Selection, Gender and the Impact of Schooling Type in the Dhaka Slums

John C. Ham [1] Saima Khan [2]

[1]NYU Abu Dhabi & NYU Wagner School of Public Service

[2]NYU Abu Dhabi

Weatherall Lecture, Queen's University

September 17, 2019

Introduction

- Numerous schools providing primary education in Dhaka slums - Govt schools, NGO schools, private schools and *madrassahs*.
  - All provide traditional Bengali medium education of varying qualities.
  - Govt and NGO schools dominate schooling provision in slums.
  - NGO schools treated as substitutes of government schools.
  - No evidence how Govt and NGO schools compare in terms of learning outcomes in urban Bangladesh.

Introduction (contd)

- In 2007, a new school-type called JAAGO started operating in two slums of Dhaka.
  - JAAGO is unique in terms of providing English medium education, strict monitoring, no corporal punishment etc;
  - This type of schooling previously available only to the elites of the country.
- There is no existing data we could use to evaluate JAAGO, so we collected our own data.

Research Questions

1. What type of students are being drawn to, and accepted by, each school-type?
   - Is there selection across school-types for boys and girls?

2. What is the impact of school-type on test scores by gender, before and after controlling for selection?
   (i). JAAGO vs. Govt;
   (ii). JAAGO vs. NGO;
   (iii). Govt vs. NGO.

Presentation Outline

Motivation and Overview of Results

In the context of Bangladesh:

1. No evidence about impact of school-type on student achievement and gender differential for urban Bangladesh.

2. High enrolment at primary level, but poor learning outcomes [World Bank, 2013].

3. Gender parity in primary school enrolment, but not in achievement [World Bank (2013), ADB Country Gender Assessment Bangladesh (2010)].

In a wider context:

4. Similar low student achievement and wide gender gap in Pakistan [Das, Pandey and Zajonc (2012), The Economist (Jan 4, 2018)].

Motivation and Overview of Results (contd)

- Given the poor learning outcomes and gender gap in both countries (combined population of 380 million), it is important to:
  - compare the two dominant school-types - Govt and NGO schools;
  - consider an alternative schooling model, JAAGO.

Motivation and Overview of Results (contd)

- Our results indicate strong evidence of gender heterogeneity across school-types.
    - Govt. vs. NGO: Boys are better off at Govt. schools, but girls perform equally at both school-types.
    - JAAGO vs. Govt: Girls are better off at JAAGO, but boys perform equally at both school-types.
    - JAAGO vs. NGO: Both genders better off at JAAGO.

Motivation and Overview of Results (contd)

- Our work can explain the gender achievement gap in Bangladesh.
- In work we won't be able to discuss today, we also find within school type gender differences.
- After controlling for the X's:
  - Boys and girls do equally well (badly) in NGO schools;
  - But boys outperform girls at Govt schools.
- Since the vast majority of students go to one of these two types of schools, the boys' aggregate achievement has to be higher.

Motivation and Overview of Results (contd)

- But introducing JAAGO should help to equalize gender outcomes.
- At JAAGO, girls do better than if they attended government or NGO schools, and girls do equally well as boys at JAAGO.
- Thus, JAAGO may help attain gender parity in terms of achievement, and reduce or eliminate aggregate gender differences in achievement.

The New Kid on the Block: JAAGO

- JAAGO Foundation is a Civil Society Organization (CSO) that started operations in 2007 with one physical school in the Rayerbazar slum of Dhaka city.
- As of 2019, JAAGO foundation has 3 projects: (i) the Education Program; (ii) the Youth Development Program; and (iii) the Rohingya Refugee Project.
- We focus on the education program in Bangladesh which consists of 3 offline (physical) schools and 9 online schools.

The New Kid on the Block: JAAGO (contd)

- Of the 3 off-line schools, 2 are located in Dhaka, while one is located in Chittagong, a southern city of Bangladesh (distinct from Dhaka in terms of distance, economic structure, income scale etc).

- JAAGO also has 9 online schools located in different parts of Bangladesh (outside Dhaka).
  - Each of these location consist of a brick and mortar structure where students come for their regular classes and learn their lessons from the ‘teacher in the TV’.
  - We do not consider these online schools since they are quite different from the physical schools.

Characteristics Across the 3 School-types

Characteristics	JAAGO School	Govt. School	NGO School
Instruction in English	✓	×	×
Minimum teacher qualification - Bachelors Degree	✓	×	×
Teachers require strong command over English	✓	×	×
High level in-service training	×	✓	×
High share of female teachers	✓	×	✓
High teacher absenteeism (low teacher effort)	×	✓	NA
High headmaster absenteeism (low monitoring)	×	✓	NA
High teacher salary	×	✓	×
Small class size	✓	×	✓
Longer school days	✓	×	×
Longer school year	✓	×	×
Corporal punishment	×	✓	×

Data Collection

- Between 2015-2016, collected our own stratified (by school type) data on 1936 slum children (aged 4 - 14) attending the 3 types of schools.
  - JAAGO schools - 607 children;
  - Government schools - 618 children;
  - NGO schools - 711 children.
- Took many steps, including 100% audio auditing, to insure data quality.

