain characteristics $\Rightarrow \beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H$
3. Persistent pattern on preference for some school characteristics $\Rightarrow \boldsymbol{R_0}$
4. Common popularity of schools with certain characteristics among students attending same type of middle school + **quasi-random variation in MS assignments by tie-breaking** $\Rightarrow \rho_\tau$

Likelihood ◂

For student i, conditional on γ_i^M, ξ_i,

$$\begin{aligned} P_i(\theta, \gamma_i^M, \xi_i) &= P(\text{observe } m_i, j_i | \gamma_i^M, \xi_i, \theta) \\ &= P\left(\begin{array}{c} U_{im_i} = \max_{m \in O_i^m} U_{im} \quad \text{and} \\ V_{ij_i} = \max_{j \in O_i^h} V_{ij} \text{ given } m_i \end{array} \middle| \gamma_i^M, \xi_i, \theta \right) \\ &= \frac{\exp(u_{im_i}(\gamma_i^M, \theta))}{\sum_{m \in O_i^m} \exp(u_{im}(\gamma_i^M, \theta))} \frac{\exp(v_{ij_i}(\gamma_i^M, \xi_i, \theta; m_i))}{\sum_{j \in O_i^h} \exp(v_{ij}(\gamma_i^M, \xi_i, \theta; m_i))} \end{aligned}$$

Then,

$$P_i(\theta) = \int_{\gamma_i^M} \int_{\xi_i} P_i(\theta, \gamma_i^M, \xi_i) \phi(\xi_i | \Sigma_\xi) \phi(\gamma_i^M | \Sigma_\gamma) d\xi_i d\gamma_i^M$$

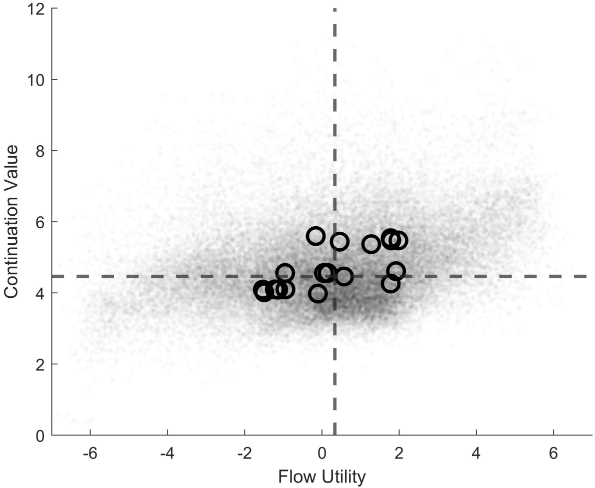
where $\phi(\cdot|\Sigma)$ is the pdf of a multivariate normal with mean zero and covariance matrix Σ, and hence

