

# Student Performance and the Effects of School Quality versus School Fit

Kehinde Ajayi

Department of Economics  
Boston University

UNU-WIDER Human Capital and Growth Conference  
June 7, 2016

# This Paper

## Question

- ▶ How do schools affect student performance?

## Potential channels

1. School quality → Access to better academic inputs
2. School fit → Access to a preferred type of school

# This Paper

## Question

- ▶ How do schools affect student performance?

## Potential channels

1. School quality → Access to better academic inputs
2. School fit → Access to a preferred type of school

	Available schools	Students' 1st choice
Private	0.063	0.001
Single sex	0.111	0.253
Colonial	0.102	0.250
Christian	0.188	0.238
Boarding	0.634	0.867
Small	0.249	0.056
Specialized	0.214	0.104
<i>N</i>	537	139073

# This Paper

## Empirical strategy

1. Track a cohort of Ghanaian high school students
2. Exploit merit-based school assignment process

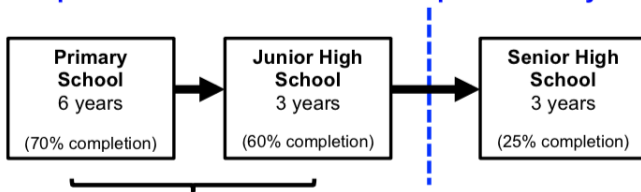
## Results

1. School quality
  - ▶ Increases likelihood of staying in same school
  - ▶ Generates modest improvements in exam performance
2. School fit
  - ▶ Increases likelihood of staying in same school
  - ▶ No significant effects on exam performance
3. Prioritizing school quality maximizes student learning

## Context and Data

# Context: Ghana

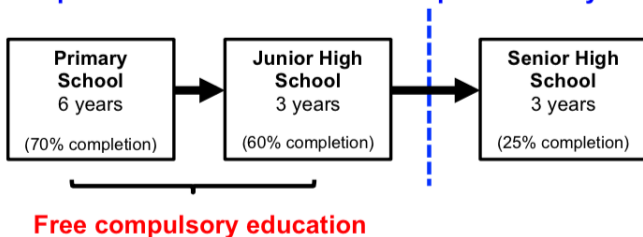
## Computerized school selection and placement system



**Free compulsory education**

## Context: Ghana

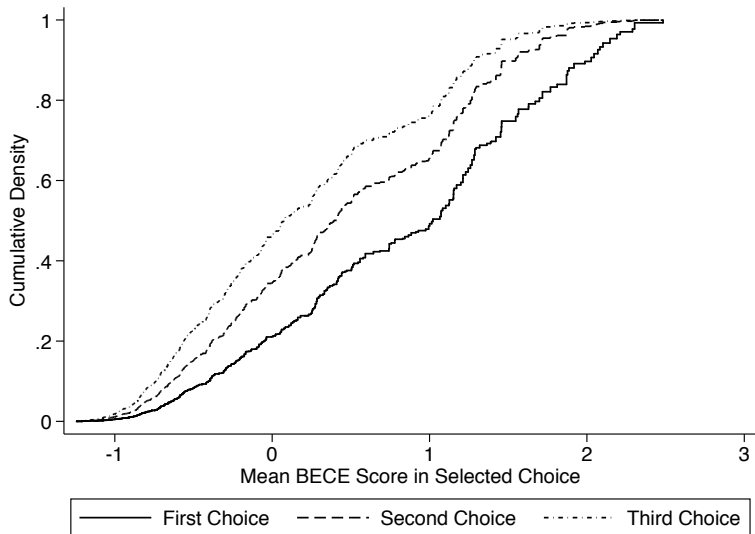
### Computerized school selection and placement system



### Coordinated school choice for students

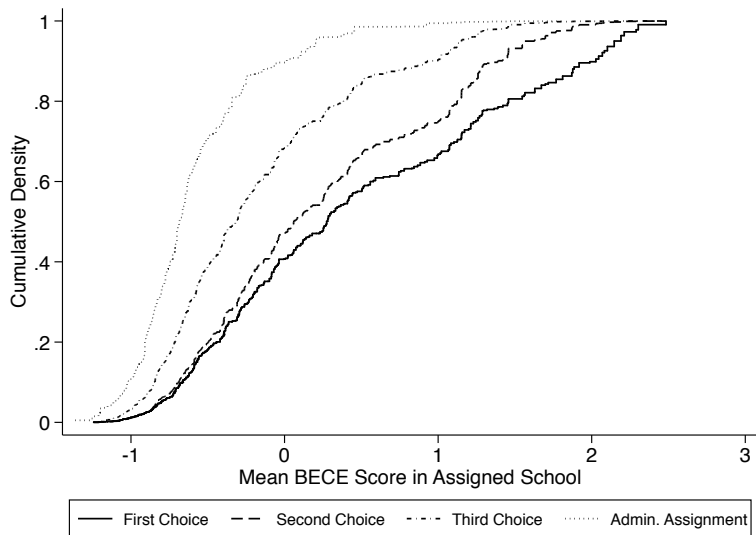
1. Submit ranked list of up to three Senior High School programs
2. Take Basic Education Certification Exam (BECE)
3. Admitted to one program based on BECE score and ranked list

