IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18214

In BRAC We Trust? Comparing Schools for Disadvantaged Students in Dhaka's Slums

John C. Ham
Saima Khan

OCTOBER 2025

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18214

In BRAC We Trust? Comparing Schools for Disadvantaged Students in Dhaka's Slums

John C. Ham
NYU Abu Dhabi, NYU Wagner School of Public Service, IRP and IZA

Saima Khan
North South University

OCTOBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

In BRAC We Trust? Comparing Schools for Disadvantaged Students in Dhaka's Slums[1]

BRAC has over 40,000 schools worldwide. It is widely praised for serving disadvantaged students and for matching or outperforming government schools. Using data that we collected from Dhaka's slums, we test these claims. We find that BRAC serves the most disadvantaged students in our survey, but contrary to popular belief, BRAC students perform significantly worse than comparable students at other school types when we control for family demographics in a matching procedure. Anticipating our need to control for selection, we collected data on family demographics and the child's fluid intelligence; since the latter affects both types of school and student performance, it unambiguously should be included in the propensity score. Once we control for fluid intelligence, the performance difference with other NGO schools disappears. The gaps between government and JAAGO schools have narrowed, but they still remain large and statistically significant.

JEL Classification: C21, C83, I21, J24

Keywords: BRAC schools, math achievement, fluid intelligence, choice-based sampling, common support, matching

Corresponding author:
John C. Ham
Department of Economics
NYU Abu Dhabi
PO Box 129188
Saadiyat Island
Abu Dhabi
United Arab Emirates
E-mail: jch18@nyu.edu

[1] We would like to thank the National University of Singapore for funding our research when we were both at that institution. We benefited from comments at seminars at New York University Abu Dhabi, New York University, New York, Queens University, the Institute for Research on Poverty's summer workshop, 11th Biennial Conference of the Czech Economic Society, 2021 Chinese Economists Society Annual Conference, 2021 China Meeting of the Econometric Society, 14th Annual Meeting of the Portuguese Economic Journal, Australian Conference of Economists 2021, Korean Economic Review International Conference 2021, 23rd INFER (International Network for Economic Research) Annual Conference, Bolivian Conference on Development Economics, and Asian and Australasian Society of Labour Economics (AASLE) 2021 Conference. Finally, we appreciated detailed comments from Sebastian Galiani, Owen Ozier, Debraj Ray, and Geert Ridder.

1. Introduction

A major challenge for developing countries is making cost-effective, quality schooling available to a large proportion of the population, particularly those from marginalized communities. Often, government schools cannot cover the entire school-going population, and non-government organizations (NGOs) have to help fill the gap. In a country like Bangladesh, BRAC schools are NGO schools that teach a considerable number of students. BRAC states that it targets out-of-school children in marginalized communities not covered by the formal schooling system. Two stylized beliefs about BRAC schools are almost universally accepted. First, BRAC teaches children from quite disadvantaged backgrounds. Second, BRAC students perform at a level equal to, or higher than, Government schools.[2]

In this paper, we use data we collected to shed considerable light on these beliefs by asking two questions. First, which students go to BRAC in terms of their family backgrounds and fluid intelligence (IQ), as compared to those who go to (i) government schools (hereafter, GOV schools), (ii) a relatively new type of school (JAAGO schools)[3] and (iii) schools run by other non-government organizations (hereafter, ONGOs). Second, how much better or worse off would BRAC students be, in terms of math achievement, if they went to one of these other schooling modes?

Regarding the first question, we find that BRAC teaches the most disadvantaged students, and the difference between BRAC pupils and other students is the largest for GOV and JAAGO schools. Students at ONGO schools are somewhat advantaged when compared to BRAC students.

---

[2] See, e.g., Rosenberg (2013), UNESCO Institute of Lifelong Learning (2016), and the 'Our Work' section of BRAC's official website (BRAC International: Education 2025)

[3] We describe JAAGO schools in more detail below.

Hence, our data is consistent with the first stylized fact. However, regarding our second question, BRAC students score substantially lower on math achievement tests than their counterparts at the other school types when we do not control for selection, and these effects are all statistically significant. We then use propensity score matching, appropriately adjusted for our choice-based sampling scheme, to estimate treatment effects. Specifically, we estimate the impact on test scores of switching BRAC students to JAAGO, GOV, and ONGO schools while controlling for selection. We first control for only the child demographics and family background variables (hereafter background characteristics) in the adjusted propensity score; these estimated impacts (average treatment effects on the treated, or ATT) on a BRAC student if they were to go to a GOV school, a JAAGO school, or an ONGO school are very similar to the simple respective differences in achievement, respectively. However, when we further condition on the students' IQ scores, the treatment effects all fall considerably. Specifically, now the estimated impact on a BRAC student of going to a GOV school, a JAAGO school, and an ONGO school is 0.259 (0.109), 0.389 (0.107), and 0.056 (0.104) of a standard deviation in achievement, respectively.

The rest of the paper is organized as follows. Section Two briefly discusses the identification and estimation of school-type impacts. Section Three provides background information on BRAC schools and then focuses on the differences between the various school types. Section Four briefly reviews the existing literature on the effect of BRAC schools on student achievement. Section Five gives a general overview of the data we collected; in Section Six we compare the differences in student characteristics across the school types. Section Seven provides an overview of propensity score matching in the presence of choice-based sampling, which we used to guarantee a sufficient number of students in each school type. It also discusses the common support issue that is very important in our application. We present our results in Section Eight while Section Nine concludes the paper.

2. Identifying and Estimating the Impact of School-Type

Studies measuring the impact of different school types often face the issue that student assignment to the schools is not random. Several studies avoid this problem by using random assignment to pay tuition for some students in private schools, which allows them to avoid the selection problem by making school type endogenous and using the offer as an instrumental variable. This allows these studies to follow the recent tradition in Development Economics of using randomization to avoid making strong identifying assumptions (Banerjee and Duflo 2011).

Unfortunately, using random assignment to evaluate BRAC schools is infeasible for three reasons. First, none of the schools we consider require tuition payments. Second, as described below, all schools conduct interviews with parents (and sometimes the children), and both the school and the parents must agree before a given student is admitted to the school. Changing this process to one that randomly assigns students to the schools would change a school's product. Third, this approach may face powerful opposition from parents if subsidies are known to be unevenly distributed among students within the same school.

Another possible approach in the BRAC context is using instrumental variable estimation, treating the school choice as endogenous. However, to obtain a sufficient number of students in each school type, like in previous studies, we had to use choice-based sampling (hereafter CBS). IV is inconsistent in the presence of CBS (Solon, Haider, and Wooldridge 2015).

Instead, we use propensity score matching to address the selection problem of school allocation. A necessary condition for matching to be appropriate is that one controls for all observables that affect the selection process and the outcome variable, math achievement. Unlike

IV, we can adjust our matching procedure to account for CBS.[4] When selection has been considered previously in this context, studies have assumed that controlling for family background characteristics will meet this necessary condition; these studies have also used inconsistent estimators in the presence of choice-based sampling. However, we were skeptical of this approach if some measure of ability is not in the conditioning variables, since it will affect the selection process and the outcome variable.[5] Instead, we administered an internationally recognized IQ test to each child and used the resulting score as a conditioning variable in our matching procedure since it affects the outcome and school assignment. Including the IQ score in the conditioning variables dramatically reduced the estimated treatment effects. We would argue that we addressed our identification issue by collecting data on variables that would make our analysis credible, and that previous studies that only used family variables as conditioning variables are likely to have produced biased results.

3. A Comparison of How the Different Schools Operate

3.1 An Introduction to BRAC Schools

Established in 1972, BRAC is now one of the largest international NGOs in the world.[6] BRAC has many programs, from microloans to education, health care, skills development, and legal and human rights services.[7] Our focus here is on BRAC's Education Program in Bangladesh;[8] BRAC runs schools in both rural and urban areas in Bangladesh and other countries.

---

[4] See Heckman and Todd (2009) and Hahn et al. (2025).

[5] We ran a logit model to describe, e.g., the outcome of attending BRAC schools vs. GOV students. IQ was strongly significant in this estimated enrollment equation. Further, we ran a reduced-form regression equation where math aptitude was as function of family characteristics and IQ. Again, IQ was very significant in this equation.

[6] See brac (2020); note that *brac* is styled in lowercase on its official website, which is why we use lowercase in citations.

[7] Further details on BRAC can be found on BRAC's websites (brac 2025a; 2025b).

[8] Note that, Banerjee et al. (2015) and Banerjee, Duflo, and Sharma (2021) provide evaluations of BRAC's program aimed at improving life for the ultra-poor. They do not, however, evaluate BRAC schools.

Started in 1985, BRAC's Education Program focuses on five areas: non-formal primary education, the pre-primary schools' program, the adolescent development program, multipurpose community learning centers, and the mainstream secondary schools support initiative (UNESCO Institute of Life Long Learning 2016). This paper focuses on the first two areas: BRAC Primary Schools (BPS) providing non-formal primary education and the preprimary schools (BPPS), both in urban Bangladesh.

The BRAC model of non-formal primary schooling (BPS) has a *one-teacher, one-classroom* setup, with student intake occurring every four years. At BRAC, the five-year cycle of primary education is provided in four years.[9] Essentially, a four-year session starts in grade 1, with a group of approximately 30 children who are paired with one teacher. These same groups of children then progress through years 2, 3, and 4 with the teacher they were initially paired with. Thus, a BRAC non-formal schooling model does not operate in multiple grades with multiple teachers, as in formal schools. Further, once the four years are complete with one batch of students, the same teacher/branch starts with a new cohort; however, if there is insufficient demand, then the BRAC branch relocates to a new area where there is demand.

BPSs use BRAC textbooks in grades 1 to 3 and NCTB (National Curriculum and Textbook Board) textbooks in grades 4 and 5. Once they finish the 4-year primary cycle (i.e., grades 1-5) at a BPS, BRAC graduates go on to sit for the national exams, the Primary School Certificate Exams (PSCE). They can then subsequently integrate into the government schooling system or ONGO schools. Note that BRAC schools were free at the time of our survey.[10]

---

[9] See Rosenberg (2013) and Ministry of Foreign Affairs of the Netherlands (2011).

[10] BRAC's literature suggests the program is aimed at students who spend some time out of school. However, in our 2015/2016 data for Dhaka, there was no BRAC student who spent time outside school after age 5.

Meanwhile, BRAC's pre-primary schools (BPPS) provide one-year programs intended to prepare children for entry into primary schools. Like BPSs, BPPSs also use a one-classroom setup with a teacher recruited from the local community.

Although the BPPS program is designed for 5–6-year-olds, actual enrollment often includes a wider age range, with children as young as four years old and as old as seven years old (or more).[11] A 2015 survey found that 36% of BPPS students were aged seven years or older, while 5.5% were age four (Nath, Hossain, and Chowdhury 2018). In contrast, BPS generally admits students aged 8 to 10. [12] Since our sample includes children aged 4 to 13, it spans the full range of both BRAC school types.

Further, in our sample, students only indicated that they attended BRAC schools (but did not distinguish between BPS or BPPS). Hence, we cannot separate the effects of these two types of BRAC schools on achievement in our data; instead, we must evaluate the BRAC program as a whole.

3.2 Comparing the Different School Types

Since this paper compares performance across different school types, we briefly compare the schools' characteristics, admission criteria, teacher qualifications, the share of female teachers, and class sizes in Table 1.[13] As mentioned earlier, BRAC provides non-formal schooling. In contrast, JAAGO and GOV schools provide formal schooling. In what follows, a formal school is

---

[11] See Shahjamal and Nath (2008) and Nath, Hossain, and Chowdhury (2018).

[12] See Khan and Samadder (2010).

[13] Part of our discussion on school-type characteristics is informed by school interviews. We interviewed 14 ONGO schools, the two JAAGO schools, and two BRAC schools (out of the many BRAC schools operating in the area.) We lacked direct access to GOV teachers and were unwilling to pursue the lengthy process that might have allowed us to interview teachers in GOV schools. For BRAC school characteristics, given the extensive literature available, we rely on published reports and BRAC literature rather than our interviews from two BRAC schools. In fact, we generally do not draw on their responses in the discussion presented here.

defined as one that uses a permanent structure, has separate classrooms, and employs separate class and subject teachers.[14] The ONGO schools in our sample are a more heterogeneous mix, with some following BRAC's non-formal setup, with others adopting a formal structure. All schools, except JAAGO, teach in Bengali, except for their English language classes. In contrast, at JAAGO, all subjects are taught in English, except for the Bengali language class.