Data Collection (contd)

- Used choice based sampling to ensure sufficient number of JAAGO students show up in the sample (common for sampling of rare events).
  - We collect the data by streets. We start with a street with a JAAGO student, then collect other students on the same street. We have 26 clusters in our sample.
  - We adjust the standard errors for this cluster sampling following Abadie, Athey, Imbens and Wooldridge (2017).

Distributions of Schools by School Type in Our Sample

Table 1: **Summary Statistics of Schools by School Type**

	(1) No. of schools	(2) Mean (no. of students)	(3) Std. Dev	(4) Total (no. of students)
Govt	13	47.54	86.39	618
JAAGO	2	303.50	84.15	607
NGO	29	24.52	46.99	711

(a) Note that due to unavailability of administrative data, we are unable to present distribution of schools per school-type in the greater population.

School Type and Selection: Sorting?

- Investigate selection across school-types in terms of 5 key variables :
  1. Monthly Family Expenditure (deflated by equivalence scale);
  2. Father's Schooling;
  3. Mother's Schooling;
  4. K-BIT (IQ/Fluid Intelligence);
  5. Raven's Coloured Progressive Matrices (IQ/Fluid Intelligence).

School Type and Selection: Sorting? (contd)

- Note that fluid intelligence, which is presumably measured by IQ tests, is defined as intelligence that is not supposed to be affected by attending school unless the schools 'teach to the test';
  - K-BIT and Raven's CPM - 2 different IQ tests that have some overlap (but not full overlap).

Selection: Means Across School-types for Boys

Table 2: **Means Across School Types (Boys)**

	JAAGO	Govt	NGO
Monthly Family Expdt (in BDT adjusted by equivalence scale)	5842.28	6262.32	5440.13
Father's schooling	4.0212	3.8987	3.0961
Mother's schooling	3.7327	3.2368	2.6275
K-BIT (IQ)	0.3596	0.0412	-0.3100
Raven's CPM (IQ)	0.2635	0.2031	-0.2510

Notes:
(a) For the Achievement Test scores and both the IQ scores, we use their respective age adjusted Z-scores. In other words for student i in age group a, we calculate, $Z_i = \frac{X_i - X_a}{\sigma_a}$, where X_a and σ_a is the mean and standard deviation in age group a.

Selection: Means Across School-types for Girls

Table 3: **Means Across School Types (Girls)**

	JAAGO	Govt	NGO
Monthly Family Expdt (in BDT adjusted by equivalence scale)	5846.97	6057.47	5315.87
Father's schooling	3.3787	3.5254	2.8093
Mother's schooling	3.8481	3.2624	2.5185
K-BIT (IQ)	0.1858	0.0875	-0.2913
Raven's CPM (IQ)	0.0443	0.0302	-0.2285

Notes:
(a) For the Achievement Test scores and both the IQ scores, we use their respective age adjusted Z-scores. In other words for student i in age group a, we calculate, $Z_i = \frac{X_i - X_a}{\sigma_a}$, where X_a and σ_a is the mean and standard deviation in age group a.

Selection: Mean Differences Across School-types for Boys

Table 4: **Mean Differences Across School Types (Boys)**

	J vs. G	J vs. N	G vs. N
Monthly Family Expdt	-420.04***	402.15 ***	822.19 ***
(in BDT adjusted by equivalence scale)	[155.4693]	[150.6422]	[139.4578]
Father's schooling	0.1225	0.9251***	0.8026***
	[0.3200]	[0.3086]	[0.3015]
Mother's schooling	0.4959*	1.1052***	0.6093**
	[0.2800]	[0.2711]	[0.2531]
K-BIT (IQ)	0.3184***	0.6696***	0.3512***
	[0.0802]	[0.0886]	[0.0819]
Raven's CPM (IQ)	0.0604	0.5145***	0.4541***
	[0.0928]	[0.0902]	[0.0834]

Notes:
(a) Standard errors in parentheses;
(b) We report the difference in means at the 10 percent, 5 percent and 1 percent significance level denoted by * , ** , and *** respectively;
(c) For the Achievement Test scores and both the IQ scores, we report their respective age adjusted Z-scores.