# Peer Quality by Rank of Selected Choices





# Peer Quality by Rank of Assigned Choices



# Data: Administrative Records on a Cohort of Students

## **BECE candidates** (senior high school applicants) in 2005

- ▶ Background information: name, age, sex, junior high school
- ▶ Ranked list of choices
- ▶ BECE performance and admission outcomes

## **SSCE candidates** (senior high school graduates) in 2008 + 2009

- ▶ Background information: name, age, sex, senior high school
- ▶ SSCE performance

**Outcomes:** Link BECE to SSCE candidates using name, age, sex

# Research Design

# 1: Basic Model

$$Y_{is} = \alpha Q_s + f(BECE_i) + \gamma \mathbf{X}_i + \epsilon_{is}$$

- ▶  $Y_{is}$  - 12th grade outcomes (exam taking and scores)
- ▶  $Q_s$  - mean BECE score of assigned SHS peers
- ▶  $BECE_i$  - 9th grade exam score
- ▶  $X_i$  - age, gender, JHS mean BECE score, indicator for public JHS
- ▶ **Identification challenge:**  
Endogenous selection of students into schools

## 2: Selection on Observables

$$Y_{is} = \alpha Q_s + f(BECE_i) + \gamma \mathbf{X}_i + \delta Z_i + \nu_{is}$$

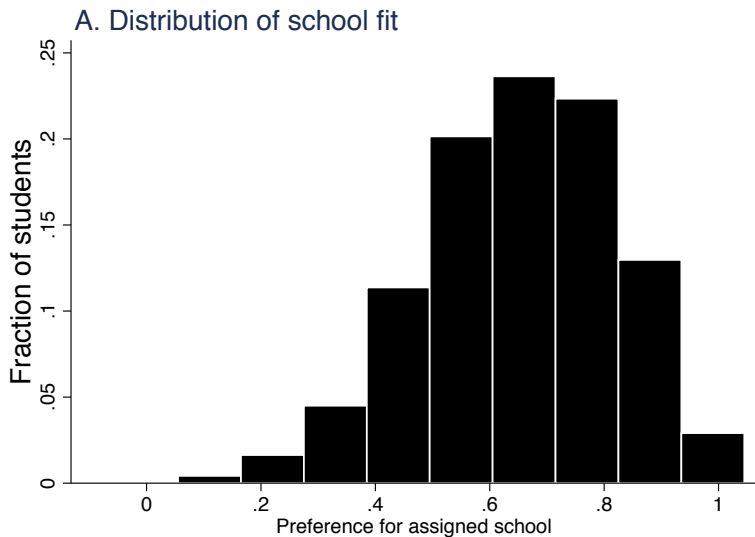
- ▶  $\delta Z_i$ : control for student application behavior
  - ▶ mean selectivity of chosen schools
  - ▶ fixed effect for ranked list of schools
  - ▶ fixed effect for ranked list of schools  $\times$  programs
- ▶ **Identification assumption:**  
Controlling for application behavior controls for unobserved student ability

### 3: School Fit

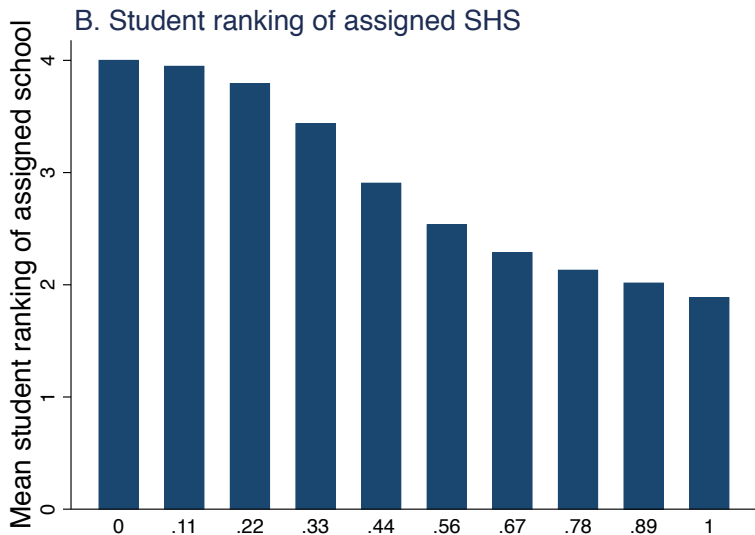
$$Y_{is} = \alpha Q_s + f(BECE_i) + \gamma \mathbf{X}_i + \delta Z_i + \phi_r + \theta Fit_s + v_{is}$$