Since selection into the different school types is central to our paper, we first give an overview of the selection process used by BRAC, GOV, JAAGO, and ONGO schools respectively Of course, for each school, part of the selection process is the parents wanting to send their children to the school. BRAC's goal is to educate children who come from poor households, remote areas (for rural students), and minority groups.[15] The BRAC literature does not mention whether BRAC students and parents are generally interviewed as part of the admission process. The two BRAC schools we spoke to interviewed both the child and the parent(s) as part of the admission process.

GOV schools follow the Ministry of Primary and Mass Education (MoPME). Such schools start at grade 1 and continue to grade 5, with an increasing number of schools including a pre-primary section since 2010. The guardian of a six- to ten-year-old child must ensure that their child enrolls in a primary educational institution in the vicinity of their place of residence, which may be a government or non-government institution (Government of Bangladesh, 1990).[16]

[14] BRAC does have a small number of formal schools (Nath et al. 2005). We omit discussing them here since they are not represented in our sample

[15] See Ministry of Foreign Affairs of the Netherlands (2011). Note that minority groups include children with disabilities and ethnic minorities (Ministry of Foreign Affairs of the Netherlands 2011, 144).

[16] Note that according to the Primary Education Compulsory Act of 1990 (Government of Bangladesh 1990), guardians are exempt from sending their child to the local school given *justifiable reasons.* One such reason includes "*the impossibility of admitting a child in a primary education institute even if applied for*". This suggests that the Compulsory Primary Education Act accounts for the fact that children may end up not attending the local government school in the event like capacity shortages.

Based on our interviews with JAAGO personnel, a child admitted to JAAGO must fulfill the following criteria:[17] (i) he or she must be from a household where each family member earns less than or equal to BDT 2,000 (approximately USD 25.5) per month;[18] and (ii) they must not be enrolled in any other school, simultaneously. Further, JAAGO gives preference to children from single-parent households. JAAGO interviews both prospective students and their parents, viewing these interviews as an integral part of the admission process. The 14 ONGO schools in our 2018 survey employ various selection criteria, including the age of the child, family income, a priority for children from single-parent households, and admission tests. Notably, 11 out of 14 of the ONGO schools interviewed both the parents and the child, while the remaining schools interviewed only the parents.

We summarize other school characteristics in Table 1. Regarding teacher qualifications, BRAC and ONGO schools tend to have less strict requirements than JAAGO and GOV schools. Both types of BRAC schools require their teachers to have at least ten years of schooling,[19] be willing to teach part-time, and live in the community being served by its school.[20] Most of the ONGO schools in our sample reported requiring their teachers to have at least an HSC degree, i.e., twelve years of schooling. In the case of GOV schools, until 2013, the requirement for female teachers was at least an SSC degree (equivalent to completing 10 years of schooling), while the requirement for male teachers was at least an HSC degree.[21] However, in 2014, the government introduced stricter requirements, raising the minimum educational qualification for female and

[17] Note that these criteria applied during our survey period.
[18] In 2016, USD 1 = BDT 78.58 ('US Dollar to Bangladesh Taka History: 2016' n.d.).
[19] See Shahjamal and Nath (2008, 26).
[20] This is true for both BRAC primary and pre-primary schools; See Chabbot (2006, 8) and Shahjamal and Nath (2008, 26).
[21] See Campaign for Popular Education (CAMPE) (2015) and the 2013 circular published by the Directorate of Primary Education (DPE 2013).

male teachers to HSC and bachelor's degrees, respectively.[22] Meanwhile, JAAGO only employs teachers with at least a bachelor's degree (16 years of schooling).

Furthermore, the schools differ significantly in the amount they pay their teachers. BRAC schools paid average monthly salaries of BDT 1,250-1,650 in 2015 (UNESCO Institute of Life Long Learning 2016).[23] Teachers at JAAGO schools earned BDT 4,000 – 5,000 monthly in 2015-2016. Finally, GOV schools paid their teachers BDT 11,000 – 26,590 per month as per the 2015 National Pay Scale.[24] Unfortunately, even after extensive searching, we could not find a documented source for the salaries of the ONGO teachers. When we attempted to find these salaries using AI apps, one suggested that the ONGO teachers' salaries were in the neighborhood of the BRAC salaries, which seems credible to us, given the qualifications the ONGO schools require. However, another AI app suggested that their salaries were closer to the JAAGO salaries, which does not seem credible to us, given that the JAAGO schools require much higher qualifications.

The GOV schools invest more in teacher training than the other school types. According to the 2015 Campaign for Popular Education (CAMPE) report[25], about 90% of GOV school teachers had some form of training, such as a Certificate in Education (C-in-Ed), a Diploma in Education (DipEd), or a Bachelor in Education (B-in-Ed). The C-in-Ed is a one-year in-service training program designed for teachers teaching at the primary level, while the Diploma in Education (DipEd) is an 18-month-long basic training program created to replace the C-in-Ed. In particular, about 70-75 percent of GOV teachers have undergone the Certificate-in-Education in-

---

[22] Refer to CAMPE (2015) and DPE (2014b).

[23] In terms of the ONGO salaries, we could only find information on them via AI apps, which placed the salaries as somewhat below those at BRAC in 2015.

[24] See Alamgir (2024) and Finance Division, Ministry of Finance (2015).

[25] See CAMPE (2015).

service training program.[26] These in-service training programs, available to all GOV teachers, are provided and funded by the Government.[27] In contrast, BRAC, ONGO, and JAAGO schools have shorter training programs. Specifically, BPS teachers usually undergo 12-15 days[28] training before starting to teach, while BPPS teachers undergo about 9 days of training before teaching.[29] In contrast, JAAGO teachers undergo a month of training before teaching. Finally, among the 14 ONGO schools we surveyed, seven provide some form of training before teaching begins, with three providing 1-5 days of training and four providing 7-14 days. The remaining seven ONGO schools provide no pre-service training at all.

All school types include refresher courses or in-service training throughout the year. GOV school teachers undergo six one-day refresher courses throughout the year.[30] BPS teachers undergo about 30-35 refresher training days annually,[31] while BPPS teachers undergo 12 days of refresher training annually.[32] JAAGO teachers participate in need-based, interactive workshops tailored to address the challenges they face. Furthermore, JAAGO's Education Coordinators (similar to school principals) and selected teachers participate in regular training and workshops hosted by organizations such as Speed, BRAC, the British Council, and Teach for Bangladesh. Among the 14 ONGO schools that responded to our survey, three provide 1-7 days of in-service training, while

[26] See DPE (2014a, 2015, 2018a).

[27] See National Academy for Primary Education (NAPE 2020), BEPS (2002), Ministry of Foreign Affairs of the Netherlands (2011) and DPE Training Calendar 2017-2018 (DPE 2018b). These in-service training courses, held at training institutes across the country, are overseen by the National Academy of Primary Education which is under the Directorate of Primary Education, Ministry of Mass and Primary Education (NAPE 2020; BEPS 2002).

[28] See Chabbott (2006, 8).

[29] See Shahjamal and Nath (2008, 10).

[30] See BEPS (2002).

[31] See Ministry of Foreign Affairs of the Netherlands (2011) and Rosenberg (2013). The annual training for BPS seems high, but we do not know if each training day occupies a full day.

[32] See Shahjamal and Nath (2008).

four provide between 12-30 days of in-service training. The remaining ONGO schools state that they do not offer any in-service training.

In terms of gender balance, the share of female teachers is about 60-66% at GOV schools;[33] approximately 80% at JAAGO schools;[34] and 98-99% at BRAC schools.[35] Our survey of ONGO schools indicates that 85% of their teachers were women.

Administrative monitoring often focuses on teacher and headmaster absenteeism. JAAGO is the strictest school type here, and its Education Coordinators oversee the day-to-day operations of the schools. They monitor teacher presence, ensure teachers sign in and out, and reduce salary if there is a lack of punctuality and absenteeism. Specifically, at JAAGO, each day of absenteeism translates into a loss of salary for that day; additionally, teachers are marked as absent for one day if they are late for 3 days. At BRAC, administrative monitoring involves program officers visiting randomly assigned schools at least twice a week.[36] On the day of the visit, the program officer attends all the classes during school hours to evaluate the lesson delivery and teacher-student interactions.[37] We could not find anything in BRAC literature or reports on BRAC that describes how BRAC schools deal with teacher absenteeism or tardiness.[38] Among our 14 surveyed ONGO schools, 10 require teachers to sign in and, in some cases, sign out. Only one-third of these schools link teacher pay to teacher absenteeism.

---

[33] See Bangladesh Bureau of Statistics (2017) and DPE (2015).
[34] This is based on our interviews with JAAGO personnel.
[35] See Ministry of Foreign Affairs of the Netherlands (2011).
[36] See Chabbott (2006).
[37] See UNESCO Institute of Life Long Learning (2016); Reports on BRAC do not mention monitoring practices such as sign-in and sign-out or linking teacher's pay to absenteeism; the only form of monitoring they mention includes random visits by program officers.
[38] Again anecdotally, the two BRAC schools we were able to interview stated that teachers were not required to sign in or sign out and were not punished for absenteeism or tardiness.

Every GOV school has a Headteacher who maintains a teacher attendance register and these records are sent to the *Upazila (subdistrict) Primary Education Officer* (UPEO) on a quarterly basis.[39] Unauthorized absenteeism and excessive tardiness result in warnings from the UPEO, a 'show-cause' letter requiring the teacher to justify their absence/unpunctuality, and in some cases, the docking of salary, transfer to a less desirable school, or loss of job. However, these policies are not strictly enforced.[40]

GOV schools reported teacher absenteeism rates of around 15.3% in 2003, 16% in 2005, and 12.7% in 2014, based on unannounced visits by independent researchers.[41] However, school authorities categorized most of this absenteeism as authorized leave, official teacher-related duty, and GOV training programs.[42] In terms of teacher tardiness, a 2014 CAMPE survey reported that about 40% of GOV school teachers were late, on average by about 25 minutes, on the day of the survey.[43] Furthermore, the high rate of absenteeism and tardiness among headteachers at GOV schools raises questions about the effectiveness of teacher monitoring. For instance, 1 in 5 headteachers were found to be absent on the day of the survey in a 2004 study; another nationwide survey found tardiness to also be a problem, with one-third of the surveyed headmasters being late by 15-45 minutes.[44]

Contact hours vary across the school types in terms of both hours per day and days per year. GOV schools, the majority of which operate in a double-shift setup,[45] hold lessons for 2

---

[39] Tietjen et al. (2004, 77) details government policies regarding teacher absenteeism and tardiness.
[40] See Tietjen et al. (2004, 92-94).
[41] See Chaudhury et al. (2004, 9); Oxford Policy Management (OPM 2006, 108); CAMPE (2015, 53).
[42] See Chaudhury et al. (2004, 14); OPM (2006, 109); CAMPE (2015, 54).
[43] See CAMPE (2015, 54).
[44] Refer to Chaudhury et al. (2004, 12) and OPM (2006, 112) for information on headmaster absenteeism and tardiness, respectively.
[45] See DPE (2017, 114); OPM (2006, 98).

hours (grades 1 and 2) to 3.5 hours (grades 3 to 5) a day,[46] while BRAC states that its students are in school three hours (grades 1 and 2) to four hours (grades 3 to 5) a day.[47] Further, our survey suggests that the ONGO schools are in session approximately 2.4 hours (morning shift) to 2.9 hours (day shift) per day. Finally, the typical school day at JAAGO is 4 hours (day shift) to 5 hours (morning shift). In terms of days-per-year, BRAC schools are open for 265 to 270 days in an academic year,[48] JAAGO schools are open for approximately 265 days, while the average school year at GOV schools is 228 days.[49] The average days per year in the ONGO schools, based on our survey, is 247 days.

Lastly, GOV schools tend to have the largest classes, with an average of 52 students per class in 2014.[50] The average class size at BRAC schools is estimated to be between 30 and 35 students.[51] According to our survey, the average class size at ONGO schools is 37. Finally, the average class size at JAAGO is 31 students.