Selection: Mean Differences Across School-types for Girls

Table 5: **Mean Differences Across School Types (Girls)**

	J vs. G	J vs. N	G vs. N
Monthly Family Expdt	-210.50	531.11***	741.61***
(in BDT adjusted by equivalence scale)	[140.1312]	[125.3314]	[123.8186]
Father's schooling	-0.1467	0.5694**	0.7162***
	[0.2853]	[0.2662]	[0.2668]
Mother's schooling	0.5857**	1.3296***	0.7439***
	[0.2557]	[0.2301]	[0.2329]
K-BIT (IQ)	0.0983	0.4771***	0.3788***
	[0.0809]	[0.0731]	[0.0733]
Raven's CPM (IQ)	0.0141	0.2728***	0.2587***
	[0.0794]	[0.0693]	[0.0708]

Notes:
(a) Standard errors in parentheses;
(b) We report the difference in means at the 10 percent, 5 percent and 1 percent significance level denoted by * , ** , & *** respectively ;
(c) For the Achievement Test scores and both the IQ scores, we report their respective age adjusted Z-scores.

Evidence of Selection Across School-Types

- J vs. G
  - Boys at Govt schools belong to wealthier families.
  - Both genders at JAAGO schools have better educated mothers.
  - Boys with higher K-BIT score (fluid intelligence) go to JAAGO schools.
  - Raven's test score (fluid intelligence) fails to pick up any significant difference between JAAGO and Govt students for both genders.
- J vs. N and G vs. N (both genders)
  - NGO boys and girls have significantly lower fluid intelligence, have less educated parents and belong to poorer families.

Outcome Variable

- Our outcome variable: Woodcock Johnson Tests of Math Achievement
  - Widely used in the the Economics, Education and Psychology Literature.
  - Internationally developed and standardized.

Outcome Variable (contd)

- Used 3 Math oral subtests.
  - Govt and NGO students taught in Bengali while JAAGO students taught in English;
  - Used Mathematics subtests since it is not as dependent on language skills;
  - However, administered the tests in Bengali to Govt and NGO students; administered the same tests to JAAGO students in "Banglish" (i.e. kept technical terms in English).

Methodology: Dealing with Endogeneity

- School-type is endogenous - children with different observed and unobserved abilities & family background sorting into different school-types.

- Simple Estimation Equatioon

$$Ach_i = c + \gamma\, Male_i + \alpha_1\, DJ_i + \alpha_2\, DN_i$$
$$+\beta_1\, [DJ_i \times Male] + \beta_2\, [DN_i \times Male] + \epsilon_i$$

where:

- Ach_i: child's z-score in the Woodcock Johnson Test;
- Govt schools (female) are the reference group;
- $DJ_i = 1\, if\ JAAGO, 0\, otherwise$; $DN_i = 1\, if\ NGO, 0\, otherwise.$

Methodology: Dealing with Endogeneity

- Due to selection, we are worried that:
  - $cov\left(DN_i, \epsilon_i\right) \neq 0$, $cov\left(DJ_i, \epsilon_i\right) \neq 0$;
  - $cov\left(DN_i \times Male, \epsilon_i\right) \neq 0$, $cov\left(DJ_i \times Male, \epsilon_i\right) \neq 0$.
- One way to deal with this selection problem - use the Instrumental Variable Approach;
- However, we do not use this approach because:
  - IV estimates are inconsistent in the presence of choice based sampling [Solon et al. (2015)].
  - Adjusting IV estimator to make it consistent infeasible given our sample.

Matching: Least Squares Version

- We will deal with this endogeneity issue by assuming that there exists observable X, such that conditional on X, what school-type they go to is a coin toss.
  - This is called the Conditional Independence Assumption (CIA).
  - It is not clear that the least squares approach this works with choice-based sampling but it is useful expostitionally.

Methodology: Least Squares Version (contd.)

- Given CIA we can run OLS with X as regressors where X consists of family background and fluid intelligence:

$$Ach_i = c + \gamma\, Male_i + \alpha_1\, DJ_i + \alpha_2\, DN_i + \pi_1\, X_i$$
$$+\beta_1\, [DJ_i \times Male_i] + \beta_2\, [DN_i \times Male_i] + \pi_2\, [\, X_i \times Male_i] + \nu_i$$

Regression Estimates: Coefficients of Interest

- Coefficients of Interest
  - J vs. G (girls): α_1
  - G vs. N (girls): $-\alpha_2$
  - J vs. N (girls): $\alpha_1 - \alpha_2$
  - J vs. G (boys): $\alpha_1 + \beta_1$
  - G vs. N (boys): $-\alpha_2 - \beta_2$
  - J vs. N (boys): $\alpha_1 + \beta_1 - \alpha_2 - \beta_2$
- However, even in the absence of choice based sampling, OLS essentially compares all treatment to all comparisons and imposes functional form assumptions.
- Additionally, there no proof that OLS conditional on X is consistent given choice based sampling.

What if school-type affects IQ? (contd)

- Our IQ measures are intended to measure:
  - *fluid* intelligence (i.e. intelligence not affected by schooling) and NOT *crytallized* intelligence, as shown by Blair and Razza (2007); Fitzpatrick et al. (2014); Swanson (2008, 2011); Dauvier et al. (2014); Font (2014); Barac and Bialystok (2012); Hastings et al. (2014).
  - It is possible to improve Raven's score by teaching to the test; however, such training is not common in the average slum schools of Bangladesh.
  - Raven's increases by age but we account for that and age effects for K-BIT.