- ▶  $\phi_r$ : fixed effect for student's ranking of assigned school
- ▶  $Fit_s$ : similarity between student's assigned school and selected choices
  1. private
  2. single-sex
  3. colonial (34 schools established by the British, pre-1957)
  4. Christian affiliated
  5. equipped with boarding facilities
  6. small (admits fewer than 185 students → lowest quartile)
  7. specialized (offers fewer than 4 programs → lowest quartile)
  8. in a student's JHS district
  9. in a student's JHS region

# School Fit Distribution

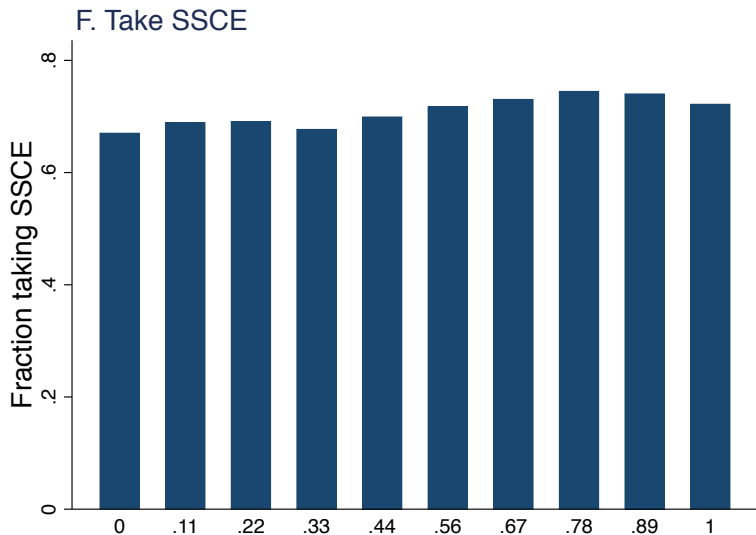


# School Fit Distribution

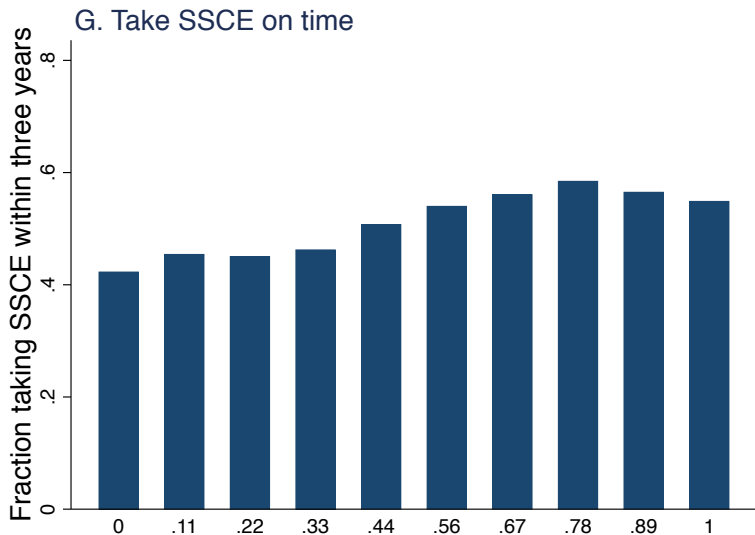




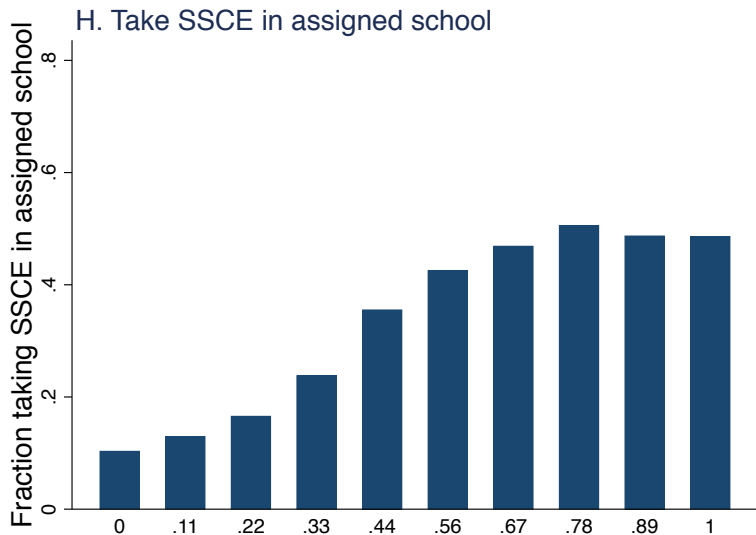
# School Fit Outcomes



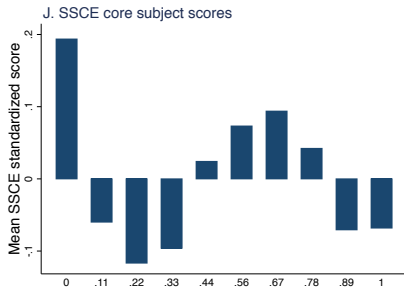
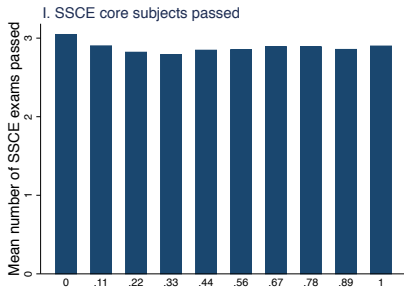
# School Fit Outcomes



# School Fit Outcomes



# School Fit Outcomes



# Summary of Descriptive Results

1. Large variation in school fit
2. Correlated with student retention but not exam performance

## Regression Results

# School Quality versus School Fit

Matched program lists	w/ group-specific score slope	w/ cubic score function
	(1)	
<i>Panel A. Take SSCE</i>		
Assigned SHS peers	0.059 (0.024)**	
School fit		
<i>Panel B. Take SSCE on time</i>		
Assigned SHS peers	0.081 (0.029)***	
School fit		
<i>Panel C. Take SSCE in assigned school</i>		
Assigned SHS peers	0.211 (0.030)***	
School fit		
Rank of Assigned SHS Observations	No 32153	

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# School Quality versus School Fit

Matched program lists	w/ group-specific score slope		w/ cubic score function
	(1)	(2)	
<i>Panel A. Take SSCE</i>			
Assigned SHS peers	0.059 (0.024)**	0.053 (0.032)	
School fit			
<i>Panel B. Take SSCE on time</i>			
Assigned SHS peers	0.081 (0.029)***	0.055 (0.038)	
School fit			
<i>Panel C. Take SSCE in assigned school</i>			
Assigned SHS peers	0.211 (0.030)***	0.093 (0.044)**	
School fit			
Rank of Assigned SHS Observations	No 32153	Yes 32153	

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .



# School Quality versus School Fit

Matched program lists	w/ group-specific score slope			w/ cubic score function
	(1)	(2)	(3)	
