Centralised School Assignment and Information

BREAD-IGC Virtual PhD Course: Education

Kehinde Ajayi

February 24, 2022

Motivation

Key questions

- How does information affect choices and outcomes?

Key questions

1. What type of information to provide?
   - School attributes
   - Probability of assignment
2. How to provide it?
   - Paper form
   - TV/video/radio
   - Online platform/text message
   - Social networks
3. Who to provide it to?
   - Students? Parents?
4. When to provide it?
   - Early on? Closer to deadline for choice submission?

Empirical analysis

1. Administrative data
2. Survey data
3. Exogenous variation
4. Structural estimation
5. Qualitative data

Empirical analysis

1. **Administrative data**
2. Survey data
3. Exogenous variation
4. Structural estimation
5. Qualitative data

Administrative data

- Advantages
  1. Automatically collected through the school choice process
  2. Covers the universe of students and schools

- Limitations
  1. Don't observe detailed background characteristics
  2. Don't observe beliefs and preferences

Ajayi (forthcoming)

- **Question**: What are the barriers to educational mobility?
- **Context**: Computerized School Selection and Placement System (CSSPS) for secondary school admission in Ghana
- **Approach**: Model the school choice problem; Empirically analyze heterogeneity in student demand

Ghana's Education System

- **Three levels of schooling**
  - 6 years Primary School (70% completion)
  - 3 years Junior High School (60% completion)
  - 3 years Secondary School (25% completion)
- **Computerised School Selection and Placement System**
  - **Feb**: Students submit ranked list of up to 6 secondary schools
  - **June**: Take Basic Education Certification Exam (BECE)
  - **Sept**: Assigned to schools based on BECE score + ranked list
- **Complications in application decision**
  - Apply *before* taking BECE
  - Can only submit a limited number of choices
  - Little information on admission standards, academic performance

Model

Setup

1. set of students $I = \{1, \cdots, K\}$ with unknown ability T_i^*
2. set of schools $S = \{1, \cdots, M\}$ with known quality threshold q_s

Each student i

1. assigns utility U_{is} to attending school s
2. has a subjective admission probability, $\Pr(T_i > q_s) \equiv \tilde{p}_{is} \in [0, 1)$
3. has an expected value of applying, $z_{is} = \tilde{p}_{is} U_{is}$
4. chooses an application portfolio, $A_i = \{1, \cdots, N\}$
5. faces cost $c(|A|)$ for applying to a portfolio of size $|A|$ schools

$$c(|A|) = \begin{cases} 0 & \text{if } |A| \leq n \\ \infty & \text{if } |A| > n \end{cases}$$

School Choice Problem

Student solves:

$$\max_{A_i \subseteq S} f(A_i) = \tilde{p}_{i1} U_{i1} + (\tilde{p}_{i2} - \tilde{p}_{i1}) U_{i2} + \cdots + (\tilde{p}_{iN} - \tilde{p}_{iN-1}) U_{iN}, \qquad N \leq n$$

- Students have 2655 alternatives and can submit 6 choices
- Identifying the optimal portfolio requires choosing between $2,655 \times 2,654 \times 2,653 \times 2,652 \times 2,651 \times 2,650 =$

$$348,282,630,021,183,832,064$$

alternative portfolios for each student

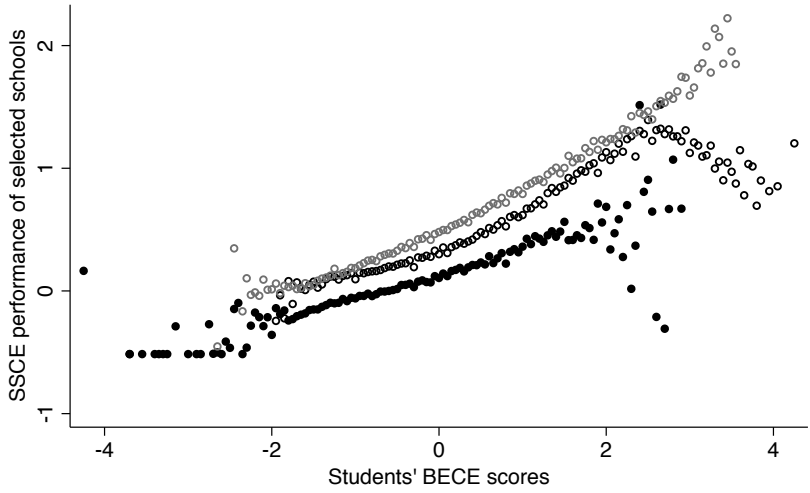
Data

Administrative data from 2005 to 2009

1. Students: 350,000 each year
   - background information (age, sex, junior high school)
   - ranked list of selected choices
   - standardized BECE score
   - admission outcome