What if school-type affects IQ?

- Suppose schooling does affect IQ, and better school types raise IQ more.
  - Then it is straight-forward to show our J vs N and G vs N effects are downward biased.
  - Intuition: IQ is taking part of the credit for school type.
  - This would mean that the school-type effect is underestimated and the selection effect is overestimated.

What if school-type affects IQ? (contd)

- Let S = index of schooling quality and $\dfrac{\partial IQ}{\partial S} > 0$

$$\frac{dACH}{dS} = \frac{\partial Ach}{\partial S}\Big|_{dIQ=0} + \frac{\partial Ach}{\partial IQ}\frac{\partial IQ}{\partial S}$$

- We estimate:

$$\frac{\partial Ach}{\partial S}\Big|_{dIQ=0} = \frac{dACH}{dS} - \frac{\partial Ach}{\partial IQ}\frac{\partial IQ}{\partial S}$$

or,

$$\frac{\partial Ach}{\partial S}\Big|_{dIQ=0} < \frac{dACH}{dS}$$

$$\text{since } \frac{\partial Ach}{\partial IQ}\frac{\partial IQ}{\partial S} > 0$$

Alternatively Use Propensity Score Matching

- Another way to control for this selection:
  - Compare school-type 1 to school-type 2 using Propensity Score Matching which can be adjusted for choice based sampling.
  - Consider 3 Treatment Effects:
    - Average Treatment on the Treated (ATT);
    - Average Treatment on the Untreated (ATU);
    - Average Treatment Effect (ATE) [today's focus]
- For expository purposes, in what follows, we let the JAAGO individuals be the treatment students and NGO individuals be the comparison students.

Use Propensity Score Matching to Obtain Treatment Effects

- Average Treatment Effect on the Treated (ATT)
  - ATT captures the average effect, on achievement, of taking all students attending JAAGO schools and placing them in NGO schools.
- Average Treatment Effect on the Untreated (ATU)
  - ATU captures the average effect, on achievement, of taking all students attending NGO schools and placing them in JAAGO schools.
- Average Treatment Effect (ATE)
  - We can aggregate the ATT and ATU to get the ATE;
  - ATE is the effect, on achievement, of switching a randomly chosen student from JAAGO schools to NGO schools (or vice versa with a change of sign).

Defining Treatment Effects: ATT and ATU

- Specifically, we get ATT by comparing each child i in JAAGO with propensity score $P_1(X_i)$ to NGO observations j with similar propensity scores $P_1(X_j)$, where $P_1(X)$ is the probability of going to JAAGO schools versus NGO schools given characteristics X.

- Then, we get ATU by comparing each child j in NGO with propensity score $P_2(X_j)$ to JAAGO observations i with similar propensity scores $P_2(X_i)$, where $P_2(X)$ is the probability of going to NGO schools versus JAAGO schools given characteristics X.

- We use local linear matching to obtain the ATT and ATU.

Local Linear Regression Matching

- When estimating the ATT, local linear regression matching methods construct the counterfactual by solving the following minimization problem for each JAAGO student i and setting the counterfactual to $\widehat{\beta}_{0i}$:

$$\min_{\beta_0,\beta_1} \sum_{j=1}^{N_2} \left\{ Y_j - \beta_{0i} - \beta_{1i} \left[\hat{p}(x_j) - p_i \right] \right\}^2 K\left(\frac{\hat{p}(x_j) - p_i}{h} \right)$$

  where
  - $K(.)$ is the kernel weighting function;
  - h is the bandwidth;
  - j refers to NGO students whose total number is N_2.
- We impose the common support condition $0 < p(x_i) < 1$.

Local Linear Regression Matching (contd)

- The Average Treatment Effect on the Treated (ATT) is:

$$\frac{1}{N_1}\sum_{D_i=1}(Y_{1i}-\widehat{Y_{0i}})=\frac{1}{N_1}\sum_{D_i=1}(Y_{1i})-\frac{1}{N_1}\sum_{D_i=1}(\widehat{Y_{0i}})$$

where

- Y_{1i} : observed test score of child i going to JAAGO;
- $\widehat{Y_{0i}}$: predicted test score of JAAGO child i if s/he had gone to NGO; note that the minimization problem on the previous slide yields $\widehat{\beta}_{0i}=\widehat{Y_{0i}}$ as the counterfactual estimate of each JAAGO student i.