<i>Panel A. Take SSCE</i>				
Assigned SHS peers	0.059 (0.024)**	0.053 (0.032)	0.050 (0.032)	
School fit			0.060 (0.088)	
<i>Panel B. Take SSCE on time</i>				
Assigned SHS peers	0.081 (0.029)***	0.055 (0.038)	0.052 (0.038)	
School fit			0.080 (0.108)	
<i>Panel C. Take SSCE in assigned school</i>				
Assigned SHS peers	0.211 (0.030)***	0.093 (0.044)**	0.082 (0.042)*	
School fit			0.306 (0.113)***	
Rank of Assigned SHS Observations	No 32153	Yes 32153	Yes 32153	

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# School Quality versus School Fit

Matched program lists	w/ group-specific score slope				w/ cubic score function
	(1)	(2)	(3)	(4)	
<i>Panel A. Take SSCE</i>					
Assigned SHS peers	0.059 (0.024)**	0.053 (0.032)	0.050 (0.032)	0.051 (0.032)	
School fit			0.060 (0.088)	0.075 (0.111)	
<i>Panel B. Take SSCE on time</i>					
Assigned SHS peers	0.081 (0.029)***	0.055 (0.038)	0.052 (0.038)	0.053 (0.038)	
School fit			0.080 (0.108)	0.149 (0.145)	
<i>Panel C. Take SSCE in assigned school</i>					
Assigned SHS peers	0.211 (0.030)***	0.093 (0.044)**	0.082 (0.042)*	0.087 (0.041)**	
School fit			0.306 (0.113)***	0.457 (0.145)***	
Rank of Assigned SHS Observations	No 32153	Yes 32153	Yes 32153	Yes 32153	

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# School Quality versus School Fit

Matched program lists	w/ group-specific score slope				w/ cubic score function	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Take SSCE</i>						
Assigned SHS peers	0.059 (0.024)**	0.053 (0.032)	0.050 (0.032)	0.051 (0.032)	0.054 (0.009)***	0.032 (0.012)***
School fit			0.060 (0.088)	0.075 (0.111)		0.059 (0.047)
<i>Panel B. Take SSCE on time</i>						
Assigned SHS peers	0.081 (0.029)***	0.055 (0.038)	0.052 (0.038)	0.053 (0.038)	0.087 (0.010)***	0.052 (0.013)***
School fit			0.080 (0.108)	0.149 (0.145)		0.143 (0.060)**
<i>Panel C. Take SSCE in assigned school</i>						
Assigned SHS peers	0.211 (0.030)***	0.093 (0.044)**	0.082 (0.042)*	0.087 (0.041)**	0.200 (0.012)***	0.092 (0.014)***
School fit			0.306 (0.113)***	0.457 (0.145)***		0.462 (0.063)***
Rank of Assigned SHS Observations	No 32153	Yes 32153	Yes 32153	Yes 32153	No 32153	Yes 32153