2. Schools: 9,000 junior highs, 650 senior highs
   - basic characteristics (location, age, single sex, public, etc)
   - selectivity
   - academic performance (SSCE exam)

Observed choices



Empirical analysis

1. Administrative data
2. **Survey data**
3. Exogenous variation
4. Structural estimation
5. Qualitative data

Survey data

- Advantages
  1. Can collect detailed background information
  2. Can directly ask about beliefs, preferences, and behavior

- Limitations
  1. More costly (not automatic)
  2. Self-report might be biased and might not reflect actual behavior and outcomes

Empirical analysis

1. Administrative data
2. Survey data
3. **Exogenous variation**
4. Structural estimation
5. Qualitative data

Exogenous variation

- Advantages
  1. Identify causal effects of a change in information
- Limitations
  1. External validity
  2. Understanding underlying mechanisms

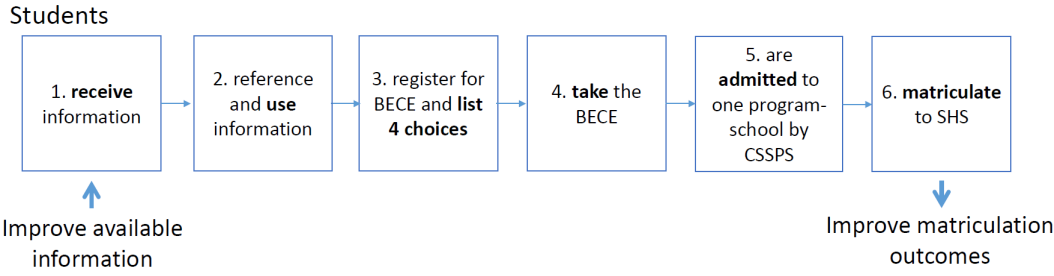
Exogenous variation

1. Randomized experiment
   - Ajayi, Friedman, and Lucas (2020)
2. Regression Discontinuity Design
   - Dustan (2018)
3. Regression Discontinuity Design + Randomized experiment
   - Arteaga, Kapor, Neilson, and Zimmerman (forthcoming)

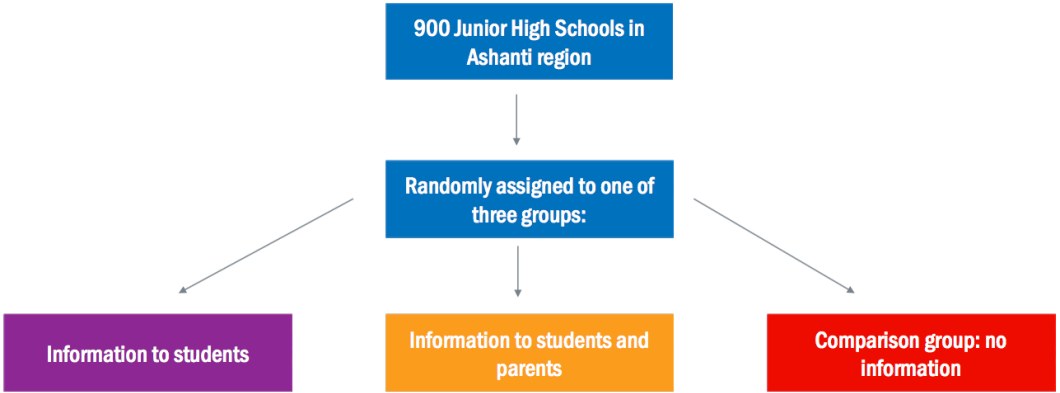
Ajayi, Friedman, and Lucas (2020)

- **Questions**: 1) What information do parents and students say they want? 2) Does receiving information affect school choices and educational outcomes? 3) Does who receives this information matter?
- **Context**: Computerized School Selection and Placement System (CSSPS) for secondary school admission in Ghana
- **Approach**: Randomized assignment to receive information booklet, video screening, Q and A workshop
- **Results**: Students whose parents were directly targeted were admitted to higher value added schools and were more likely to be attending the school to which they were admitted, conditional on attending any school.