4. Reviewing previous work on BRAC schools

In this section, we focus on evaluation studies that explicitly compare BRAC schools to other types of schools in terms of their impact on student achievement. We identified a limited number of such studies (Ahmad and Haque 2011; Nath, Sylva and Grimes 1999; Nath, Shahjamal, Yasmin, Zafar and Kabir 2005; Nath, Roy and Hossain 2007; Nath and Hossain 2017), all of which focused on rural schools. Below, we discuss the insights from these evaluations more closely.

---

[46] See DPE (2017, 115); OPM (2006, 173-174).
[47] See Ministry of Foreign Affairs of the Netherlands (2011, 57) and Chabbott (2006, 8).
[48] See Ministry of Foreign Affairs of the Netherlands (2011, 57).
[49] For GOV schools, officially the length of the school year is 242 days, but in practice it is 228 days (DPE 2017, 155).
[50] Refer to CAMPE (2015, 56).
[51] The standard class size in rural BRAC schools is about 25-30; but it was increased to 30-35 for urban BPS (Kielland 2015, 29). Similarly, in rural BPPS, class size ranged from 26-30 (Shahjamal and Nath 2008, 10) but the BRAC literature does not specify class size for urban schools.

In a study that is in some ways closest to ours in terms of methodology, Ahmad and Haque (2011) examine the impact of receiving BRAC pre-primary (BPPS). They focus on students who received formal schooling in the later years, i.e., those who attended GOV or private schools. Their outcome interests are (i) final exam scores and (ii) getting first division in the primary completion exam. They began with a sample of primary school students in grade five who had received BRAC pre-primary education. Then they created a comparison group consisting of grade five primary school students who had not been exposed to BRAC pre-primary education.[52] They did not explore the impact of BPPS on students who went to non-formal schools.

They used propensity score matching based on family characteristics to examine the impact of BPPS schooling on outcomes (i) and (ii). [53] Regarding outcome (i), they found that attending BPPS increased final exam scores by a statistically significant 7.7%.[54] Concerning (ii), receiving BPPS schooling raised the probability of being in the first division by a statistically significant 13.3%. Their approach has three serious econometric differences with ours. First, it does not adjust for using choice-based sampling, so its matching estimates will be inconsistent. Second, it does not use a measure of ability as a conditioning variable, and we show below that a failure to control for ability leads to upwardly biased treatment effects. Finally, they fail to investigate the common support problem, which we find to be of great importance.

Nath et al. (1999) evaluated the BPS. They used logistic regression on five outcomes—basic education, life skills, numeracy, reading, and writing skills, including only statistically significant background characteristics as regressors in their final specification. Their explanatory

---

[52] They do not say the extent to which their control group had preschool training outside of BPPS.
[53] They included family income. Below, we argue that family income may well be a post-treatment variable if family income is affected by the length of the school day and hence should not be included as a conditioning variable.
[54] It is unclear how they chose the BPPS and non-BPPS samples.

variable of interest was a dummy variable coded one if a student attended a BRAC school and zero if the student went to a formal school, which could be a GOV or private school.[55] After controlling for family income,[56] age, access to media, and enrollment, BRAC students outperformed formal students in basic education, life skills, and numeracy. It is unclear how they chose the BRAC and non-BRAC samples, but given that taken together, GOV and private schools teach many more students than BRAC, they must have used some type of choice-based sampling.

Nath et al. (2005) compare BRAC BPS students to those attending community schools, with the latter being established and run by local communities with government financial support. These schools resemble formal schools in structure—they have separate classrooms for grades 1 to 5, employ 3–4 teachers, follow a five-year primary cycle, and use NCTB textbooks throughout. They use the results from the *Achievement of Basic Competencies* (ABC) test to measure achievement —a locally developed assessment tool designed by CAMPE and featured prominently in its *Education Watch* series (CAMPE 2000).[57] The ABC test measures 27 competencies (subtests) in Bangla, English, Mathematics, Social Studies, General Science, and Religious Studies. To achieve a competency/subtest, a student does not need to get every answer correct, but instead they need answer at least a designated number of questions correctly.

Nath et al. (2007) also use the number of competencies achieved (out of 27) on the ABC tests as their outcome variable. They compare 600 students across 30 BPS to 581 students from 30 government primary schools (GOV schools) in rural Bangladesh. Specifically, 30 BPSs were

---

[55] Refer to Nath, Sylva and Grimes (1999, 8).

[56] In this study, we do not use family income as a conditioning variable since family income may be affected by school choice through the length of the school day and/or school year,

[57] Note that Education Watch comprises of a series of studies, on literacy and basic education in Bangladesh, conducted by an advocacy and campaign network called Campaign for Popular Education (CAMPE) (Nath and Chowdhury 2008,3). Established in 1991, CAMPE is comprised of NGOs, researchers, educators, and Civil Society Organizations (CAMPE n.d.).

randomly selected from 30 areas. Then, for each selected BPS, a nearby GOV school was selected. This sampling scheme has several implications for their econometric model. First, it is an example of choice-based sampling and, as noted above, is likely to lead to inconsistent estimates. Second, they should treat the observations between students in the BPS and the paired GOV school as correlated when calculating their standard errors, since this is an example of cluster sampling. Finally, there remains the problem of selection bias in terms of the school attended; in our sample we find that the observable student characteristics differ significantly between BRAC schools and GOV schools in the same Dhaka slum. Nath et al. (2007) found that a higher percentage of BPS students (2.2%) achieved all 27 competencies compared to GOV students (0.9%). In terms of the mean number of competencies achieved, BPS students achieved 19.1 competencies, while GOV students achieved 18.6 competencies. However, it is not clear whether these differences are statistically significant since the authors do not report standard errors.

Nath and Hossain (2017) compare school completers at BPS to a national sample of primary school graduates from a national literacy survey conducted by *Education Watch* in 2016. All students were 6th graders, aged 11 or above, at the time of the survey. The study includes 472 students from 30 randomly selected BRAC schools and 328 students from the national sample, implying choice-based sampling. (Neither Nath and Hossain (2017) nor the 2016 Education Watch report[58] specify what school types are included in the national sample or how the national sample was chosen.) Using a locally developed Literacy Assessment (Ahmed, Nath, and Ahmed 2002), they found that the literacy rate was 95.7% among BRAC students and 91% among the national sample. They use multivariate logistic regression to predict literacy with conditioning variables including gender, age, mother's education, father's education, ethnicity, religion, having electricity

[58] See Nath and Chowdhury (2016).

at home, and household food security status. They do not control for IQ, which we find below is crucial to avoid selection bias.

Nath and Hossain (2017) present interesting results on the family characteristics of both types of students. They show that a higher share of BRAC students have parents with no schooling than those in the national sample. For instance, 38% of BPS students have mothers with no schooling as compared to 34% for the national sample. Similarly, 45% of BPS students have fathers with no education compared to 38.7% of national sample students. Additionally, BRAC households were found to be poorer as measured by food security status. These figures are comparable to those found in our data.

Finally, one could argue that Romero, Sandefur and Sandholtz (2020,2022) provide a randomized evaluation of BRAC schools. In 2016, the Liberian government delegated management of 93 public schools to eight contractors, one of whom was BRAC. Each contractor had a randomized comparison group. BRAC did not significantly affect the performance of their students. However, this evidence is far from conclusive since in this case BRAC did not hire its own staff but rather kept the public-school teachers. Given that, teacher recruitment and training is a substantial component of the BRAC experience in Bangladesh , we argue that these results are not comparable to ours.

5. Our Data Set and the Choice of Survey Instruments

5.1 Our Data Collection

In 2015-2016, we collected our own data on 1936 children (aged 4-13) attending school in the two slums of Dhaka where JAAGO operated. We avoided selecting students from outside our two slums, as NGO students in outside areas may differ from those in the slums. Here, we are guided by the findings of Heckman, Ichimura, and Todd (1997) that propensity score matching works better when treatments and comparisons are from a common economic environment.[59] Our aim was to collect a stratified sample of approximately 600 students in each of GOV schools, JAAGO schools, and NGO schools from these two slums. We also collected data on the family background and IQ of each student. Our collection of IQ data was based on the knowledge that we would need convincing conditioning variables to address the inherent selection problem in student assignments. In other words, we addressed our selection problem by collecting richer conditioning variables in addition to those used in previous studies in both the labor and development literature. Finally, our outcome variable was the score on an internationally recognized math achievement test described below.

To ensure a sufficient number of students in each school type, we used choice-based sampling (CBS). We chose CBS over random sampling, since the latter would have resulted in very few JAAGO students. We collected the data by 'streets', where a street is a long road and it includes smaller roads off it. We started on a street with at least one JAAGO student,[60] and then collected data on more JAAGO students, GOV students, and NGO students (which included both ONGO and BRAC students) on that street. We continued to add students until we had approximately an equal number of students from the three school types in each age group on the street.[61] We collected data from 26 of these 'streets'; hence, for constructing standard errors, we have 26 clusters (Abadie et al. 2023). In actuality, we collected data on 607 JAAGO students, 618 GOV students, and 711 NGO students (who are either BRAC or ONGO students).

[59] To ensure that our evaluations of the school types are not affected by a given type of school having more migrant families, we required that students be residents of their given slum since 2010.

[60] JAAGO gave us the addresses of their students.

[61] Hence, we constructed choice-based sampling on school type and age; since selecting on age is an example of exogenous sampling, we need only be concerned with the selection on school type.

When conducting our survey, we took many steps to ensure that our data was high-quality. First, we hired experienced enumerators and invested heavily in monitoring. Second, we conducted interviews with children outside of school hours. Since schools operate on morning and day shifts, children attending morning shifts were interviewed in the afternoon, and those attending day shifts were interviewed in the morning. Children were not interviewed during meals. Third, enumerators were randomly assigned to different sub-areas and across the times of the day. Fourth, since our internationally recognized tests (described below) are tedious to administer, we recorded this part of the interview and hired a team of audio auditors, who listened to each interview to identify mistakes and/or 'cheating' on the survey instruments. Enumerators who were caught cheating were immediately dismissed, and a nontrivial number of such enumerators were involved. All interviews with these enumerators were deleted and redone by trusted enumerators. Fifth, we hired a team of data editors who specialized in reviewing the information in household questionnaires and verifying contradictory or missing information with the enumerators and surveyed households. Sixth, all survey instruments were graded twice by two separate data editors. Finally, to minimize data entry errors, we used a double-entry system for our data.

5.2 Our Survey Instruments

When choosing tests to measure mathematics achievement and IQ, respectively, we had two requirements. Given that our sample includes children aged between 4 and 13, with younger children more likely to be illiterate, we needed a test that (i) would not require literacy (i.e., tests that could be verbally administered) and that (ii) would be suitable for the above age range. Note that the GOV, BRAC, and ONGO students are taught in Bengali, while the JAAGO students are taught in English. We chose to use only the mathematics subtests as our achievement measures, since performance on these tests is less dependent on language skills and the medium of instruction compared to tests in other subjects. Furthermore, we administered the tests in Bengali to the GOV, BRAC, and ONGO students, and administered the same tests to the JAAGO students in "Banglish," i.e., we retained technical terms in English.

To choose our achievement test, we reviewed those used in both the developed and developing country literature. Tests used in developing country studies are usually developed locally, based on local curriculum, and are not standardized across countries (Newhouse and Beegle 2006; French and Kingdon 2010; Goyal and Pandey 2009; Goyal 2009; Hsieh and Urquiola 2006; Lucas and Mbiti 2014; Pal and Saha 2014). Note that the studies discussed in Section 3 above use Bangladesh-specific tests.

In the context of developed countries, certain standardized achievement tests are frequently used in the fields of Economics, Education, Child Psychology, and Developmental Psychology. These include the Woodcock-Johnson Test of Achievement III (hereafter WJ-III), the Peabody Individual Achievement Test (PIAT), the Wechsler Individual Achievement Test (WIAT), and the Wide Range Achievement Test (WRAT). We chose the WJ-III test since it is most frequently used in the literature,[62] is applicable for our age range (4 to 13) and class range (pre-school to grade 5), and can be administered verbally.[63] Specifically, we used three Math subtests: Test 10 (Applied Problems), Test 18A (Quantitative Concepts), and Test 18B (Number Series). We calculated the average score for each student on the three tests as our measure of achievement.