Notes: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

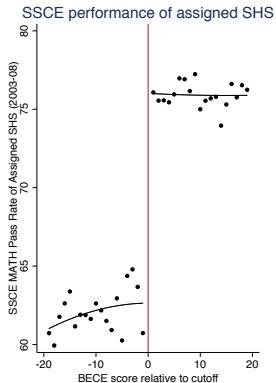
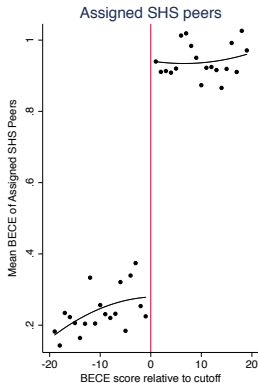
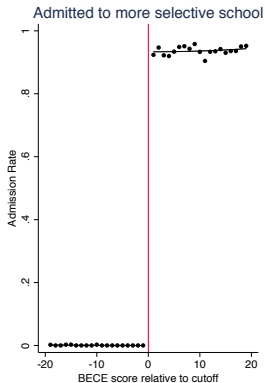
# School Fit

Matched program lists	w/ group-specific score slope				w/ cubic score function	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel D. SSCE passes</i>						
Assigned SHS peers	0.013 (0.064)	0.076 (0.102)	0.076 (0.103)	0.076 (0.102)	0.034 (0.025)	0.084 (0.034)**
School fit			0.003 (0.267)	-0.022 (0.348)		-0.145 (0.141)
Rank of Assigned SHS	No	Yes	Yes	Yes	No	Yes
$R^2$	0.799	0.799	0.799	0.799	0.492	0.492
Observations	20941	20941	20941	20941	20941	20941
<i>Panel E. SSCE score</i>						
Assigned SHS peers	-0.003 (0.065)	-0.017 (0.097)	-0.013 (0.097)	-0.010 (0.094)	0.055 (0.024)**	0.128 (0.030)***
School fit			-0.089 (0.235)	-0.448 (0.287)		-0.405 (0.121)***
Rank of Assigned SHS	No	Yes	Yes	Yes	No	Yes
$R^2$	0.884	0.884	0.884	0.885	0.709	0.710
Observations	20941	20941	20941	20941	20941	20941

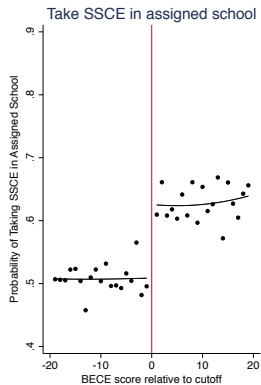
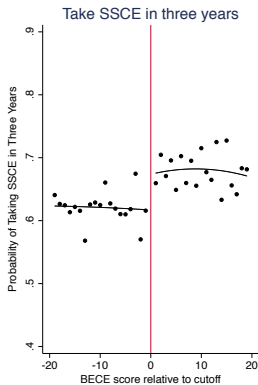
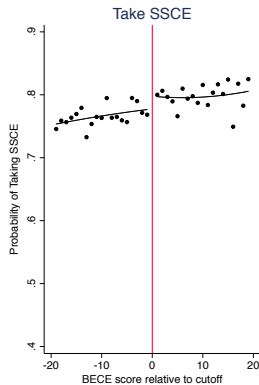
Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# Alternative Identification Strategy

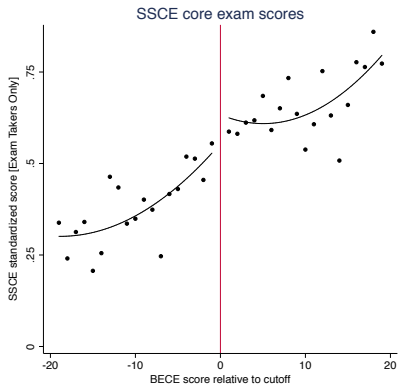
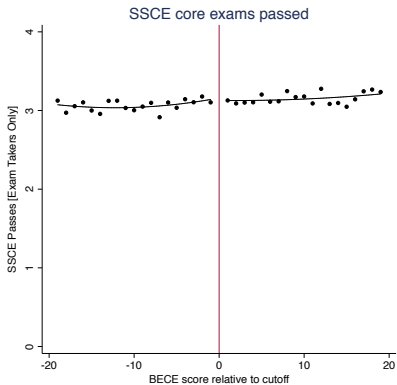
# Regression Discontinuity Design: First Stage