Local Linear Regression Matching (contd)

- When estimating the ATU, local linear regression matching methods construct the counterfactual by solving the following minimization problem for each NGO student j and setting the counterfactual to $\widehat{\alpha}_{0j}$:

$$\min_{\alpha_0, \alpha_1} \sum_{i=1}^{N_1} \left\{ Y_i - \alpha_{0j} - \alpha_{1j} \left[\hat{p}(x_i) - p_j \right] \right\}^2 K\Big(\frac{\hat{p}(x_i) - p_j}{h} \Big)$$

  where
  - $K(.)$ is the kernel weighting function;
  - h is the bandwidth;
  - i refers to JAAGO students whose total number is N_1.
- We impose the common support condition $0 < p(x_j) < 1$.

Local Linear Regression Matching (contd)

- The Average Treatment Effect on the Untreated (ATU) is:

$$\frac{1}{N_2}\sum_{D_j=0}(\widehat{Y_{1j}} - Y_{0j}) = \frac{1}{N_2}\sum_{D_j=0}(\widehat{Y_{1j}}) - \frac{1}{N_2}\sum_{D_j=0}(Y_{0j})$$

  - Y_{0j} : observed test score of child j going to NGO;
  - $\widehat{Y_{1j}}$: predicted test score of NGO child j if s/he had gone to JAAGO; note that the minimization problem on the previous slide yields $\widehat{\alpha}_{0j} = \widehat{Y_{1j}}$ as the counterfactual estimate of each NGO student j.

Local Linear Matching (contd)

- We combine the ATT and the ATU to obtain the ATE;
- Recall that ATE refer to the effect, on achievement, of switching a randomly chosen student from JAAGO schools to NGO schools (or vice versa with a change of sign).
- Note that we need similarly sized treatment and comparison groups to obtain a relatively precise ATE.

Propensity Score Matching (with Choice Based Sampling)

- Do matching while accounting for choice based sampling;
  - However, with choice based sampling, standard propensity score matching does not yield consistent estimates;
  - This problem can be addressed by the Heckman and Todd (2009) approach;
    - match on log odds ratio (LOR) of the estimated propensity score to obtain consistent estimates.
    - Note that LOR replaces p in the previous slides.
- Note that for the treatment effects, we again use bootstrapped standard errors clustered at the street level.

Trimming, Bandwidth Choice and Kernel-Type

- We need to trim the data to achieve common support.
  - Continuing with the example of *J vs. N* students, we do not want to estimate the ATT for *J vs. N* where there are no N students.
  - Similarly, we do not want to estimate ATU for *J vs. N* where there are no J students.

- We use 2 methods to obtain common support and I can talk about them after my presentation.

Trimming, Bandwidth Choice and Kernel-Type (contd)

- When choosing the bandwidth, we again use 2 alternative approaches - one of which is a fixed ex-ante bandwidth, while the other bandwidth is data driven.
- Bandwidth choice is considered important in this literature.
- The data driven approach of bandwidth choice is supposed to be optimal (but it is optimal for LLR, not LLR matching).
- We also use a Normal Kernel and an Epanechnikov Kernel.
- Our results are quite robust to all these choices.

Estimating ATE using Matching to Control for Selection (Case 3.1)

Table 6: **Estimating ATE using Matching to Control for Selection (Case 3.1: Common and Adjusted LPOLY)**

	Dependent Variable - Achievement Test Z-Score					
	Under-12/ Smaller Sample					Full Sample
	(1) **Mean Difference (no controls)**	(2) **Only Family Background (no IQ)**	(3) **Family Background & Raven's**	(4) **Family Background & K-BIT**	(5) **Family Background & Both IQ**	(6) **Family Background & K-BIT**
J vs. G (girls)	0.2459***	0.2252***	0.2234***	0.1902**	0.2123**	0.1843**
	(0.0752)	(0.0860)	(0.0828)	(0.0838)	(0.0861)	(0.0797)
bandwidth		0.23	0.25	0.28	0.49	0.25
J vs. G (boys)	0.1201	0.1194	0.1183	-0.0026	0.0145	-0.0209
	(0.0880)	(0.1035)	(0.0848)	(0.0960)	(0.0882)	(0.0878)
bandwidth		0.42	0.33	0.22	0.25	0.23
J vs. N (girls)	0.4937***	0.4685***	0.4012***	0.3057***	0.2910***	0.2490***
	(0.0931)	(0.0867)	(0.1050)	(0.0880)	(0.0944)	(0.0821)
bandwidth		0.44	0.44	0.27	0.28	0.36
J vs. N (boys)	0.5754***	0.5096***	0.3497***	0.2322*	0.2258*	0.2044
	(0.1457)	(0.1296)	(0.1264)	(0.1372)	(0.1309)	(0.1244)
bandwidth		0.23	0.34	0.34	0.58	0.41
G vs. N (girls)	0.2477***	0.1935**	0.1106	0.0376	0.0184	0.0048
	(0.0949)	(0.0981)	(0.0986)	(0.1006)	(0.1003)	(0.0947)
bandwidth		0.41	0.43	0.42	0.40	0.50
G vs. N (boys)	0.4554***	0.3981***	0.2276**	0.2828**	0.1970*	0.2695**
	(0.1175)	(0.1267)	(0.1151)	(0.1161)	(0.1157)	(0.1106)
bandwidth		0.37	0.50	0.43	0.56	0.70

(a) Bootstrapped standard errors in parentheses clustered at street level; (b) * $p < .1$, ** $p < .05$, *** $p < .01$; (c) Family background matching covariates consists of to child's age, family size, father absence dummy, father's schooling and mother's schooling; (d) By smaller sample, we refer to students aged between 5 years and 11 years and 11 months for whom the Raven's CPM IQ test is designed.