[62] The following papers use the Woodcock Johnson Test: Burchinal et al. (2014), Cameron et al. (2012), Crosnoe et al. (2010), Davis-Kean (2005), Denton et al. (2013), Duncan et al. (2007), El Nokali, Bachman, and Votruba-Drzal 2010, Espy et al. (2004), Ferrer and McArdle (2004), Fitzpatrick et al. (2014), Floyd, Bergeron, and Alfonso (2006), and Gormley et al. (2005).
[63] See Woodcock, McGrew, and Mather (2001), Mather and Woodcock (2001), and Dean (2011) for applicable age and grade specifications.

To capture fluid (innate) intelligence or IQ, we use the Kaufman Brief Intelligence Test-2 (K-BIT), as it can be administered verbally and meets our age requirements.[64] K-BIT is extensively used in the psychology literature to measure fluid intelligence,[65] i.e., intelligence that is not affected by schooling unless schools teach to the test. K-BIT is a more comprehensive measure of fluid intelligence than the Raven's Matrices measure. Some readers may worry that school type affects the K-BIT scores, with better schools potentially raising the K-BIT scores more than poorer schools. We show below that our estimated treatment effects will be understated in such a case.

Since the Woodcock-Johnson Test and the K-BIT scores tend to increase with age, we use their age-adjusted z-scores for both tests. In other words, for student *i* in age group a for test *k,* we use $Z_{ik} = (X_{iak} - X_{ak})/\sigma_{ak}$, where $X_{iak},\ X_{ak},\ and\ \sigma_{ak}$ are the student's achievement in test *k,* the mean value of achievement for students aged a in test *k*, and the standard deviation in achievement for students in age group *a* for test *k.*

6. Average Student Characteristics Across the Different Schools

We investigate selection across different school types in terms of four key variables. First, we consider average monthly family expenditure (deflated by the square root of family size) across school types. We asked for family expenditure as a proxy for family income since families are often reluctant to reveal their monthly earnings but more readily share their average monthly expenditures with enumerators. The second and third variables of interest are the father's average schooling and the mother's average schooling, measured in terms of the highest grade completed across school types. Finally, we compare the students' K-BIT mean scores across school types.

---

[64] The K-BIT is designed for ages 4 to 90 years (Kaufman and Kaufman 2004).

[65] For example, the following papers use the K-BIT test: Arenas et al. (2024), Barac and Bialystok (2012), Cheng et al. (2024), Font (2014), Hastings, Kahle, and Nuselovici (2014), Pitts and Mervis (2016), Rodriguez-Martínez et al. (2023), and Waber et al. (2001).

In Table 2, we present the means of these four variables across school types. The deflated average monthly expenditure ranges from approximately BDT 5,182 to BDT 6,145 across school types.[66] The average father's schooling by school type ranges between 2.5 and 3.7 years of schooling, while the average mother's schooling ranges between 2.1 to 3.8 years of schooling. Further, the average normalized K-BIT Z-scores range from -0.17 to 0.27.

For expositional ease, we show the differences in means between the BRAC students and students from the other school types for our four variables in Table 3. For example, column (1) of this table is obtained by subtracting the estimated GOV mean from the estimated BRAC mean for our four variables of interest. Column (2) subtracts the estimated JAAGO means from the estimated BRAC means, while column (3) subtracts the estimated ONGO means from the BRAC means. Thus, in the table a negative value indicates that the average BRAC student has a lower value for the variable compared to the respective school type. All of the entries in Table 3 are negative and statistically significant (except for the difference in family consumption between BRAC and ONGO students),[67] indicating that the BRAC students are, in general, disadvantaged compared to students from the other school types. Furthermore, the differences are more pronounced when comparing the JAAGO and GOV students to the BRAC students.

7. Estimation Strategy

Here we discuss how we address the selection issue. Since, as noted above, none of the schools charge tuition, the usual practice of eliminating selection bias by randomly assigning tuition subsidies would not make sense here. One could conduct an RCT by randomly assigning families a subsidy for a specific school type. Still, it is not clear how much impact this would have

[66] In 2016, USD 1 = BDT 78.58 ('US Dollar to Bangladesh Taka History: 2016' n.d.).

[67] When calculating standard errors for all of the tables in this section we account for our cluster sampling design.

given that the allocation of students into the different school types reflects both the preferences of the parents and of the schools. Such a scheme may also run into trouble if some students at a given school type are not offered the subsidy while others are, and the students compare their subsidies (or the lack of them). Instead, we turn to matching, which has been employed in numerous evaluation studies in both developing and developed countries. It is one of the four quasi-experimental methods discussed in the World Bank's DIME Wiki,[68] and has recently gained traction in the epidemiological literature and in medical research.[69]

7.1 Propensity Score Matching in General

Let $D_i = 0$ if student *i* attends a BRAC school and let $D_i = 1$ if student *i* attends a GOV school. The outcome variable is defined as the student's *Z* score on the WJ math achievement test. Define Y_i^0 as student *i*'s achievement in a BRAC school and define Y_i^1 as their achievement if the student were to attend a GOV school.[70] Our first goal is to estimate the effect on Z scores for achievement if a BRAC student *i* were to move to a GOV school

$$ATET = E(Y^1 - Y^0 \big| D = 0) = E(Y^1 \big| D = 0) - E(Y^0 \big| D = 0). \tag{1a}$$

We can observe the second term on the right-hand side of (1a) since we observe the actual performance of this student in a BRAC school. But we cannot directly observe how the student would perform in a GOV school – the (counterfactual) first term on the right-hand side of (1a). We use matching to estimate this counterfactual.

---

[68] From https://dimewiki.worldbank.org/, (material in italics added), DIME Wiki aims to fulfill a unique role as a publicly available, easily searchable, up-to-date resource focused on practical implementation *(of program evaluation)* rather than theory (The World Bank, n.d.).

[69] See, e.g., Austin (2009), Wu et al. (2015) and studies cited therein, as well as the studies listed at https://pubmed.ncbi.nlm.nih.gov/25643113/.

[70] For ease of exposition, we will drop the individual subscript, where doing so does not create confusion.

Also, we will estimate the (hypothetical) effect on achievement if a GOV student were to go to a BRAC school

$$ATU = E(Y^1 - Y^0 \mid D = 1) = E(Y^1 \mid D = 1) - E(Y^0 \mid D = 1). \quad (1b)$$

Again, we will have to impute the second term in (1b). For matching to be a valid estimator here, certain assumptions must hold. In particular, the fundamental assumption underlying matching estimators is known as the Ignorable Treatment Assignment Assumption (ITA), as described by Rosenbaum and Rubin (1983), or the Conditional Independence Assumption (CIA) as discussed in Lechner (2000)

$$\left(Y^0, Y^1\right) \perp D \mid X. \quad (2)$$

If (2) holds, we can treat the allocation of a child to a school type as random, conditional on X. The crucial factor is that X contains any variable that affects both selection and the potential outcomes. Since matching on X becomes impractical as the number of conditioning variables increases, Rosenbaum and Rubin (1983) show that (2) implies

$$\left(Y^0, Y^1\right) \perp D \mid p(X), \quad (3)$$

i.e., if the CIA holds when we condition on X, then it also holds when we only condition on the propensity score $p(X) = \Pr(D = 1 \mid X)$.

For a matching estimator to make sense, there should be at least one variable that affects school allocation but not math achievement; otherwise, there is no way to explain why some students attend BRAC schools while observationally equivalent students attend GOV schools. However, the researcher does not have to observe the excluded variable(s). For example, in our application, the excluded variable may be the distance from the student's home to the closest

BRAC school versus the distance to the closest GOV school. However, whether we observe these distances is irrelevant.

To estimate both the ATT and ATU, we need the common support condition

$$0 < \Pr(D = 1 | X) < 1. \quad (4)$$

Equation (4) requires a positive probability (less than one) of observing a GOV student at each level of $X = x'$. Note that if $\Pr(D = 1 | X = x') = 0$, then at $X = x'$ we will observe only BRAC students; this is ruled out by (4).

We would note that if there are substantially more students of type T1 (e.g., BRAC) than type T2 (e.g., GOV) for a given, $X = x'$, we may encounter problems when estimating the ATT for the T1 students, as a relatively small number of T2 students will be repeatedly used to construct these estimates. For expositional ease, we will follow the literature and refer to avoiding this situation as part of imposing the common support condition.

7.2 Choice of Specific Matching Estimator

Most studies use either local (polynomial) regression (LPR) matching or inverse propensity score (IPS) matching.[71] The former is advocated by Heckman, Ichimura, and Todd (1998) while the latter is suggested by Hirano, Imbens, and Ridder (2003). In standard LPR matching, one generally uses a probit or logit model to estimate the propensity score $\hat{p}(x)$ for each student. Then, for the ATT discussed above, one solves the following optimization problem for each BRAC student i

[71] Some studies use nearest neighbor matching, where only the non-treatment individual with the closest propensity score is used to form the counter factual. This approach is inefficient in terms of the small amount of data used for the counterfactual.

$$V_i^* = \min_{\gamma_{0i},\gamma_{1i},\ldots,\gamma_{Li}} \sum_{j=1}^{N_0} \left\{ Y_j^1 - \sum_{l=0}^{L} \gamma_{li} \left(\hat{p}(x_j) - \hat{p}(x_i) \right)^l \right\}^2 K\left(\frac{\left(\hat{p}(x_j) - \hat{p}(x_i) \right)}{h_i} \right), \tag{5}$$

where N_0 is the number of GOV students. The value of $\hat{\gamma}_{0i}$ from (5) is the counterfactual estimate of what achievement would be for BRAC student *i* in a GOV school. We can construct the ATT by averaging $\hat{\gamma}_{0i}$ over the BRAC students. In (5), $K(\bullet)$ is a kernel weighting function, such that $K(w) = 0$ if $\mathrm{w} \geq 1$, and h_i is the bandwidth. To calculate the ATU, for each government student *j* we use

$$W_j = \min_{\alpha_{0ji},\alpha_{1j},\ldots,\alpha_{lj}} \sum_{i=1}^{N_1} \left\{ Y_j^0 - \sum_{l=0}^{L} \alpha_{lj} \left(\hat{p}(x_j) - \hat{p}(x_i) \right)^l K\left(\frac{\left(\hat{p}(x_j) - \hat{p}(x_i) \right)}{h_j} \right) \right\}^2, \tag{6}$$

where N_1 is the number of BRAC students. We then set the estimated ATU equal to the average of the $\hat{\alpha}_{oi}$ over the GOV students.

Researchers usually find that their results are insensitive to the choice of a standard kernel. The bandwidth *h* is an important parameter, as it defines the set of GOV students used as comparisons for each BRAC student in (5).[72] Here, one faces a tradeoff in that a larger value of *h* will result in additional, but less suitable, matches for a given BRAC student, while a smaller value of *h* will lead to fewer, possibly more suitable, matches. Below, we consider both data-driven and fixed bandwidths when implementing matching.

Finally, L in (5) and (6) is the order of the polynomial: $L=0$ corresponds to local kernel regression (LKR) matching, $L = 1$ corresponds to Local Linear Regression (LLR) matching, $L =$

[72] We note below that one can also affect the number of comparisons by imposing a trimming rule to meet the common support condition

2 corresponds to local quadratic regression (LQR) matching and $L = 3$ corresponds to local linear cubic regression (LCR) matching. Ham, Li, and Reagan (2011) investigated the issue of the degree of the polynomial in (5). They note that Fan and Gijbels (1996, Section 3.3.2) show that, *ex-ante,* LLR matching dominates LKR matching, and LCR matching dominates LQR matching.[73] They also find that LLR dominates LCR in their application, since using LCR overfits their data. Thus, we use LLR matching.

We will use the block (by street) bootstrap to calculate standard errors to account for our cluster sampling. Ham et al. (2011) show that the bootstrap performs well for a matching estimator based on LLR. Where we use a data-driven bandwidth, we recalculate the bandwidth for each bootstrap sample to ensure that our standard errors are consistent.