# Effects: School Retention

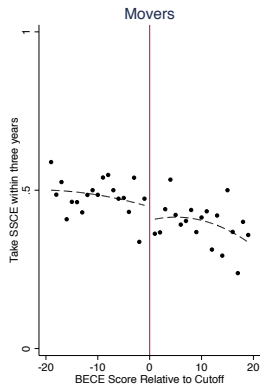
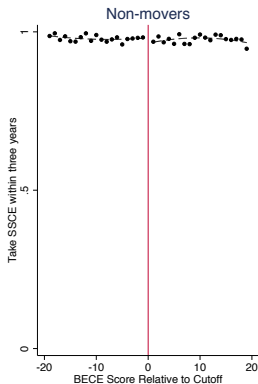
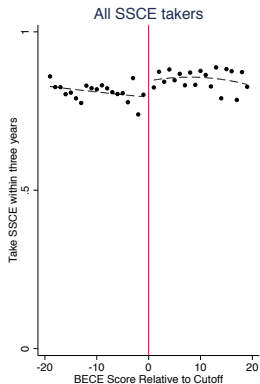


# Effects: SSCE Performance

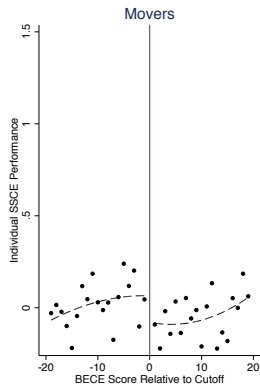
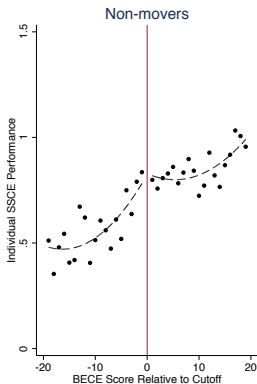
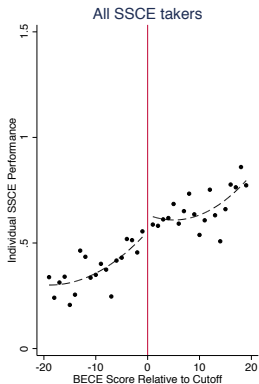




# Implications of Moving: School Completion



# Implications of Moving: SSCE Performance



# Summary of Additional Results

- ▶ Similar results using alternative approaches
- ▶ Schools have much larger effects on student retention than on exam performance
- ▶ Negative correlation between switching schools and exam performance

# Conclusions

- ▶ Prioritizing school quality over school fit maximizes gains in student performance
- ▶ Implications:
  1. Information emphasizing school quality over other attributes might better enable families to reap the academic benefits of school choice programs
  2. Public education investments more likely to raise student performance by expanding access to high quality schools instead of expanding range of attributes of available schools
- ▶ Limitations: cannot observe nonacademic outcomes

# Appendix

# Context: Merit-based assignment

## Deferred acceptance algorithm

1. All students apply to first choice
  - ▶ Schools conditionally accept highest scorers, reject others
2. Rejected students apply to second choice
  - ▶ Schools consider all applicants (conditionally accepted and new) and conditionally accept highest scorers, reject others
3. Rejected students apply to next choice on their list
4. Algorithm stops when students have exhausted their choices
5. Unsuccessful applicants assigned to undersubscribed programs

# Student Characteristics

	Analysis sample		Matched applicants same school list		Matched applicants same program list	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
<i>Student Characteristics</i>						
Age	16.633	1.727	16.577	1.714	16.533	1.727
Male	0.592		0.585		0.598	
JHS Public	0.751		0.743		0.730	
Standardized BECE score	0.011	1.005	0.083	1.025	0.195	1.055
Mean BECE of JHS peers	0.004	0.780	0.053	0.797	0.116	0.818
Mean BECE of assigned SHS peers	0.016	0.857	0.089	0.885	0.155	0.921
<i>Admission Outcomes</i>						
First choice program	0.306		0.297		0.299	
Second choice program	0.206		0.207		0.210	
Third choice program	0.261		0.266		0.275	
Administrative assignment	0.227		0.230		0.217	
<i>Secondary School Performance</i>						
Take SSCE	0.725		0.738		0.750	
Take SSCE in three years	0.549		0.568		0.589	
Take SSCE in assigned school	0.441		0.464		0.488	
SSCE core subjects passed	2.080	1.604	2.138	1.597	2.205	1.592
Standardized SSCE score	0.012	1.016	0.059	1.031	0.134	1.066
<i>Selectivity Range within Matched Sets</i>						
Range in schools applied to			0.999	0.502	0.961	0.464
Range in schools admitted to			1.297	0.911	0.786	0.793
<i>N</i>	139073		94918		32153	