$$\sum_i \log P_i(\theta)$$

is the log-likelihood function

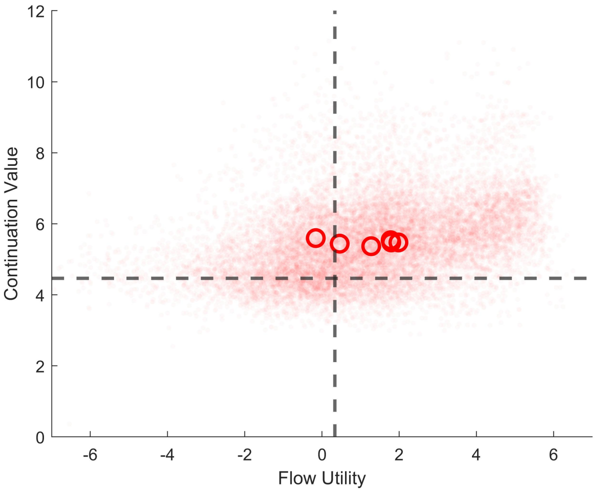
Flow Utility and Continuation Value

All Middle Schools



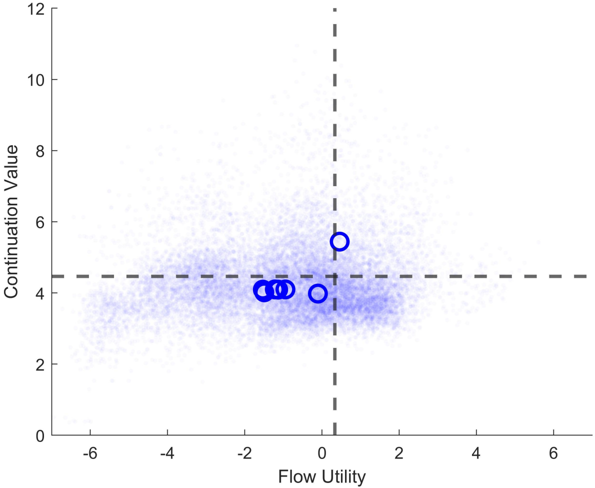
Flow Utility and Continuation Value

High Achievement Middle Schools



Flow Utility and Continuation Value

High Minority Middle Schools



Goodness-of-Fit ◂

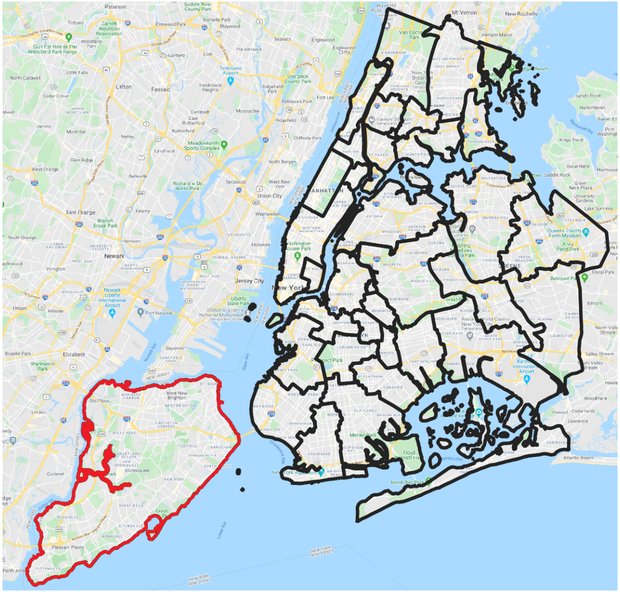
Goodness-of-Fit ◂

Characteristics of assigned schools by student type:

	Middle Schools				High Schools			
	% High Performer		% FRL		% High Performer		% FRL	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian	36%	37%	62%	60%	33%	33%	58%	58%
Black	27%	32%	74%	69%	25%	24%	69%	70%
Hispanic	31%	34%	67%	65%	29%	27%	64%	66%
White	45%	45%	48%	49%	39%	37%	50%	52%
ELL	27%	30%	71%	67%	24%	24%	70%	69%
FRL	35%	37%	62%	61%	31%	30%	61%	62%

More

Staten Island



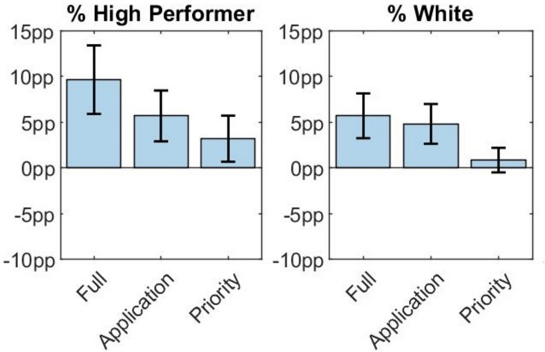
Staten Island

Affluent county with more White and better performing students compared to citywide average

	All NYC		Staten Island	
	Mean	Std	Mean	Std
5th Grade Math Score	311.26	(37.31)	315.49	(32.40)
5th Grade ELA Score	300.56	(34.96)	306.13	(31.10)
English Language Learner	0.12		0.05	
Disability	0.21		0.25	
Free or Reduced Lunch	0.73		0.54	
Asian	0.19		0.09	
Black	0.23		0.10	
Hispanic	0.41		0.24	
White	0.17		0.56	
N	54017		2626	

Decomposition of Middle School Effects

Change in Characteristics of Assigned High School (Lowest- → Highest Performing MS)



Full Estimates ◂

	Middle Schools est	Middle Schools se		High Schools est	High Schools se	
Panel A: Preference Estimates						
Fraction of High Performer	4.944	(1.144)	***	0.795	(0.272)	***
Asian	-1.267	(1.947)		0.827	(0.39)	**
Black	6.82	(1.961)	***	-0.199	(0.462)	
Hisp	1.781	(1.288)		-0.275	(0.33)	
Poverty	-0.881	(1.13)		-0.922	(0.271)	***
ELL	-1.804	(2.309)		0.342	(1.177)	
5th Gr Test Score	1.088	(0.581)	*	1.652	(0.141)	***
Fraction of White	3.056	(0.875)	***	4.931	(0.343)	***
Asian	0.976	(1.588)		-2.011	(0.599)	***
Black	-6.444	(1.721)	***	-1.52	(0.613)	**
Hisp	-1.666	(1.047)		-1.06	(0.421)	**
Poverty	-0.565	(0.886)		0.162	(0.346)	
ELL	0.752	(1.954)		-0.24	(1.202)	
5th Gr Test Score	-0.951	(0.468)	**	0.341	(0.126)	***
1(STEM)	0.281	(0.198)		-0.676	(0.123)	***
Asian	0.157	(0.324)		-0.174	(0.2)	
Black	-0.42	(0.269)		0.09	(0.196)	
Hisp	0.121	(0.213)		0.083	(0.144)	
Poverty	-0.122	(0.198)		0.257	(0.126)	**
ELL	0.062	(0.345)		1.005	(0.326)	***
5th Gr Test Score	-0.159	(0.096)	*	0.003	(0.044)	