One concern regarding our approach is that a student's K-BIT score is necessarily recorded after the student has attended a given school type. The K-BIT test is intended to be independent of instruction unless schools specifically 'teach to the test'. It is essential to consider how our estimated impacts might change if this assumption is not valid. First, consider a scenario where all school types increase the K-BIT score by the same amount. In this instance, the estimated treatment effects will remain unchanged, as this would not alter which treatment and comparison students are compared. Next, consider the more realistic situation where better schools, such as the GOV schools, yield greater improvements in K-BT scores than less-effective ones, like BRAC. In this case, we would likely match a given BRAC student with a lower IQ GOV student, as the GOV school would artificially inflate measured fluid intelligence. Consequently, we would underestimate the *potential* performance of a BRAC student in a GOV school, which suggests that

[73] Specifically, they show that using an even order polynomial of degree k, as opposed to one of $k+1$, increases bias in, but does not reduce the variance of, local regression estimators.

we would also underestimate the ATT for a BRAC student moving to a GOV school. Hence our ATT may well be conservative. In this case, we are likely to underestimate the ATU as we would be comparing a GOV student to a more able BRAC student, given the GOV schools inflate their students fluid intelligence.

7.3 Matching with a Choice-Based Sample

We must adjust the matching estimator in two ways to account for the use of choice-based sampling when collecting our data. First, Heckman and Todd (2009) show that with choice-based sampling, in (5) and (6) one should condition on the estimated log odds ratio $\hat{q}_i(\underline{X}_i) = \hat{p}(\underline{X}_i)/[1-\hat{p}(\underline{X}_i)]$ instead of on the estimated propensity score, $\hat{p}(\underline{X}_i)$, from the choice-based sample. Secondly, Hahn et al. (2025) show that the correct version of the ATT with choice-based sampling is based on solving the following minimization problem for each BRAC student *i*

$$V_i^* = \min_{\gamma_{0i},\gamma_{1i},\ldots,\gamma_{Li}} \sum_{j=1}^{N_0} \left\{ Y_j^1 - \sum_{l=0}^{L} \mu_{li} \left(\hat{q}(x_j) - \hat{q}(x_i) \right)^l \right\}^2 K\left(\frac{\left(\hat{q}(x_j) - \hat{q}(x_i) \right)}{h_i} \right). \qquad (5)'$$

Further, a simple manipulation of their results implies that one can obtain an estimate of the ATU by carrying out the following for each GOV student j

$$W_j^* = \min_{\delta_{0j},\delta_{1j},\ldots,\delta_{Lj}} \sum_{i=1}^{N_1} \left\{ Y_i^0 - \sum_{l=0}^{L} \delta_{lj} \left(\hat{q}(x_j) - \hat{q}(x_i) \right)^l \right\}^2 K\left(\frac{\left(\hat{q}(x_j) - \hat{q}(x_i) \right)}{h_j} \right). \qquad (6)'$$

Finally, their results imply that researchers should use the following expression for the average treatment effect [74]

$ATE = [\pi ATE] + [(1-\pi)ATU],$

[74] See also Zhang, Hu, and Liu (2019).

where π is the true fraction of BRAC students relative to [BRAC students + GOV students] in the population. In our application we had difficulty estimating π for the different school types. We were able to find data on the share of some types of schools at the national level.

7.4 Imposing Common Support

To obtain our estimates of the respective treatment effects, we need to impose the common support condition. Consider Figures 1-3, which illustrate the Log Odds Ratios (LORs) for BRAC students compared to Government (GOV) students, BRAC students compared to JAAGO students, and BRAC students compared to ONGO students, respectively. In each Figure, we have created 30 bins for the LORs, with the BRAC students in grey and the other students in red. As noted above, when estimating an ATT, we can only sensibly estimate the impact of attending a GOV school for a BRAC student when there is a sufficient number of GOV students with similar LOR values. Analogously, to calculate the ATU, we need to estimate the effect of going to a BRAC school for each GOV student *j*, and thus we will need an adequate number of BRAC students with similar values of $\hat{q}_j(\underline{X}_j)$.

[*Insert Figures 1, 2 and 3 here.]*

One trimming method is given by Stata's *psmatch2's common trim* routine (hereafter *common trim*). For calculating the ATT of moving a BRAC student to a GOV school, this routine drops BRAC observations whose LORs are higher than the maximum or lower than the minimum of the GOV students' LORs. Conversely, for calculating the ATU of moving a GOV student to a BRAC school, *common trim* drops GOV observations whose LORs are higher than the maximum or lower than the minimum of the LORs of the BRAC students.[75] For example, in Figure 1, when

[75] See Leuven and Sianesi (2003).

using *common trim,* we would delete only bins 29 and 30 when calculating the ATT, and only bin 1 when calculating the ATU. Thus, it will produce different trimmed samples for the ATT and ATU, respectively. Following this approach, in Figure 2, when considering BRAC vs JAAGO, one would drop bins 27-30 for the ATT and bin 1 for the ATU. When comparing the BRAC vs. ONGO students in Figure 3, common trim will delete bin 30 for calculating the ATT; for estimating the ATU, common trim will delete bins 1, 5, and 6.

However, it may be better to use the same sample to calculate the ATTs and ATUs; otherwise, we will use different samples to calculate each term in the ATE's formula (7). Thus, one could consider a symmetric *common trimmed* sample for BRAC vs. GOV by deleting bins 1, 29, and 30.[76] A symmetric *common trim* of the BRAC vs. JAAGO students would delete bins 1 and 27-30. Finally, for the BRAC and ONGO students, a symmetric *common trim* procedure would delete bins 1, 5, 6, and 30.

The major drawback with the *common trim* procedure is that it will allow bins with only one type of student, such as 3, 26, and 27 in Figure 1, to remain. In this case, there are no suitable comparisons when calculating the ATT and ATU.[77] A more serious issue with *common trim* (which also occurs in our data) is that it creates bins where we have many GOV students but only a few BRAC students (or vice versa). Loosely speaking, this would describe bins 2 and 4-11.[78] Furthermore, one would also have to be concerned about bins with many BRAC students but only a few GOV students; again, one might argue that this is the case for bins 23-25 and 28.[79] Our

---

[76] Heckman, Ichimura, and Todd (1997) and Ham, Li, and Reagan (2011) both use a symmetric trimming algorithm.
[77] Computationally, this problem could be mitigated by the use of a reasonable bandwidth, which would delete most of the observations in these bins.
[78] Note that we discussed the removal of bins 1 and 3 above.
[79] Again, we discussed deleting bins 26, 27, 29, 30 above.

experience suggests that matching when only a small number of one group are left in some bins as comparisons is likely to produce quite noisy treatment effects.

The upshot is that one needs a rule defining the bins to be dropped where there is an insufficient number of students from one of the groups. We do this by proposing a new method of obtaining common support: *the 1:5 trimming method.* We continue to break down the LOR distribution into 30 equal-sized bins, as shown in Figures 1-3. We delete all observations if they are in a bin where the ratio of treated to comparison observations is more than 5.0 or less than 0.2. The result of using this approach for the BRAC vs GOV students is also shown in Figure 1, where we would eliminate bins 1, 3-6, 8-11, 24, 26, 27, 29, and 30 (these are denoted by blue lines and/or blue dots). Note that the *1:5 trimming method* automatically deletes bins with only one type of student.[80]

However, given that there is a significant correlation of -0.355 between the log odds ratio and student achievement scores (before trimming), eliminating GOV students with the lowest log odds ratio results in the deletion of a disproportionate number of top-performing GOV students. While this may not pose a significant problem when estimating ATTs, it is a concern when calculating the ATUs, the average effect of switching GOV students to BRAC schools. In such a case, to treat the estimated ATU as representing the overall ATU, one would need to assume that the treatment effects of moving a GOV student to BRAC are homogeneous across all GOV students, including those in the dropped bins. We find this assumption to be unrealistic and cannot recall it being made previously in the literature.

[80] One could possibly mitigate this problem by considering individuals in neighboring bins. We do not pursue that here.

A similar concern arises in the BRAC vs. JAAGO comparison depicted in Figure 2, where the *1:5 trimming algorithm* eliminates bins 1, 3–14, and 26–30 (again highlighted with blue lines and dots). In contrast, this issue is much less pronounced in the BRAC vs. ONGO comparison in Figure 3, where only bins 1, 5–6, 9, 10, 13, 26, and 30 are excluded (also marked with blue lines and dots).

7.5 Balancing tests

Many matching papers use balancing tests as a diagnostic for whether the CIA holds. If one finds a 'treatment effect' on one or more of the conditioning variables, this suggests that one has not achieved the CIA since such a 'treatment effect' could only be caused by remaining selection. Our balancing tests involve LLR matching based on the LOR, where the outcome variable is the respective conditioning variable. This allows us to formally test the null hypothesis of no 'treatment' effect on that conditioning variable after matching. This has the advantage that we can use the matching framework to obtain standard errors that account for the fact that the propensity scores are estimated. [81]

Recall that matching is based on the assumption that one includes all conditioning variables that affect both the outcome and the propensity to be treated. Here, we argue that a balancing test would be informative for a variable not in the LOR if it potentially affects our outcome, specifically achievement Z-scores (see Austin 2009) and school assignment. In our case, this would involve conducting a balancing test for IQ, even when we only use family background variables in the LOR. We know that IQ affects achievement due to numerous previous studies.[82] Further, we know

---

[81] Our tests are in the same spirit as the third test suggested by Smith and Todd (2005), in response to Dehejia (2005).

[82] This is also evident in our data, as shown in a reduced-form regression of achievement Z-scores on family background and IQ.

from the equations used to construct the propensity score that it affects school assignment. If it fails the balancing test, it belongs in the LOR.[83]

8. Empirical Results

8.1 Estimated ATT Effects

In this section, we discuss the ATTs for switching BRAC students to GOV, JAAGO, and ONGO schools, respectively. Table 4, column (1) shows the respective differences in the math achievement Z-scores in the untrimmed sample (when we do not use any conditioning variables). Line 1 of column (2) shows the estimated ATT for switching BRAC students to GOV schools when we do not condition on IQ. To obtain this estimated ATT, we use the Epanechnikov kernel. We first estimate the log odds ratios on the untrimmed data to impose the *1:5 trimming method.* We then re-estimate the log odds ratios (i.e., re-estimate the logit model on the trimmed sample) to obtain the data-driven bandwidth, which enables us to estimate the ATT. We then repeat this process for the case where we also condition on fluid intelligence to estimate a different ATT in column (3).

Table 4, Line 1 shows that the mean value of GOV students' math achievements Z-scores minus the mean value of BRAC students' achievements Z-scores equals 0.488 (0.102) of a standard deviation when we do not use matching or trimming.[84] Column (2) shows that the estimated ATT of moving a BRAC student to a GOV school is a statistically significant 0.496 (0.104) of a standard deviation in Z-scores when we condition on only the student demographics and family variables using the trimmed sample. However, when we also control for K-BIT Z-scores using the trimmed sample in column (3), this ATT falls by over 50 percent to 0.259 (0.109).

[83] Ham et al. (2011) also conducted a balancing test for some variables not in the propensity score. They do not reject the null hypothesis that these variables are balanced.

[84] In what follows, the numbers in parentheses are standard errors.

Table 4, line 5, provides the analogous estimates when comparing students from BRAC and JAAGO schools. From column 1, the mean difference between the BRAC and JAAGO students is 0.671 (0.106) standard deviations in the math achievement Z-score. When we control only for background characteristics, the ATT of moving a BRAC student to a JAAGO school increases to 0.770 (0.112). However, once we also control for K-BIT Z-scores in the LOR, the estimated ATT falls to 0.389 (0.107) of a standard deviation, i.e., the estimated treatment effect falls by almost 50%.

Table 4, Line 9 considers the ATT for switching a BRAC student to an ONGO school. The mean of the ONGO students' achievement Z-scores is 0.286 (0.100) of a standard deviation higher than the mean in BRAC schools. The ATT of moving a BRAC student to an ONGO school when we do not control for K-BIT Z-scores is 0.253 (0.101), but this falls to a statistically insignificant 0.056 (0.104) when we control for K-BIT Z-scores.