Conceptual Framework



Research Design



Timeline and Data

Timeline:

- Academic year 2015-2016: intervention and data collection (JHS3)
- Academic year 2016-2017: data collection (SHS1)

Data:

- Administrative data—42,228 students across 900 schools
- Survey data (3 rounds)—11,454 students and 9000 guardians across 450 schools

Information was Received (Step 1)

Students saw the booklet and REALLY saw the video

	Saw Booklet	Saw Video	Believe CSSPS fair
Student Info	0.140***	0.780***	0.048
	(0.017)	(0.018)	(0.030)
Parent Info	0.151***	0.762***	0.032
	(0.017)	(0.021)	(0.029)
Student = Parent p-value	0.14	0.39	0.60
Observations	11692	11663	11671
R^2	0.10	0.59	0.06
Control Group Mean	0.82	0.14	0.53

- Other (less informative) booklets were around–may have REALLY seen our booklet.
- 50% attendance at parent meetings

Information was Referenced (Step 2)

Information shifted towards booklet and away from other sources

Table: Source of Information

	Booklet (1)	Radio/TV Newspaper (2)	Internet (3)	People (4)	Other (5)
Student Info	0.168***	-0.060***	-0.032*	-0.052***	-0.008*
	(0.021)	(0.017)	(0.017)	(0.013)	(0.005)
Parent Info	0.173***	-0.045***	-0.012	-0.035***	-0.012***
	(0.020)	(0.016)	(0.017)	(0.013)	(0.004)
Student = Parent p-value	0.804	0.381	0.231	0.159	0.315
Observations	11679	11678	11678	11678	11679
R^2	0.056	0.020	0.035	0.026	0.011
Control Group Mean	0.211	0.681	0.344	0.199	0.032

Parents also more likely to be involved (according to both parents and students).

Intervention Changed Application Characteristics (Step 3)

Students

- Were more likely to select schools **in Ashanti** (4 pp)
- Selected schools with marginally **easier admissions** (6 percent of a SD)
- Selected schools with **equivalent WASSCE scores**
- **Did not improve strategies**

Admission Outcomes (Step 5)

Did not change the likelihood of being admitted overall, but

Table: Admissions Characteristics

	Admitted Ashanti (1)	Admitted SHS BECE mean (2)	Admitted SHS Value Added (3)
Student Info	0.027	-0.050	0.033
	(0.019)	(0.030)	(0.025)
Parent Info	0.042***	-0.069**	0.086***
	(0.015)	(0.031)	(0.023)
Student = Parent p-value	0.35	0.49	0.03
Observations	42577	29914	29164
R^2	0.24	0.04	0.08
Control Group Mean	0.75	0.02	0.00

Admitted to schools with **lower admissions standards** and **higher value added** and maybe better matches on margins we cannot see.

Matriculation (Step 6)

Matriculation decisions are **revealed preferences** and **match quality**...

And...conditional on attending more likely to go to their assigned school in parent arm

Table: Matriculation

	Currently Attending SHS (1)	Started on Time if Attending (2)	Attending Admitted SHS (3)	Attending Admitted (if attending) (4)	Attending SHS BECE mean (5)
Student Info	-0.008	-0.018	-0.003	0.002	0.010
	(0.023)	(0.024)	(0.018)	(0.025)	(0.035)
Parent Info	-0.024	-0.040	0.033	0.075***	0.083*
	(0.022)	(0.025)	(0.020)	(0.028)	(0.043)
Student = Parent p-value	0.52	0.40	0.09	0.01	0.11
Observations	7654	3751	7087	3956	3591
R^2	0.06	0.07	0.03	0.02	0.13
Control Group Mean	0.60	0.44	0.23	0.41	0.08

Targeting students and parents

Intervention affected the intermediate steps of the causal chain

1. Students and parents wanted information
2. Students and parents received information
3. Students acted on information—changed types of schools to which they applied
4. Students admitted to higher value added schools, closer to home

But...