Full Estimates

	Middle Schools		High Schools		
	est	se	est	se	
Panel B: Middle School Type Effects					
Type 1 (High Achievement MS)					
Fraction of High Performer			0.546	(0.276)	**
Fraction of White			1.6	(0.318)	***
1(STEM)			-0.322	(0.137)	**
Type 2 (High Minority MS)					
Fraction of High Performer			0.875	(0.301)	***
Fraction of White			-1.447	(0.378)	***
1(STEM)			0.198	(0.136)	

Full Estimates ◀

	Middle Schools			High Schools		
	est	se		est	se	
Panel C: Unobservable Tastes						
ρ_0				0.074	(0.044)	*
				0.429	(0.127)	***
				-0.035	(0.118)	
(1,1) of Σ_γ	18.461	(10.853)	*			
(1,2)	-17.93	(9.653)	*			
(1,3)	-0.186	(1.626)				
(2,2)	23.168	(10.222)	**			
(2,3)	2.765	(2.018)				
(3,3)	1.163	(0.697)	*			
(1,1) of Σ_ξ				0.447	(0.316)	
(1,2)				-2.184	(0.95)	**
(1,3)				0.411	(0.163)	**
(2,2)				10.67	(2.877)	***
(2,3)				-2.006	(0.512)	***
(3,3)				0.377	(0.193)	*

Full Estimates ◄

	Middle Schools			High Schools		
	est	se		est	se	
Panel D: Other Parameters						
Outside option	2.698	(0.367)	***	-0.371	(0.175)	**
Distance	0.655	(0.038)	***	0.509	(0.018)	***
Discount Factor	0.877	(0.064)	***			

MS Effects by Race

White/Asian

High achievement MS makes students willing to travel $\begin{cases} +\ 0.12 \text{ miles} \\ +\ 0.41 \text{ miles} \end{cases}$ for 10pp increase of $\begin{cases} \text{\% high performer} \\ \text{\% White} \end{cases}$

High minority MS makes students willing to travel $\begin{cases} +\ 0.24 \text{ miles} \\ -\ 0.34 \text{ miles} \end{cases}$ for 10pp increase of $\begin{cases} \text{\% high performer} \\ \text{\% White} \end{cases}$

Black/Hispanic

High achievement MS makes students willing to travel $\begin{cases} +\ 0.12 \text{ miles} \\ +\ 0.08 \text{ miles} \end{cases}$ for 10pp increase of $\begin{cases} \text{\% high performer} \\ \text{\% White} \end{cases}$

High minority MS makes students willing to travel $\begin{cases} +\ 0.11 \text{ miles} \\ -\ 0.22 \text{ miles} \end{cases}$ for 10pp increase of $\begin{cases} \text{\% high performer} \\ \text{\% White} \end{cases}$

Goodness-of-Fit

Characteristics of assigned students by school type:

	Middle Schools				High Schools			
	High Achievement		High Minority		High Achievement		High Minority	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian (%)	9%	9%	9%	9%	9%	9%	7%	8%
Black (%)	4%	3%	25%	24%	4%	4%	30%	23%
Hispanic (%)	12%	12%	41%	39%	18%	15%	42%	40%
White (%)	74%	75%	25%	27%	68%	71%	20%	28%
ELL (%)	2%	1%	10%	9%	3%	3%	11%	9%
FRL (%)	41%	40%	77%	73%	46%	44%	78%	74%
5th grade Math	322.6	322.6	304.2	307.5	320.0	322.3	301.9	303.7
From High Achievement MS (%)					57%	61%	10%	9%
From High Minority MS (%)					10%	10%	62%	49%

Note: The scale of 5th grade math score is from 125 to 402.

Goodness-of-Fit

Assignment and ROL prediction:

	Dynamic Model	
	MS Application	HS Application
Panel A. Simulated versus observed assignment (100 simulated samples)		
Mean predicted fraction of students assigned to observed assignments	0.5709 (0.0053)	0.2022 (0.0049)
Panel B. Predicted versus observed partial preference order		
Mean predicted probability that a student's partial preference order among the programs in her ROL coincides with the submitted rank order	0.3848	0.1395

Benchmark for assignment prediction (Panel A):

- Lower bound (random assignment): 5.9% (MSAP), 2.3% (HSAP)

Decomposition Exercise

Specifically,

1. Assign students to *School A*, and **one student at a time**, change the assignment to *School B*
2. Keep track of how the student's high school assignment change in alternative scenarios
   i. Both channels are active
   ii. Turn off the priority channel i.e., priorities don't depend on middle schools
   iii. Turn off the application channel i.e., applications don't depend on middle schools

Racial Gap in Assigned High School Characteristics