Hence, adding the K-BIT Z-scores to the log odds ratio dramatically reduces all of the above ATTs. *Our results suggest that previous Labor and Development Economics studies, which only condition on child characteristics and family background variables to eliminate selection bias, may produce substantially biased treatment effects.*

To put the size of the ATT effects in Table 4 into perspective, we compared them to a reduced-form estimate of the impact of a one standard deviation increase in K-BIT Z-scores on math achievement. (See Appendix Table A1 for the results of regressing math scores on our conditioning variables.) The estimated effect of a full standard deviation increase in IQ on math Z-scores is 0.433 (0.025) of a standard deviation, which is only slightly larger than the impact of moving their child from a BRAC to a JAAGO school (equal to 0.389 of a standard deviation). In other words, the ATTs for BRAC vs. GOV and BRAC vs. JAAGO in Table 4 are substantial.

8.2 Estimated ATU Effects

We present the ATUs in columns (4) of Table 4 when we control for background characteristics only, whereas column (5) shows the ATUs when we also control for K-BIT. As discussed above, we can use only a limited number of observations when calculating the ATUs for BRAC vs. GOV and BRAC vs. JAAGO because of the common support issue. The estimated ATU is representative of moving *all* GOV (JAAGO) students to BRAC schools only if we assume it is constant across the LOR distribution of GOV (JAAGO) students. Thus, one would want to treat our ATU estimates with caution.

The estimated ATU of moving a GOV student to BRAC school is statistically significant at –0.389 (0.172) when we omit K-BIT in the LOR, but insignificant at –0.122 (0.167) when we include K-BIT. The estimated ATUs of switching a JAAGO student to BRAC are –0.665 (0.141) when we do not control for K-BIT and –0.370 (0.154) when we do. Thus, controlling for K-BIT has a significant impact on our estimated ATUs. Finally, the ATU for moving an ONGO student to a BRAC school when we do not control for K-BIT is –0.172 (0.116), but –0.065 (0.128) when we do; both estimates are statistically insignificant.

8.3 Estimated ATE Effects

As noted above, the estimated ATEs are weighted averages of the estimated ATTs and ATUs,

$$ATE = (\pi ATET) + ((1-\pi)ATU),$$

where π is the true proportion of BRAC students relative to BRAC and GOV students in the population (Hahn et al. 2025). We ran into trouble finding data on the relative shares of each type of school. We did find these shares for the country[85], but these differed drastically from the shares for several slums in Dhaka, as reported by other researchers.[86] Given this, we deferred estimating ATEs for now.

8.4 Robustness to Changes in the Matching Procedure

Tables 5–7 show how changing the kernels and bandwidths affects our matching estimates. Consider the estimates in Table 5, where, compared to Table 4, we change the bandwidth from a data-driven one to a fixed ex ante one. Comparing Table 6 to Table 4, we change the Epanechnikov kernel to one based on the normal distribution, while continuing to use a data-driven bandwidth. Finally, in Table 7, we investigate the impact of changing both the kernel and the bandwidth compared to those used in Table 4. None of these modifications changes our estimated treatment effects.

8.5 Balancing tests

Table 8 presents the results of our balancing tests when we use an Epanechnikov kernel. Rows 1, 2, and 3 of Panel A display our balancing test results for BRAC vs. GOV, BRAC vs. JAAGO, and BRAC vs. ONGO schools, respectively, when using a data-driven bandwidth. Columns (1) and (3) contain our ATT balancing test results for the cases where we do, and do not, include K-BIT scores in the LORs, respectively, using our base specification. Columns (2) and (4) present analogous results when we estimate the ATUs. Columns (1) and (2) indicate that we pass the balancing test for all but one of the variables using the full model (including IQ) for the three different school-type comparisons. However, when we use only the background

---

[85] See DPE (2015).

[86] See e.g., Cameron (2011).

characteristics to estimate the LOR, we find many rejections in columns (3) and (4). It is worth emphasizing that most rejections come from the K-BIT variable, and only one rejection comes from the background characteristics. Hence, our model would have looked quite good if we had only checked for balance with the family variables. This suggests that the LOR is mis-specified when it does not include K-BIT. These results are consistent with our understanding that K-BIT impacts both school-type and math achievement, and therefore should be included in the LOR.

Panel B of Table 8 presents the corresponding results when a fixed bandwidth with an Epanechnikov kernel is used. Our test results for BRAC vs. GOV and BRAC vs. ONGO are similar to those in Panel A. However, when comparing BRAC and JAAGO schools, we observe a higher rejection rate than in Panel A for the model that includes K-BIT in the LOR. Table 9 presents the results of our balancing tests when a normal kernel is used. These results are close to those in Panel A of Table 8.

8.6 Discussion of the Results

In this section, we relate our estimated treatment effects to the characteristics of the schools in Table 1. We will focus on the ATTs for two reasons. First, the ATUs for GOV and JAAGO are likely to be significantly impacted by common support issues, which led us to drop or trim many high-achieving students from these schools. Furthermore, it seems credible that the parents of GOV or JAAGO students are less likely to switch them to BRAC schools (the effect of which is measured by the ATU) than BRAC parents are to change their children to GOV or JAAGO schools (the impact of which is measured by the ATT). One reason is that the GOV and JAAGO schools are harder to enter. Furthermore, GOV schools are perceived as more prestigious than BRAC schools. Moreover, JAAGO schools were modelled after elite private schools, which could have enhanced their reputation. These factors suggest that our estimated ATUs for comparing GOV and JAAGO schools to BRAC schools are less useful for policy than our estimated ATTs.

This leads to the natural question of why GOV and JAAGO schools are successful in increasing achievement even after we control for student quality. Considering Table 1, GOV and JAAGO schools both provide formal schooling, while BRAC and some ONGO schools implement the non-formal schooling model.

Furthermore, both GOV and JAAGO schools likely have significantly better teachers for several reasons. First, as noted above, JAAGO and, in particular, the GOV teachers have significantly higher salaries than the BRAC teachers. Second, GOV and JAAGO schools have much higher teacher qualifications. Moreover, GOV schools are likely more attractive to teachers since they offer relatively generous civil service benefits. Further, teachers at GOV schools may be of higher quality than those at BRAC schools because their teachers receive one year of in-service professional training. Finally, the fact that JAAGO teachers are required to be fluent in English likely results in a positive selection process.

We can consider whether other characteristics of JAAGO, GOV, and BRAC schools can potentially explain the achievement differences. For example, GOV schools have corporal punishment, while JAAGO and BRAC schools do not. Furthermore, GOV schools have a shorter school day than BRAC schools, while JAAGO schools have a longer school day than BRAC schools. Hence, we cannot use the presence of corporal punishment or the length of the school day to explain why GOV and JAAGO schools both have higher value-added than BRAC schools. Instead, we believe that the differences in teacher quality and BRAC's reliance on an informal schooling model explain the differences in achievement at GOV and JAAGO schools compared to BRAC schools.

9. Conclusions

Over the past 30 years, BRAC has emerged as a global leader in education for developing countries, with a mission to deliver effective schooling to marginalized populations. Its model of non-formal primary education has been widely replicated by NGOs and adopted by the Government of Bangladesh. Sources, such as The New York Times and BRAC itself, claim that BRAC schools outperform government schools.[87] Others, including UNESCO, suggest BRAC students perform at par with their GOV counterparts.[88] However, our findings contradict these claims. A likely reason is that previous studies have employed estimation strategies that do not account for the use of choice-based sampling, which is prevalent in this literature.

Our key results can be summarized as follows. First, BRAC serves the most disadvantaged students, both in terms of family background and fluid intelligence. As shown in Table 3, GOV and JAAGO students have substantially higher IQ scores and more educated parents than BRAC students; the profiles of ONGO students are somewhat stronger than those of BRAC students. To our knowledge, this is the first study to rigorously document the extent of BRAC students' disadvantages relative to students at other school types, especially in terms of fluid intelligence.

Second, accurate estimation of BRAC's impact requires careful control for selection. This is unsurprising given the severe disadvantages faced by BRAC students. Conditioning only on standard demographic and family background variables does little to adjust for selection bias, as it produces impacts that are similar to the simple mean differences in achievement. In contrast, controlling for both fluid intelligence and background characteristics substantially reduces the estimated treatment effects. Nonetheless, even after accounting for these factors, BRAC students

[87] See Rosenberg (2013) and the 'Our Work' section of BRAC's official website (BRAC International: Education 2025).

[88] See UNESCO Institute of Lifelong Learning (2016).

continue to perform significantly worse than their counterparts in GOV and JAAGO schools. Notably, once IQ is taken into account, BRAC students show no significant difference in outcomes compared to those in ONGO schools.

Although BRAC schools possess several potentially positive features, particularly when compared to GOV schools, many of these characteristics are also found in JAAGO schools. We argue that differences in educational models (i.e., formal versus non-formal structures) and disparities in teacher quality primarily drive the observed differences in achievement. GOV and JAAGO schools impose higher teacher qualification standards and offer more competitive salaries, enabling them to recruit more capable teachers. BRAC, by contrast, restricts its teacher recruitment to the local community, pays lower wages, and imposes lower minimum qualifications, thereby reducing staff quality. While BRAC is often praised for its low per-student cost and for having learning outcomes at least as good as those of GOV schools, our findings suggest that this narrative is far too optimistic.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. ‘When Should You Adjust Standard Errors for Clustering?’ *The Quarterly Journal of Economics* 138 (1): 1–35.

Ahmad, Alia, and Iftekharul Haque. 2011. *Economic and Social Analysis of Primary Education in Bangladesh: ·A Study of BRAC Interventions and Mainstream Schools*. Research Monograph Series No. 48. Dhaka: Research and Evaluation Division, BRAC, 2011.

Ahmed, Manzoon, Samir Ranjan Nath, and Kazi Saleh Ahmed. 2002. *Literacy in Bangladesh: Need for a New Vision*. Education Watch 2002. Dhaka: Campaign for Popular Education (CAMPE), 2002. https://www.campebd.org/page/Generic/0/6/18.

Alamgir, Mohiuddin. 2024. ‘Separate Pay Scale for Teachers Still Not in Sight’. *The Daily Star*, 5 October 2024. https://www.thedailystar.net/news/bangladesh/education/news/separate-pay-scale-teachers-still-not-sight-3719826.

Arenas, Diego, Miranda Bodi-Torralba, Andrea Oliver, Jaume Cantallops, Francisco J. Ponseti, Pere Palou-Sampol, Juan A. Collado, et al. 2024. ‘Effects of Active Breaks on Educational Achievement in Children with and without ADHD: Study Protocol and Rationale of the Break4Brain Project’. *Frontiers in Psychology* 15 (October): 1451731.

Austin, Peter C. 2009. ‘Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-score Matched Samples’. *Statistics in Medicine* 28 (25): 3083–107.

Banerjee, Abhijit, and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. 1st ed. New York: PublicAffairs, 2011.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. ‘A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries’. *Science* 348 (6236): 1260799.

Banerjee, Abhijit, Esther Duflo, and Garima Sharma. 2021. ‘Long-Term Effects of the Targeting the Ultra Poor Program’. *American Economic Review: Insights* 3 (4): 471–86.

Bangladesh Bureau of Statistics. 2017. *Education Scenario in Bangladesh: Gender Perspective*. Bangladesh Bureau of Statistics, 2017.

Barac, Raluca, and Ellen Bialystok. 2012. ‘Bilingual Effects on Cognitive and Linguistic Development: Role of Language, Cultural Background, and Education’. *Child Development* 83 (2): 413–22.

Basic Education and Policy Support (BEPS) Activity. 2002. *Teachers and Teacher Training (Formal and Nonformal)*. No. 4. Bangladesh Education Sector Review. Washington, D.C.: United States Agency for International Development, 2022.

brac. 2020. ‘BRAC Ranked #1 NGO in the World for the Fifth Consecutive Year’. Brac, 5 March 2020. https://www.brac.net/latest-news/item/1269-brac-ranked-1-ngo-in-the-world-for-the-fifth-consecutive-year?fbclid=IwAR1XjavY5tOKnUfSRJ7NZKPiEko3gkXkv1lxLtrIdiMZ_XErcuwWg8uQrSg.

———. 2025a. ‘About Us – BRAC International’. 2025. https://bracinternational.org/aboutus/.

———. 2025b. ‘Social Development’. What We Do, 2025. https://www.brac.net/what-we-do.