5. Information at most marginally improved match quality
   - Students no more likely to matriculate at all, on time, or to their admitted school
   - In parent arm: conditional on attending, more likely to attend admitted school

Dustan (2018)

- **Question**: How do family networks affect school choice?
- **Context**: centralised public high school choice system in Mexico City (COMIPEMS), assigns students to schools based on listed preferences and standardized exam score
- **Approach**: Regression discontinuity design (RDD) to estimate effects of older siblings' admission outcomes on younger siblings' choices and outcomes
- **Results**: Older siblings' school assignments strongly affect younger siblings' listed preferences and assignment outcomes

Data and Background

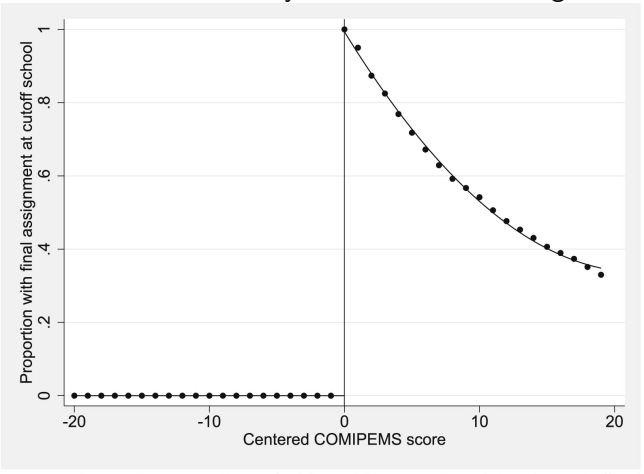
Administrative data: compiled by COMIPEMS for each student registered for exam from 1998 to 2011 (14 cycles)

1. demographic information (name, date of birth, phone number, address, middle school identifier, grade point average)
2. context survey (parental education, family composition)
3. list of up to 20 ranked school preferences
4. assignment results (exam score and school assigned)

Sample: sibling pairs where older sibling attended a public middle school and each sibling took exam at the end of middle school.

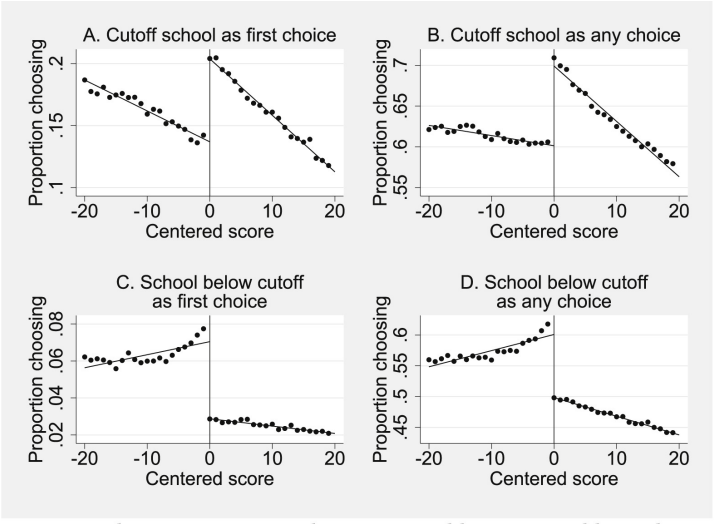
Empirical Strategy

compare older siblings wanting to attend a certain school, some score barely enough to be assigned there, others score barely too low and are assigned to another school



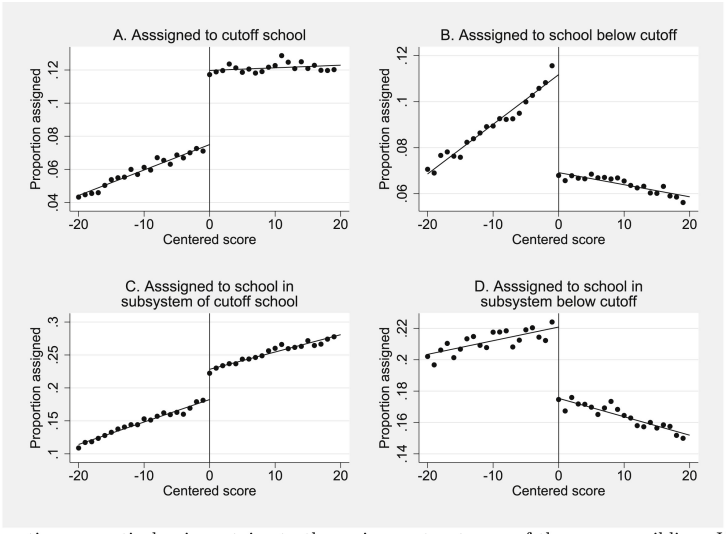
Note. Variable on vertical axis is proportion of older siblings assigned to the cutoff school. Variable on horizontal axis is older sibling's COMIPEMS exam score, centered to be 0 at the corresponding cutoff score.