Impact of School-Type by Gender : Matching ATT, ATU and ATE Results

Table 7: Matching ATT, ATU and ATE Results with Both IQ Specification (Case 3.1: Common Trim with Adjusted LPOLY Bandwidth)

	(1) J vs G (F)	(2) J vs G (M)	(3) J vs. N (F)	(4) J vs. N (M)	(5) G vs. N (F)	(6) G vs. N (M)
ATT	0.2050**	0.0269	0.1774*	0.1912	-0.0334	0.1364
	(0.0993)	(0.0914)	(0.0997)	(0.1240)	(0.0974)	(0.1183)
ATU	0.2200***	0.0029	0.3895***	0.2590*	0.0626	0.2626*
	(0.0813)	(0.0969)	(0.0915)	(0.1490)	(0.1190)	(0.1434)
ATE	0.2123**	0.0145	0.2910***	0.2258*	0.0184	0.1970*
	(0.0861)	(0.0882)	(0.0944)	(0.1309)	(0.1003)	(0.1157)

Notes:
(a) Bootstrapped standard errors in parentheses are clustered at street level;
(b) * , ** , and *** denote statistical significance at the 10%, 5% and 1% level;
(c) Other matching covariates include child's age, family size, father absence dummy, father's schooling and mother's schooling

Estimating ATE using Matching to Control for Selection (Case 1.1)

Table 8: **Estimating ATE using Matching to Control for Selection (Case 1.1: Common and Rescaled Bandwidth)**

	Dependent Variable - Achievement Test Z-Score					
	Under-12/ Smaller Sample					Full Sample
	(1) **Mean Difference (no controls)**	(2) **Only Family Background (no IQ)**	(3) **Family Background & Raven's**	(4) **Family Background & K-BIT**	(5) **Family Background & Both IQ**	(6) **Family Background & K-BIT**
J vs. G (girls)	0.2459***	0.2228***	0.2227***	0.1964**	0.2053**	0.1835**
	(0.0752)	(0.0835)	(0.0815)	(0.0815)	(0.0837)	(0.0777)
bandwidth		0.10	0.10	0.12	0.13	0.12
J vs. G (boys)	0.1201	0.1106	0.1046	-0.0021	0.0247	-0.0225
	(0.0880)	(0.1034)	(0.0833)	(0.0969)	(0.0878)	(0.0876)
bandwidth		0.09	0.09	0.13	0.13	0.12
J vs. N (girls)	0.4937***	0.4637***	0.3951***	0.3011***	0.2852***	0.2402***
	(0.0931)	(0.0854)	(0.1033)	(0.0880)	(0.0943)	(0.0834)
bandwidth		0.21	0.20	0.21	0.21	0.28
J vs. N (boys)	0.5754***	0.5085***	0.3520***	0.2352*	0.2095	0.2069
	(0.1457)	(0.1325)	(0.1303)	(0.1386)	(0.1328)	(0.1260)
bandwidth		0.16	0.16	0.21	0.21	0.23
G vs. N (girls)	0.2477***	0.2008**	0.1087	0.0559	0.0463	0.0036
	(0.0949)	(0.0956)	(0.0987)	(0.1002)	(0.1011)	(0.0944)
bandwidth		0.19	0.21	0.23	0.24	0.26
G vs. N (boys)	0.4554***	0.4164***	0.2474**	0.2887**	0.2103*	0.2925**
	(0.1175)	(0.1259)	(0.1215)	(0.1199)	(0.1208)	(0.1136)
bandwidth		0.17	0.17	0.19	0.21	0.20

Notes: (a) Bootstrapped standard errors in parentheses clustered at street level; (b) * $p < .1$, ** $p < .05$, *** $p < .01$; (c) Family background matching covariates consists of child's age, family size, father absence dummy, father's schooling and mother's schooling; (d) By smaller sample, we refer to students aged between 5 years and 11 years and 11 months for whom the Raven's CPM IQ test is designed.