BRAC International: Education. 2025. ‘Our Work’. BRAC International: Education, 2025. https://bracinternational.org/education/.

Burchinal, Margaret, Lynne Vernon-Feagans, Virginia Vitiello, and Mark Greenberg. 2014. ‘Thresholds in the Association between Child Care Quality and Child Outcomes in Rural Preschool Children’. *Early Childhood Research Quarterly* 29 (1): 41–51.

Cameron, Claire E., Laura L. Brock, William M. Murrah, Lindsay H. Bell, Samantha L. Worzalla, David Grissmer, and Frederick J. Morrison. 2012. ‘Fine Motor Skills and Executive Function Both Contribute to Kindergarten Achievement’. *Child Development* 83 (4): 1229–44.

Cameron, Stuart. 2011. ‘Whether and Where to Enrol? Choosing a Primary School in the Slums of Urban Dhaka, Bangladesh’. *International Journal of Educational Development* 31: 357–66.

Campaign for Popular Education (CAMPE). 2000. *A Question of Quality: State of Primary Education in Bangladesh. Volume 1: Major Findings – A Synthesis*. Education Watch 2000. Dhaka: Campaign for Popular Education (CAMPE), 2000. https://www.campebd.org/page/Generic/0/6/18.

———. 2015. *Moving from MDG to SDG Accelerate Progress for Quality Primary Education*. Education Watch 2015. Dhaka: Campaign for Popular Education (CAMPE), 2015. https://www.campebd.org/page/Generic/0/6/18.

———. n.d. ‘About CAMPE’. Campaign for Popular Education (CAMPE). Accessed 4 June 2025. https://www.campebd.org/page/Generic/0/3/1.

Chabbott, Colette. 2006. *Meeting EFA: Bangladesh Rural Advancement Committee (BRAC) Primary Schools*. EQUIP2. United States Agency for International Development, 2006. https://www.epdc.org/sites/default/files/documents/BRAC%20Primary%20Schools.pdf.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and Halsey Rogers. 2004. *Roll Call: Teacher Absence in Bangladesh*. World Bank, 2004. https://documents1.worldbank.org/curated/en/443701468014360854/pdf/470430WP0Box331her1Absence01PUBLIC1.pdf.

Cheng, Alison S.M., Jolene Q.J. To, Noor Hafizah Wahianuar, Yiong Huak Chan, and Kalyani Vijaykumar Mulay. 2024. ‘Concurrent Validity of Intelligence Assessments in Children with Developmental Disabilities in an Asian Setting: Comparison of the Kaufman Brief Intelligence Test – Second Edition with the Wechsler Intelligence Scales'. *Pediatrics & Neonatology* 65 (4): 341–47.

Crosnoe, Robert, Tama Leventhal, R. J. Wirth, Kim M. Pierce, Robert C. Pianta, and NICHD Early Child Care Research Network. 2010. 'Family Socioeconomic Status and Consistent Environmental Stimulation in Early Childhood'. *Child Development* 81 (3): 972–87.

Davis-Kean, Pamela E. 2005. 'The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment.' *Journal of Family Psychology* 19 (2): 294–304.

Dean, Raymon S. 2011. 'Woodcock-Johnson III Tests of Achievement'. In *Encyclopedia of Child Behavior and Development*. Boston, MA: Springer, 2011.

Dehejia, Rajeev. 2005. 'Practical Propensity Score Matching: A Reply to Smith and Todd'. *Journal of Econometrics* 125: 355–64.

Denton, Carolyn A., Tammy D. Tolar, Jack M. Fletcher, Amy E. Barth, Sharon Vaughn, and David J. Francis. 2013. 'Effects of Tier 3 Intervention for Students with Persistent Reading Difficulties and Characteristics of Inadequate Responders.' *Journal of Educational Psychology* 105 (3): 633–48.

Directorate of Primary Education (DPE). 2013. *Primary Assistant Teacher Circular: Apply Online*. Dhaka: Directorate of Primary Education (DPE), 2013.

———. 2014a. *Annual Primary School Census 2014*. Monitoring and Evaluation Division, Directorate of Primary Education, Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh, 2014. http://www.dpe.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/35332e71_7177_43e4_acad_69855daff795/APSC-2014_for%20National%20Committee_11032015.pdf.

———. 2014b. *Primary Assistant Teacher Circular: Apply Online*. Dhaka: Directorate of Primary Education (DPE), 2014.

———. 2015. *Annual Primary School Census - 2015*. Monitoring and Evaluation Division, Directorate of Primary Education, Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh', 2015. http://www.dpe.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/f2501e16_4f57_467b_8c67_48a5855f88fe/Final%20Draft%20APSC2015.pdf.

———. 2016. *Bangladesh Primary Education Annual Sector Performance Report - 2016*. Monitoring and Evaluation Division, Directorate of Primary Education, Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh, 2016. http://www.dpe.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/7402e8d5_9a8b_43c1_9d4b_2c785c18c066/Final%20%20ASPR%202016,%2028%20September%202016.pdf.

———. 2017. *Bangladesh Primary Education Annual Sector Performance Report - 2017*. Monitoring and Evaluation Division, Directorate of Primary Education, Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh, 2017. http://www.dpe.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/86f777e9_ecf0_43e2_a582_d3df9ebaa0bd/ASPR%202017%20Report_20-11-2017%20(1).pdf.

———. 2018a. *Annual Primary School Census 2018*. Monitoring and Evaluation Division, Directorate of Primary Education, Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh, 2018. http://www.dpe.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/97463005_c127_444f_979b_721b315e71d3/Final%20Draft%20Annual%20Primary%20School%20Census-2018.pdf.

———. 2018b. *Training Calendar 2017 - 2018*. Training Division Directorate of Primary Education (DPE), Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh, 2018.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, et al. 2007. 'School Readiness and Later Achievement'. *Developmental Psychology* 43 (6): 1428–46.

El Nokali, Nermeen E., Heather J. Bachman, and Elizabeth Votruba-Drzal. 2010. 'Parent Involvement and Children's Academic and Social Development in Elementary School'. *Child Development* 81 (3): 988–1005.

Espy, Kimberly Andrews, Melanie M. McDiarmid, Mary F. Cwik, Melissa Meade Stalets, Arlena Hamby, and Theresa E. Senn. 2004. 'The Contribution of Executive Functions to Emergent Mathematic Skills in Preschool Children'. *Developmental Neuropsychology* 26 (1): 465–86.

Fan, Jianqing. 2018. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability 66*. New York: Routledge, 2018.

Fan, Jianqing, and Irene Gijbels. 1996. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability 66*. 1st Edition. Monographs on Statistics and Applied Probability 66. New York: Chapman & Hall/CRC, 1996. https://www.routledge.com/Local-Polynomial-Modelling-and-Its-Applications-Monographs-on-Statistics-and-Applied-Probability-66/Fan-Gijbels/p/book/9780412983214.

Ferrer, Emilio, and John J. McArdle. 2004. 'An Experimental Analysis of Dynamic Hypotheses About Cognitive Abilities and Achievement From Childhood to Early Adulthood.' *Developmental Psychology* 40 (6): 935–52.

Finance Division, Ministry of Finance. 2015. *National Payscale 2015*. S.R.O No. 369-Law/2015. Dhaka: Government of the People's Republic of Bangladesh, 2015. https://mof.portal.gov.bd/sites/default/files/files/mof.portal.gov.bd/page/ebb460ed_b1f5_4fa1_bd1b_e08f87400002/Civil.pdf.

Fitzpatrick, Caroline, Rachel D. McKinnon, Clancy B. Blair, and Michael T. Willoughby. 2014. 'Do Preschool Executive Function Skills Explain the School Readiness Gap between Advantaged and Disadvantaged Children?' *Learning and Instruction* 30 (April): 25–31.

Floyd, Randy G., Renee Bergeron, and Vincent C. Alfonso. 2006. 'Cattell–Horn–Carroll Cognitive Ability Profiles of Poor Comprehenders'. *Reading and Writing* 19 (5): 427–56.

Font, Sarah A. 2014. 'Kinship and Nonrelative Foster Care: The Effect of Placement Type on Child Well-Being'. *Child Development* 85 (5): 2074–90.

French, Rob, and Geeta Kingdon. 2010. *The Relative Effectiveness of Private and Government Schools in Rural India: Evidence from ASER Data*. Nos 10–03. DoQSS Working Papers. Quantitative Social Science - UCL Social Research Institute, University College London, 2010.

Gormley, William T., Ted Gayer, Deborah Phillips, and Brittany Dawson. 2005. 'The Effects of Universal Pre-K on Cognitive Development.' *Developmental Psychology* 41 (6): 872–84.

Government of Bangladesh. 1990. *Primary Education (Compulsory) Act 1990*. No. 5. Bangladesh Gazette. Dhaka: Government of Bangladesh, 1990.

Goyal, Sangeeta. 2009. 'Inside the House of Learning: The Relative Performance of Public and Private Schools in Orissa'. *Education Economics* 17 (3): 315–27.

Goyal, Sangeeta, and Priyanka Pandey. 2009. 'How Do Government and Private Schools Differ? Findings from Two Large Indian States'. *South Asia Human Development Sector Discussion Paper Series - No. 30*, December 2009.

Hahn, Jinyong, John Ham, Geert Ridder, and Shuyang Sheng. 2025. 'Stratifying on Treatment Status'. *Economics Letters* 253 (June): 112370.

Ham, John C., Xianghong Li, and Patricia B. Reagan. 2011. 'Matching and Semi-Parametric IV Estimation, a Distance-Based Measure of Migration, and the Wages of Young Men'. *Journal of Econometrics* 161 (2): 208–27.

Hastings, Paul D., Sarah Kahle, and Jacob M. Nuselovici. 2014. 'How Well Socially Wary Preschoolers Fare Over Time Depends on Their Parasympathetic Regulation and Socialization'. *Child Development* 85 (4): 1586–600.

Heckman, J. J., H. Ichimura, and P. E. Todd. 1997. 'Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme'. *The Review of Economic Studies* 64 (4): 605–54.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. 'Matching As An Econometric Evaluation Estimator'. *Review of Economic Studies* 65 (2): 261–94.

Heckman, James J., and Petra E. Todd. 2009. 'A Note on Adapting Propensity Score Matching and Selection Models to Choice Based Samples'. *The Econometrics Journal* 12 (S1,): S230–34.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. ‘Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score’. *Econometrica* 71 (4): 1161–89.

Hsieh, Chang-Tai, and Miguel Urquiola. 2006. ‘The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile’s Voucher Program’. *Journal of Public Economics* 90 (8–9): 1477–503.

Kaufman, Alan S., and Nadeen L. Kaufman. 2004. *Kaufman Brief Intelligence Test, Second Edition: Manual*. 2nd edn. Bloomington, MN: PsychCorp, Pearson, 2004.

Kawser, Rumi. 2019. ‘Graduate Degree Mandatory for Aspiring Female Government Teachers’. *Dhaka Tribune*, 10 April 2019. https://www.dhakatribune.com/bangladesh/nation/173862/graduate-degree-mandatory-for-aspiring-female.

Khan, Ashrafuzzaman, and Mrinmoy Samadder. 2010. *Beyond Dropout: A Study on BRAC Primary School*. No. 16. BRAC Working Paper. Dhaka: Research and Evaluation Division, BRAC, 2010. https://bigd.bracu.ac.bd/wp-content/uploads/2020/03/Beyond-Dropout-A-Study-on-BRAC-Primary-School_red_wp16.pdf.

Kielland, Anne. 2015. *Evolution in Approaches to Improve Access to Education for Children Living in Urban Slums*. Paper commissioned for the EFA Global Monitoring Report 2015, Education for All 2000–2015: Achievements and Challenges. Paris: UNESCO, 2015.

Lechner, Michael. 2000. ‘An Evaluation of Public-Sector-Sponsored Continuous Vocational Training Programs in East Germany’. *Journal of Human Resources* 35 (2): 347–75.

Leuven, Edwin, and Barbara Sianesi. 2003. ‘PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing’. 2003. http://ideas.repec.org/c/boc/bocode/s432001.html.

Lucas, Adrienne M., and Isaac M. Mbiti. 2014. ‘Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya’. *American Economic Journal: Applied Economics* 6 (3): 234–63.