# School Characteristics

	Available schools		All students		SSCE takers		
	Un-weighted (1)	Weight by vacancies (2)	1st choice school (3)	Assigned school (4)	1st choice school (5)	Assigned school (6)	SSCE school (7)
Private	0.188	0.063	0.001	0.065	0.001	0.056	0.078
Single sex	0.090	0.111	0.253	0.111	0.282	0.130	0.138
Colonial	0.052	0.102	0.250	0.104	0.273	0.122	0.127
Christian	0.208	0.188	0.238	0.191	0.251	0.201	0.210
Boarding	0.525	0.634	0.867	0.640	0.882	0.667	0.673
Small	0.532	0.249	0.056	0.238	0.049	0.208	0.224
Specialized	0.457	0.214	0.104	0.212	0.102	0.195	0.209
JHS region			0.773	0.790	0.764	0.781	0.628
JHS district			0.448	0.419	0.441	0.421	0.342
<i>N</i>	648	537	139073	139073	100240	100240	100240



### 3: School Fit

$$Y_{is} = \alpha Q_s + f(BECE_i) + \gamma \mathbf{X}_i + \delta Z_i + \phi_r + \theta Fit_s + v_{is}$$

- ▶  $\phi_r$ : fixed effect for student's ranking of assigned school
- ▶  $Fit_s$ : similarity between student's assigned school and:
  1. First choice school
  2. Application portfolio

### 3: School Fit

A. Similarity to first choice:

$$Fit_s = \frac{1}{C} \sum_c^C \mathbb{1}(c_s = c_1)$$

- ▶ Does student's assigned school  $s$  have the same characteristic  $c$  as her first choice school ( $c_s = c_1$ )?
- ▶  $Fit_s$  ranges from 0 to 1

### 3: School Fit

B. Similarity to portfolio:

$$Fit'_s = \frac{1}{C} \sum_{c=1}^C [(\mathbb{1}(c_s = 1) \times \mathbb{1}(c_p = 3)) + (\mathbb{1}(c_s = 0) \times \mathbb{1}(c_p = 0))]$$

- ▶ Was student assigned to a given type of school ( $c_s = 1$ ) she listed for all three choices in her portfolio  $p$ ? OR
- ▶ Was student not assigned a type of school ( $c_s = 0$ ), she didn't list for any of her choices ( $c_p = 0$ )?
- ▶  $Fit'_s$  ranges from 0 to 1

# Regression Results

$$1. Y_{is} = \alpha Q_s + \beta BECE_i + \gamma \mathbf{X}_i + \epsilon_{is}$$

$$2. Y_{is} = \alpha Q_s + \beta BECE_i + \gamma \mathbf{X}_i + \delta Z_i + \nu_{is}$$

$$Y_{is} = \alpha Q_s + \beta_1 BECE_i + \beta_2 BECE_i^2 + \beta_3 BECE_i^3 + \gamma \mathbf{X}_i + \delta Z_i + \mu_{is}$$

$$Y_{is} = \alpha Q_s + \beta BECE_i + \gamma \mathbf{X}_i + \delta Z_i + \pi (BECE_i \times Z_i) + \eta_{is}$$

$$3. Y_{is} = \alpha Q_s + f(BECE_i) + \gamma \mathbf{X}_i + \delta Z_i + \phi_r + \theta Fit_s + v_{is}$$

# Selection on Observables

	Basic model (1)	Control for selectivity of choices (2)	Matched school lists (3)	Matched program lists (4)	Prog. list w/ cubic score (5)	Prog. list w/ group slope (6)
<i>Panel A. Take SSCE</i>						
Assigned SHS peers	0.021 (0.004)***	0.019 (0.004)***	0.027 (0.005)***	0.042 (0.009)***	0.054 (0.009)***	0.059 (0.024)**
$R^2$	0.057	0.057	0.260	0.432	0.433	0.765
$N$	139073	139073	94918	32153	32153	32153
Mean Dep. Variable	0.725	0.725	0.738	0.750	0.750	0.750
<i>Panel B. Take SSCE on Time</i>						
Assigned SHS peers	0.042 (0.006)***	0.045 (0.006)***	0.057 (0.006)***	0.073 (0.010)***	0.087 (0.010)***	0.081 (0.029)***
$R^2$	0.087	0.087	0.295	0.463	0.464	0.773
$N$	139073	139073	94918	32153	32153	32153
Mean Dep. Variable	0.549	0.549	0.568	0.589	0.589	0.589
<i>Panel C. Take SSCE in Assigned School</i>						
Assigned SHS peers	0.121 (0.009)***	0.140 (0.009)***	0.165 (0.008)***	0.185 (0.012)***	0.200 (0.012)***	0.211 (0.030)***
$R^2$	0.121	0.130	0.343	0.502	0.503	0.785
$N$	139073	139073	94918	32153	32153	32153
Mean Dep. Variable	0.441	0.441	0.464	0.488	0.488	0.488