Treatment Effects for Different Bandwidths, Trimming and Kernel

Table 9: **Treatment Effects for Different Bandwidths, Trimming and Kernel (Both IQ Specification)**

	(1.1) **Common & Rescaled Bandwidth (Epanechnikov)**	(1.2) **Common & Rescaled Bandwidth (Normal)**	(2.1) **Manual Trim & Rescaled Bandwidth (Epanechnikov)**	(2.2) **Manual Trim & Rescaled Bandwidth (Normal)**	(3.1) **Common & Adjusted LPOLY Bandwidth (Epanechnikov)**	(3.2) **Common & Fixed LPOLY Bandwidth (Epanechnikov)**	(4.1) **Manual Trim & Adjusted LPOLY Bandwidth (Epanechnikov)**	(4.2) **Manual Trim & Fixed LPOLY Bandwidth (Epanechnikov)**
J vs. G (girls)	0.2053**	0.2117 **	0.1977**	0.2073**	0.2123**	0.2123**	0.209**	0.2090**
	(0.0837)	(0.0841)	(0.0819)	(0.0830)	(0.0861)	(0.0860)	(0.0841)	(0.0843)
bandwidth	0.13	0.13	0.11	0.11	0.49	0.49	0.25	0.25
J vs. G (boys)	0.0247	0.0132	0.0257	0.0143	0.0145	0.0145	0.0152	0.0152
	(0.0878)	(0.0868)	(0.0881)	(0.0869)	(0.0882)	(0.0870)	(0.0885)	(0.0873)
bandwidth	0.13	0.13	0.14	0.14	0.25	0.25	0.25	0.25
J vs. N (girls)	0.2852 ***	0.3096***	0.2757***	0.3013***	0.2910***	0.2910***	0.2830***	0.2830***
	(0.0943)	(0.0933)	(0.0955)	(0.0928)	(0.0944)	(0.0940)	(0.0955)	(0.0953)
bandwidth	0.21	0.21	0.24	0.24	0.28	0.28	0.35	0.35
J vs. N (boys)	0.2095	0.2310*	0.2024	0.2329*	0.2258*	0.2258*	0.2211*	0.2211*
	(0.1328)	(0.1285)	(0.1368)	(0.1314)	(0.1309)	(0.1289)	(0.1344)	(0.1334)
bandwidth	0.21	0.21	0.23	0.23	0.58	0.58	0.53	0.53
G vs. N (girls)	0.0463	0.0291	0.0194	0.0107	0.0184	0.0184	0.0141	0.0141
	0.1011	0.0924	0.0943	0.0901	0.1003	0.1007	0.0938	0.0938
bandwidth	0.24	0.24	0.20	0.20	0.40	0.40	0.52	0.52
G vs. N (boys)	0.2103*	0.1916	0.2153*	0.1934*	0.1970* 0	0.1970*	0.2027*	0.2027*
	0.1208	0.1179	0.1198	0.1159	0.1157	0.1147	0.1146	0.1131
bandwidth	0.21	0.21	0.2	0.2	0.56	0.56	0.57	0.57

Notes: (a) Bootstrapped standard errors in parentheses clustered at street level; (b) * $p < .1$, ** $p < .05$, *** $p < .01$; (c) Family background matching covariates consists of to child's age, family size, father absence dummy, father's schooling and mother's schooling.

Diagnostics

- Can we get a signal if matching is appropriate here, i.e., if the CIA holds - balancing tests.
- Given the propensity scores, look for a treatment effect on the Xs since we shouldn't see one.
  - Many ways of doing balancing tests, see, e.g. Smith and Todd (2005), Dehija (2005) etc;
  - Ours is another approach which has the advantages that it takes into account the fact that $p(X)$ is estimated and we do not use an 'eyeball' test.

Diagnostics (contd)

- Similar to obtaining the treatment effects, for the balancing tests we use local linear matching and adjust for the choice based sampling by matching on log odds ratio of the estimated propensity score.

- Note that we use the same trimmed sample, bandwidth and kernel-type in the balancing tests as in the matching exercises.

- The balancing tests again use bootstrapped standard errors clustered at the street level.

- Recall that, without matching, there are big differences in the conditioning variables, i.e. raw values do not balance.

Balancing Tests

Table 10: **Balancing Tests for Matching Estimations with Common Trim (Case 3.1 and Case 1.1)**

	Balance Test for Case 3.1: Matching Estimations using Common Trim and Adjusted lpoly bandwidth		Balance Test for Case 1.1: Matching Estimations using Common Trim and Rescaled bandwidth	
	(1) **LOR estimated using 7 covariates (including IQ)**	(2) **LOR estimated using 5 covariates (excluding IQ)**	(3) **LOR estimated using 7 covariates (including IQ)**	(4) **LOR estimated using 5 covariates (excluding IQ)**
J vs. G (girls)	0	0	0	0
J vs. G (boys)	0	1	0	1
J vs. N (girls)	0	2	0	2
J vs. N (boys)	0	2	0	2
G vs. N (girls)	0	2	0	2
G vs. N (boys)	0	2	0	2

Notes:
(a) LOR etimated using 5 covariates includes only family background matching covariates, i.e., child's age, family size, father absence dummy, father's schooling and mother's schooling.
(b) LOR estimated using 7 covariates includes the standard set of family background variables mentioned in (a) along with Raven's and K-BIT Z-scores.