Younger siblings' choices



Note. Proportion on vertical axis pertains to the younger siblings. Variable on horizontal axis is older sibling's COMIPEMS exam score, centered to be 0 at the corresponding cutoff score. Fitted lines are from a linear fit.

Younger siblings' assignment



Note. Proportion on vertical axis pertains to the assignment outcome of the younger sibling. Variable on horizontal axis is older sibling's COMIPEMS exam score, centered to be 0 at the corresponding cutoff score. Fitted lines are from a linear fit.

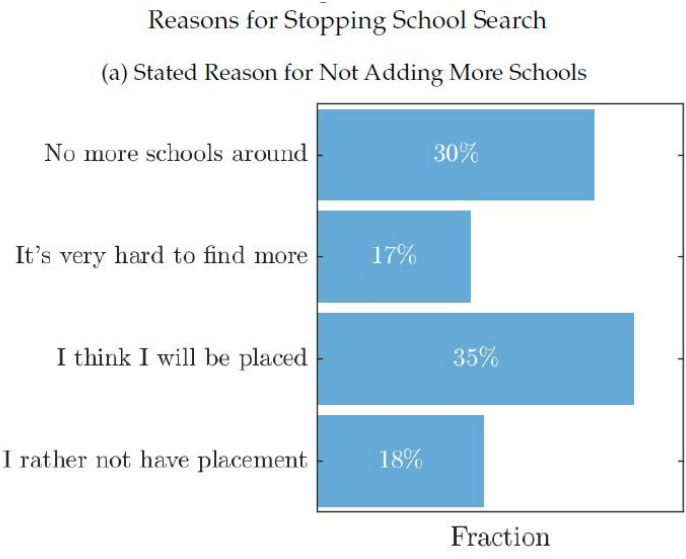
Family Networks

- Older siblings' experiences affected younger siblings' choices and assignment
- Effects do not appear to act through channels specific to siblings—convenience, commuting, or sibling rivalry
- What's the role of broader peer networks (older friends, classmates, etc)?
- How much do information-sharing and path dependence due to the perceived difficulty of navigating the school choice process affect choice?

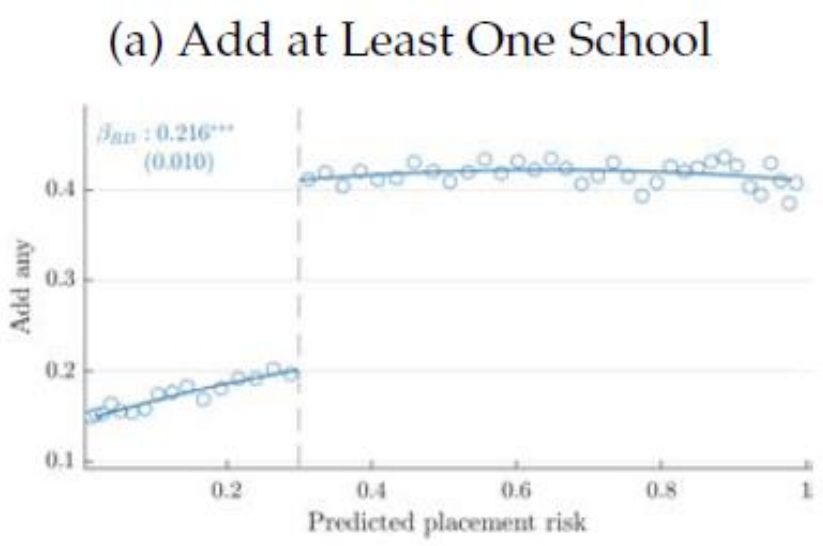
Arteaga, Kapor, Neilson, and Zimmerman (forthcoming)

- **Question**: How do search costs affect school choice? Can an information intervention expand search and improve admission outcomes?
- **Context**: centralised public high school choice system in Chile, smart matching platform with warnings (a pop-up in the application platform + off-platform text messages)
- **Approach**: RDD – looking at students at the threshold for receiving warning messages (>30% non-placement risk); Randomized assignment to receive text messages
- **Results**: Receiving a warning caused 21.6% of students to add schools to their applications, reduce their non-placement risk, and enroll in higher quality schools. Text messages without personalized risk information have no effects.