Notes: Robust standard errors clustered at the level of assigned SHS reported in parentheses, \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# Selection on Observables

	Basic model (1)	Control for selectivity of choices (2)	Matched school lists (3)	Matched program lists (4)	Prog. list w/ cubic score (5)	Prog. list w/ group slope (6)
<i>Panel D. SSCE Passes</i>						
Assigned SHS peers	0.002 (0.013)	-0.008 (0.013)	-0.009 (0.014)	-0.006 (0.025)	0.034 (0.025)	0.013 (0.064)
$R^2$	0.126	0.127	0.336	0.490	0.492	0.799
$N$	100842	100842	66259	20941	20941	20941
Mean Dep. Variable	2.869	2.869	2.905	2.962	2.962	2.962
<i>Panel E. SSCE Score</i>						
Assigned SHS peers	0.169 (0.020)***	0.154 (0.020)***	0.102 (0.017)***	0.138 (0.027)***	0.055 (0.024)**	-0.003 (0.065)
$R^2$	0.405	0.406	0.575	0.697	0.709	0.884
$N$	100842	100842	66259	20941	20941	20941
Mean Dep. Variable	0.032	0.032	0.089	0.197	0.197	0.197

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

# Validity of Selection on Observables

- ▶ Is school assignment independent of observable student characteristics, after controlling for application behavior  $\delta Z_i$ ?

	Dependent variable: Mean BECE Score of Assigned SHS Peers					
	Basic model	Control for selectivity of choices	Matched school lists	Matched program lists	Prog. list w/ cubic score	Prog. list w/ group slope
	(1)	(2)	(3)	(4)	(5)	(6)
Male	-0.014 (0.015)	-0.005 (0.015)	-0.032 (0.005)***	-0.012 (0.009)	-0.010 (0.009)	0.008 (0.014)
Age	-0.014 (0.002)***	-0.006 (0.002)***	-0.001 (0.001)	0.001 (0.002)	0.001 (0.002)	0.001 (0.003)
JHS public	-0.018 (0.008)**	0.008 (0.007)	0.005 (0.006)	-0.003 (0.010)	-0.002 (0.009)	-0.001 (0.017)
JHS peers	0.146 (0.009)***	0.113 (0.009)***	0.097 (0.008)***	0.081 (0.011)***	0.040 (0.008)***	0.020 (0.015)
$R^2$	0.741	0.749	0.835	0.903	0.911	0.973
$N$	139073	139073	94918	32153	32153	32153

Notes: Regressions also include controls for individual BECE score. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

- ▶ Does student ranking of assigned school affect outcomes?

	School Retention			Exam Performance	
	Take SSCE (1)	Take SSCE on Time (2)	Take SSCE in Assigned School (3)	SSCE passes (4)	SSCE score (5)
Assigned SHS peers	0.053 (0.032)	0.055 (0.038)	0.093 (0.044)**	0.076 (0.102)	-0.017 (0.097)
Admitted to first choice	0.005 (0.063)	0.046 (0.074)	0.257 (0.090)***	-0.198 (0.207)	0.083 (0.167)
Admitted to second choice	0.028 (0.049)	0.073 (0.058)	0.252 (0.069)***	-0.028 (0.148)	-0.069 (0.124)
Admitted to third choice	0.033 (0.042)	0.082 (0.049)*	0.250 (0.060)***	0.002 (0.118)	-0.064 (0.105)
$R^2$	0.765	0.773	0.787	0.799	0.884
Observations	32153	32153	32153	20941	20941
Mean dep. variable	0.750	0.589	0.488	2.962	0.197

Notes: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .



# Regression Discontinuity Design

Compare students on opposite sides of admission cutoffs

$$Q_s = \gamma 1\{BECE_i \geq \underline{BECE}_p\} + a(BECE_i) + \lambda_p + \eta_i$$

$$Y_{is} = \delta E(Q_s | BECE_i) + a(BECE_i) + \lambda_p + \mu_i$$

- ▶  $1\{BECE_i \geq \underline{BECE}_p\}$ : indicator for scoring above admission cutoff
- ▶  $a(BECE_i)$ : control function for BECE scores
- ▶  $\lambda_p$ : cutoff fixed effects
- ▶ Restrict sample to narrow bandwidth of binding admission cutoffs. Normalize cutoff scores to 0 and pool data across cutoffs ( $p = 257$ )

**Identification Assumption:** Students with similar scores would have had the same outcomes if there were no admission cutoffs

# Robustness Checks

- ▶ Alternative functional forms for BECE scores
- ▶ Varying RD bandwidth
- ▶ Bounding estimates to correct for missing data