Balancing Tests for Matching Covariates

- Results
  - We pass the balancing test for all variables when we use the full model to estimate the propensity score;
  - On the other hand, if we use only family background variables (excluding IQ) to estimate the propensity score, we fail the balancing test for variables like Raven's and K-BIT.

Key Findings

- Students in urban slums of Dhaka sorted across school-types.
  - 'Better' students sorted into JAAGO and government schools;
  - 'Weaker' students sorted into NGO schools.
- Fluid Intelligence plays a crucial part in controlling for selection.
  - Including fluid intelligence, especially K-BIT, which most developing country studies fail to account for, substantially reduces bias.
  - Note that K-BIT plays a larger role in reducing selection bias than Raven's.
- Family Background does not play much of a role in controlling for selection.

Key Findings (contd)

- Matching Results are insensitive to changing the trimming, bandwidth or kernel when doing Local Linear Regression Matching.
- Our sampling design allowed us to obtain relatively precise estimates when adopting good econometric practice. We took this econometric practice into account when collecting our data.
- Our empirical models pass balancing tests; it seems like these tests have some power.

Key Findings (contd)

- Boys and girls are differentially affected across the 3 school-types:
  - Govt. vs. NGO: Boys are better off at Govt. schools, but girls perform equally well at both school-types.
  - JAAGO vs. Govt: Girls are better off at JAAGO, but boys perform equally well at both school-types.
  - JAAGO vs. NGO: Both genders better off at JAAGO.
- School-types like JAAGO could reduce the gender gap in achievement.
  - If all girls going to Govt. schools could switch to JAAGO, gender gap would fall substantially.

Understanding the Gender Heterogeneity

- What school-type characteristics may be driving the gender difference in achievement between JAAGO and Govt schools?
  - Pro-male gender bias at Govt schools;
  - Share of female teachers lower at Govt schools;
  - Corporal punishment common at Govt schools.
- Find evidence of a gender differential for all above components in the literature.

Thank You

Distributions of Schools by School Type in Our Under-12/ Smaller Sample

Table 11: **Summary Statistics of Schools by School Type (Under-12/ Smaller Sample)**

	(1) No. of schools	(2) Mean (no. of students)	(3) Std. Dev	(4) Total (no. of students)
Govt	13	45.08	81.71	586
JAAGO	2	288	80.61	576
NGO	28	22.89	43.80	641

(a) that due to unavailability of administrative data, we are unable to present distribution of schools per school-type in the greater population.
(b) By smaller sample, we refer to students aged between 5 years and 11 years and 11 months for whom the Raven's CPM IQ test is designed.

Impact of School-Type by Gender: Approximate OLS Results

Table 12: Impact of School-type by Gender

	Dependent Variable - Achievement Test Z-Score					
	(1) Mean Difference (no controls)	(2) Only Family Background (exc. parents' educ)	(3) Only Family Background (inc (inc. parents' educ)	(4) Family Background & Raven's	(5) Family Background & K-BIT	(6) Family Background & Both IQ
J vs. G (girls)	0.246***	0.253***	0.241***	0.234***	0.203**	0.207**
	(0.0740)	(0.0719)	(0.0768)	(0.0777)	(0.0783)	(0.0792)
J vs. G (boys)	0.094	0.096	0.090	0.084	-0.028	-0.006
	(0.083)	(0.088)	(0.092)	(0.075)	(0.089)	(0.08)
J vs. N (girls)	0.487***	0.489***	0.432***	0.349***	0.255***	0.247***
	(0.089)	(0.093)	(0.085)	(0.096)	(0.072)	(0.082)
J vs. N (boys)	0.579***	0.581***	0.523***	0.374***	0.267**	0.239*
	(0.143)	(0.141)	(0.134)	(0.126)	(0.135)	(0.131)
G vs N (girls)	0.240**	0.236**	0.191**	0.114	0.0523	0.0395
	(0.0941)	(0.0970)	(0.0910)	(0.0928)	(0.0902)	(0.0911)
G vs. N (boys)	0.484***	0.485***	0.433***	0.290**	0.294**	0.245**
	(0.115)	(0.120)	(0.118)	(0.122)	(0.124)	(0.124)

Notes:
(a) Standard errors in parentheses clustered at street level;
(b) We report the estimates at the 10 percent, 5 percent and 1 percent significance level denoted by * , ** , & *** respectively;
(c) This table is derived from regressing achievement on family and child characteristics; school-type effects for the six comparison cases is calculated from the estimated coefficients of the OLS results;
(d) All regressions include standard set of controls (age, gender, father absent and family size).