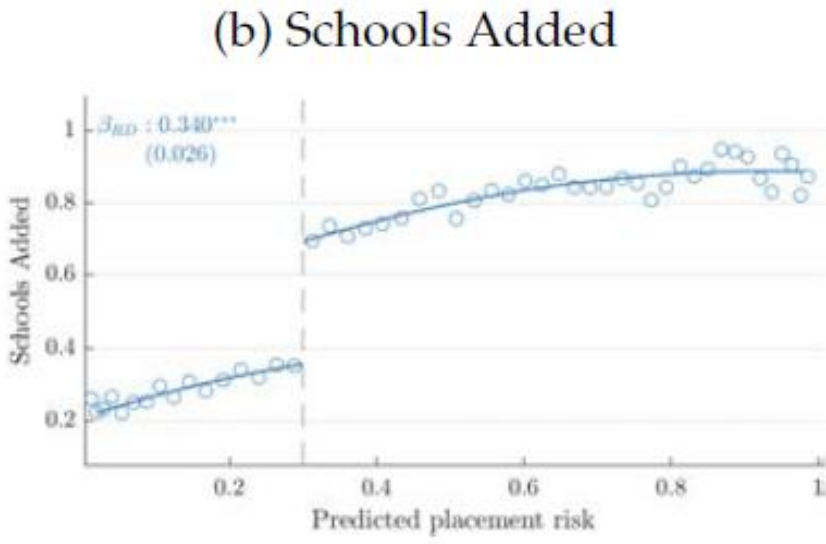
Survey evidence



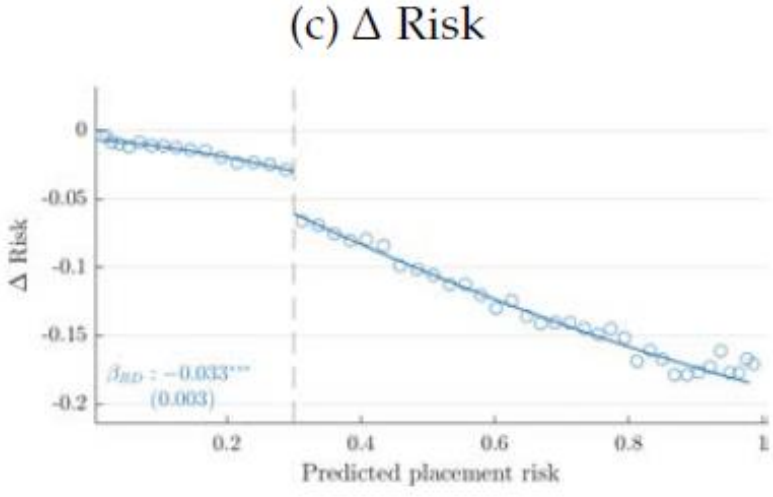
Effects on choices



Effects on choices

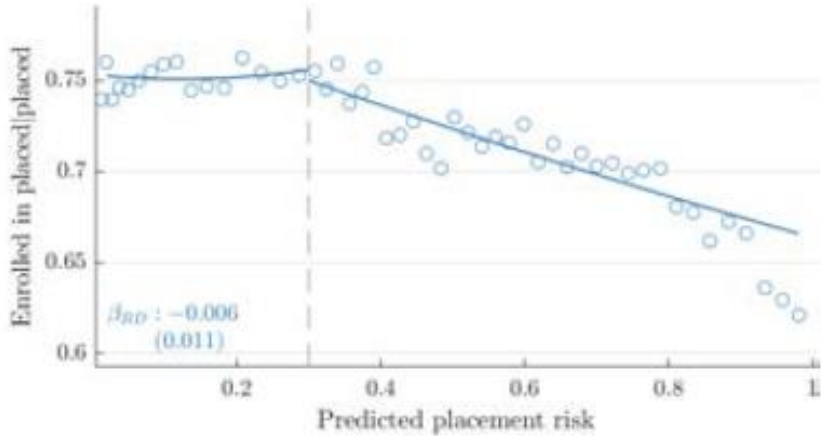


Effects on choices

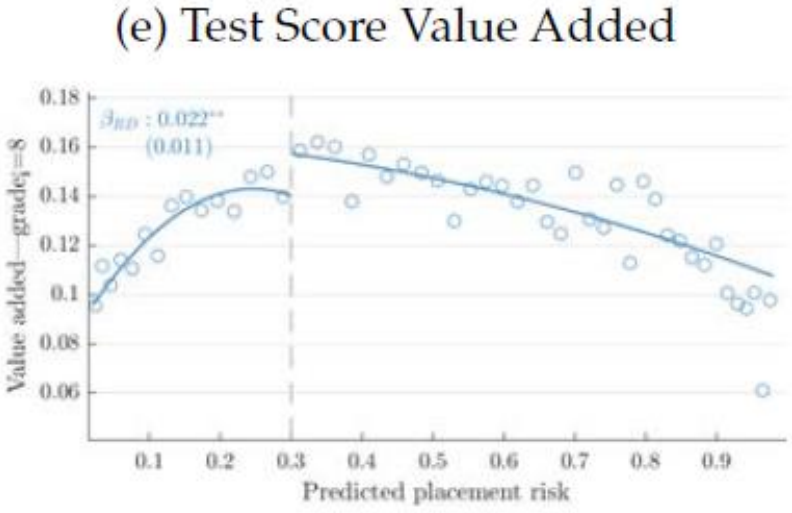


Effects on assignment

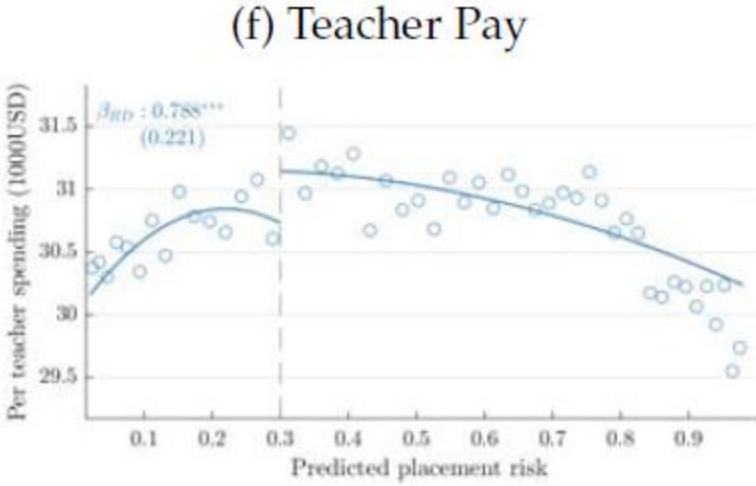
(d) Enrolled in Placed Conditional on Placed



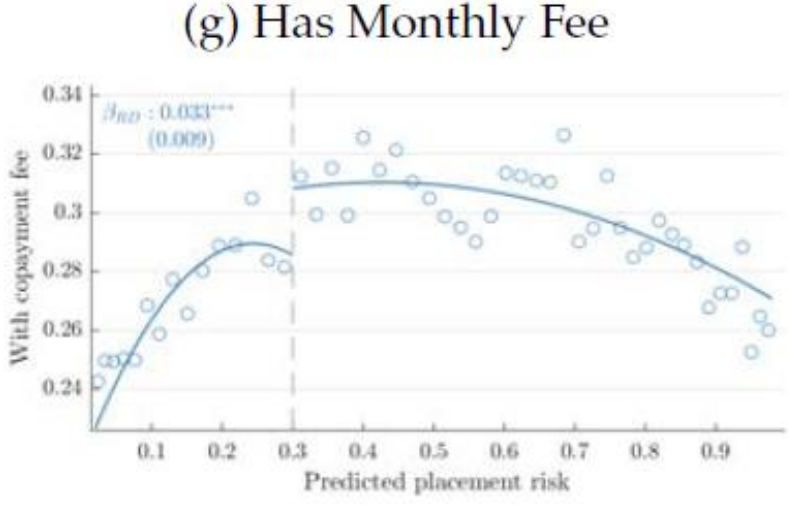
Effects on assignment



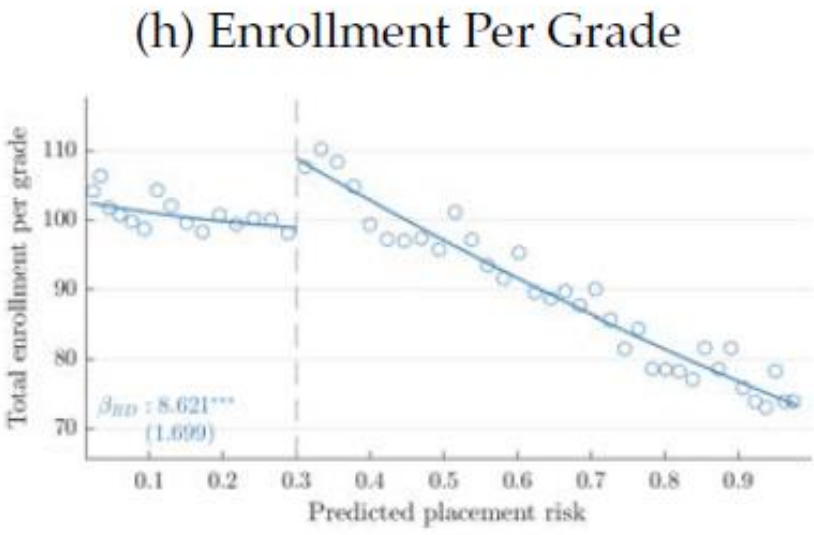
Effects on assignment



Effects on assignment



Effects on assignment



Search Costs

- Most students are overly optimistic about their admission chances
- Optimism about school placement chances makes students search too little
- Providing personalized support to reduce search costs for schools could be helpful
- Strategy-proof assignment mechanisms can still be challenging to navigate

Empirical analysis

1. Administrative data
2. Survey data
3. Exogenous variation
4. **Structural estimation**
5. Qualitative data

Structural estimation

- Advantages
  1. Counterfactual policy simulation
  2. Allows for heterogeneous treatment effects
- Limitations
  1. Relies on modeling assumptions

Ainsworth, Dehejia, Pop-Eleches and Urquiola (2020)

- **Question**: Do information and preferences determine whether households choose schools with high value added?
- **Context**: Centralised public high school choice system in Romania
- **Approach**: Randomized assignment to receive information on value added; Structural estimation to simulate effects of choice with fully accurate beliefs
- **Results**: Beliefs account for 18 (11) percent of the value added that households with low- (high-) achieving children leave unexploited.

Information on Value Added

1. Households could select schools with 1 s.d. worth of additional value added

2. Informing randomly selected households about the value added of the schools in their towns improves accuracy of households' beliefs and leads low-achieving students to attend higher-value added schools

3. For households with low-achieving children, information provision appears to affect both beliefs and preferences.

Empirical analysis

1. Administrative data
2. Survey data
3. Exogenous variation
4. Structural estimation
5. **Qualitative data**

Qualitative data

- Understand mechanisms behind observed effects
- Validate modeling assumptions

Final Thoughts

1. Centralised school assignment offers a unique policy instrument and learning opportunity
2. Blending multiple data sources and empirical strategies sheds valuable insights
3. Information is one of many factors determining school choice
4. Potentially heterogeneous impacts of information provision, often more beneficial for low-achieving students

References

- Ainsworth, Robert G., Rajeev Dehejia, Cristian Pop-Eleches and Miguel S. Urquiola (2020). "Information, Preferences, and Household Demand for School Value Added." NBER Working paper.
- Ajayi, Kehinde (forthcoming). "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana." Journal of Human Resources.
- Ajayi, Kehinde, Willa Friedman, and Adrienne Lucas (2020). "When Information is Not Enough: Evidence from a Centralized School Choice System." NBER Working Paper 27887.
- Arteaga, Felipe, Adam Kapor, Christopher Neilson, and Seth Zimmerman (forthcoming). "Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice." Quarterly Journal of Economics.
- Dustan, Andrew (2018). "Family Networks and School Choice", Journal of Development Economics, 134(2018): pp. 372-391.