Mather, Nancy, and Richard W. Woodcock. 2001. *Examiner’s Manual: Woodcock-Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside Publishing, 2001.

Ministry of Foreign Affairs of the Netherlands. 2011. *IOB Evaluation: The Two-Pronged Approach: Evaluation of Netherlands Support to Primary Education in Bangladesh*. 2011.

Nath, Samir R, Mirja M. Shahjamal, Rosie Nilufar Yasmin, Tata Zafar, and Mahbubul Kabir. 2005. *Achievement of Competencies of the BRAC School Students*. BRAC Research Report. Dhaka: Research and Evaluation Division, BRAC, 2005.

Nath, Samir R., Kathy Sylva, and Janice Grimes. 1999. 'Raising Basic Education Levels in Rural Bangladesh: The Impact of a Non-Formal Education Programme'. *International Review of Education* 45 (1): 5–26.

Nath, Samir Ranjan. 2008. 'Private Supplementary Tutoring among Primary Students in Bangladesh'. *Educational Studies* 34 (1): 55–72.

———. 2012. 'Factors Influencing Primary Students' Learning Achievement in Bangladesh'. *Research in Education* 88 (1): 50–63.

Nath, Samir Ranjan, and A Mushtaque R. Chowdhury. 2008. *State of Primary Education in Bangladesh Progress Made, Challenges Remained*. Education Watch 2008. Dhaka: Campaign for Popular Education (CAMPE), 2008. https://www.campebd.org/page/Generic/0/6/18.

———. 2016. *Literacy, Skills, Lifelong Learning: SDG4 in Bangladesh: Where Are We*. Education Watch 2016. Dhaka: Campaign for Popular Education (CAMPE), 2016. https://www.campebd.org/page/Generic/0/6/18.

Nath, Samir Ranjan, and Mohammad Anwar Hossain. 2017. *Literacy Status of BRAC Primary School Completers*. Dhaka: Research and Evaluation Division, BRAC, 2017. https://bigd.bracu.ac.bd/wp-content/uploads/2021/11/Literacy_Status_BRAC_Primary_School-completers.pdf.

Nath, Samir Ranjan, Mohammad Anwar Hossain, and Tanjeeba Chowdhury. 2018. *Slum Pre-Primary Graduates in Primary Education: A Tracer Study on BEHTRUWC-Sibling Project*. Dhaka: Research and Evaluation Division, BRAC, 2018. https://bigd.bracu.ac.bd/wp-content/uploads/2021/11/Slum-Pre-primary-Graduates-in-Primary-Education-A-Tracer-Study-on-BEHTRUWC-Sibling-Project.pdf.

Nath, Samir Ranjan, Goutam Roy, and Anwar Hossain. 2007. *Achievement of Primary Competencies: A Comparison between Government and BRAC Schools*. BRAC Research Report. 2007. http://www.ssrn.com/abstract=2919748.

National Academy for Primary Education (NAPE). 2020. *Annual Report 2019–2020*. Mymensingh: National Academy for Primary Education, 2020. http://www.nape.gov.bd/sites/default/files/files/nape.portal.gov.bd/annual_reports/bc144930_27fe_46e7_94eb_67a634453563/2020-09-16-20-21-f68c114f50461f5e5d92b3f404bf8e07.pdf.

Newhouse, David, and Kathleen Beegle. 2006. 'The Effect of School Type on Academic Achievement: Evidence from Indonesia'. *Journal of Human Resources* 41 (3): 529–57.

Oxford Policy Management (OPM). 2006. *Social Sector Performance Review – Primary Education*. Dhaka: Oxford Policy Management and Financial Management Reform Program (FMRP), 2006.

Pal, Sarmistha, and Bibhas Saha. 2014. 'In "Trusts" We Trust: Socially Motivated Private Schools in Nepal'. *IZA Discussion Paper No. 8270*, June 2014.

Pitts, C. Holley, and Carolyn B. Mervis. 2016. 'Performance on the Kaufman Brief Intelligence Test-2 by Children With Williams Syndrome'. *American Journal on Intellectual and Developmental Disabilities* 121 (1): 33–47.

Rodriguez-Martínez, Elena I., Raquel Muñoz-Pradas, Antonio Arjona, Brenda Y. Angulo-Ruiz, Vanesa Muñoz, and Carlos M. Gómez. 2023. 'Neuropsychological Assessment of the Relationship of Working Memory with K-BIT Matrices and Vocabulary in Normal Development and ADHD Children and Adolescents'. *Brain Sciences* 13 (11): 1538.

Romero, Mauricio, Justin Sandefur, and Wayne Sandholtz. 2020. 'Outsourcing Education: Experimental Evidence from Liberia'. *American Economic Review* 110(2): 364–400.

———. 2022. 'Beyond Short-Term Learning Gains: The Impact of Outsourcing Schools in Liberia After Three Years'. *The Economic Journal* 132 (644)*:* 1600–1619.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. 'The Central Role of the Propensity Score in Observational Studies for Causal Effects'. *Biometrika* 70 (1): 41–55.

Rosenberg, Tina. 2013. 'Where Private School Is Not a Privilege'. Opinion. *Opinionator*, 8 May 2013. https://archive.nytimes.com/opinionator.blogs.nytimes.com/2013/05/08/where-private-school-is-not-a-privilege/.

Shahjamal, Mirja Mohammad, and Samir Ranjan Nath. 2008. *An Evaluation of BRAC Pre-Primary Education Programme*. BRAC Research Report. Research and Evaluation Division, BRAC., 2008.

Smith, Jeffrey A., and Petra E. Todd. 2005. 'Rejoinder'. *Journal of Econometrics* 125 (1–2): 365–75.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. 'What Are We Weighting For?' *The Journal of Human Resources* 50 (2): 301–16.

The World Bank. n.d. 'Quasi-Experimental Methods'. Dimewiki. Accessed 4 June 2025. https://dimewiki.worldbank.org/Quasi-Experimental_Methods.

Tietjen, Karen, Arifur Rahman, and Seth Spaulding. 2004. *Teachers' and Students' Use of Time in Government Primary Schools in Bangladesh*. Basic Education and Policy Support (BEPS) Activity No. 4. Bangladesh Educational Assessment. Washington, D.C.: United States Agency for International Development, 2004.

UNESCO Institute of Life Long Learning. 2016. 'BRAC Education Programme, Bangladesh'. BRAC Education Programme, Bangladesh - Effective Literacy and Numeracy Practices Database – LitBase, 31 August 2016. https://www.uil.unesco.org/en/litbase/brac-education-programme-bangladesh.

‘US Dollar to Bangladesh Taka History: 2016’. n.d. Accessed 22 June 2025. https://www.exchangerates.org.uk/USD-BDT-spot-exchange-rates-history-2016.html.

Waber, Deborah P., Michael D. Weiler, Peter H. Wolff, David Bellinger, David J. Marcus, Raya Ariel, Peter Forbes, and David Wypij. 2001. ‘Processing of Rapid Auditory Stimuli in School-Age Children Referred for Evaluation of Learning Disorders’. *Child Development* 72 (1): 37–49.

Woodcock, Richard W., Kevin S. McGrew, and Nancy Mather. 2001. *Woodcock Johnson III Test of Achievement Standard Test Book (Form A)*. Itasca, IL: Riverside Publishing, 2001.

Wu, Shunquan, Yingying Ding, Fuquan Wu, Jun Hou, and Panyong Mao. 2015. ‘Application of Propensity-Score Matching in Four Leading Medical Journals’. *Epidemiology (Cambridge, Mass.)* 26 (2): e19-20.

Zhang, Zhiwei, Zonghui Hu, and Chunling Liu. 2019. ‘Estimating the Population Average Treatment Effect in Observational Studies with Choice-Based Sampling’. *The International Journal of Biostatistics* 15 (1).

Table 1: A Comparison of School Types

Characteristics	BRAC	GOV	JAAGO	Other NGOs
Informal Schooling	Yes	No	No	Mixed
Instruction in English	No	No	Yes	No
Minimum teacher qualification	10 years of schooling	HSC[a] (Female teachers) and Bachelor's degree (Male teachers) [b]	Bachelor's degree	HSC
Teachers required to pass Civil Service Exams	No	Yes	No	No
Teachers require English Proficiency	No	No	Yes	No
Pre-service training	12–15 days	0 days	30 days	4 days
In-service training/ Refresher Courses	30–35 days annually	12-18 months training (C-in-Ed/B-in-Ed)	need-based workshops	6 days
Proportion of female teachers	98–99%	60–66%	80%	85%
Monthly teacher salary	BDT 1,250 – BDT 1,650	BDT 11,000–26,590	BDT 4,000–5,000	NA
Class size	33	51	31	37
Contact hours per day	3.6	2.9	4.5	3.85
School year (days per annum)	265–270	228	265	247
Corporal punishment	No	Yes	No	No

Note: (a) Completing the Higher Secondary Certificate (HSC) means that 12 years of schooling has been completed. (b) Before 2013, the minimum teacher qualification was SSC (completing 10 years of schooling) for female teachers and HSC (completing 12 years of schooling) for male teachers.

Table 2: Means by School Type

	BRAC	GOV	JAAGO	Other NGOs
Achievement Test Z-Score	-0.4320	0.0556	0.2387	-0.1459
	(0.0977)	(0.0439)	(0.0307)	(0.0875)
Monthly Family Expenditure	5.1819	6.1451	5.8223	5.3928
(in BDT 1000 adjusted by equivalence scale)	(0.1278)	(0.1178)	(0.1334)	(0.1365)
Father's Schooling	2.4956	3.6785	3.6216	3.2606
	(0.1950)	(0.2312)	(0.1906)	(0.3121)
Mother's Schooling	2.1010	3.2451	3.7765	2.8423
	(0.1260)	(0.2341)	(0.1858)	(0.2337)
K-BIT (IQ)	-0.4918	0.0628	0.2773	-0.1729
	(0.0607)	(0.0695)	(0.0708)	(0.0842)
Observations	264	618	607	447

Notes: (a) Standard errors in parentheses, for the mean values, are clustered at the street level; (b) For the IQ score, we use age-adjusted Z-scores, i.e., for student i in age group a, we calculate, $Z_i = \frac{X_i - X_a}{\sigma_a}$, where X_a and σ_a are the mean and standard deviation in age group a, respectively.

Table 3: Mean Differences Across School Type

	(1) BRAC vs GOV	(2) BRAC vs JAAGO	(3) BRAC vs Other NGOs
Achievement Test Z-Score	-0.4876***	-0.6708***	-0.2861***
	(0.1024)	(0.1057)	(0.0998)
Monthly Family Expenditure	-0.9632***	-0.6404***	-0.2109
(in BDT 1000 adjusted by equivalence scale)	(0.1851)	(0.1720)	(0.1928)
Father's Schooling	-1.1830***	-1.1261***	-0.7650***
	(0.3082)	(0.3347)	(0.2864)
Mother's Schooling	-1.1441***	-1.6755***	-0.7413***
	(0.2645)	(0.2336)	(0.2292)
K-BIT (IQ)	-0.5546***	-0.7691***	-0.3190***
	(0.0820)	(0.1064)	(0.0992)

Notes: (a) Standard errors in parentheses, for the differences in means, are clustered at the street level; (b) We report the difference in means at the 1 percent, 5 percent, and 10 percent significance level denoted by *** , ** , and *, respectively; (c) For KBIT (IQ score), we report age-adjusted Z-scores.

Table 4: Estimating ATET and ATU Using Matching to Control for Selection (1:5 Trim, Data-Driven Bandwidth Using Epanechnikov Kernel)

	Dependent Variable: Achievement Test Z-Score				
		ATET		ATU	
	(1)	(2)	(3)	(4)	(5)
	Mean Difference (no controls)	**Only Family Background**	**Family Background & K-BIT**	**Only Family Background**	**Family Background & K-BIT**
BRAC vs GOV	-0.4876***	-0.4956***	-0.2590**	-0.3836**	-0.1220
	(0.1024)	(0.1039)	(0.1091)	(0.1721)	(0.1696)
p-value		0.0000	0.0176	0.0259	0.4721
bandwidth		0.16	0.23	0.16	0.23
BRAC vs JAAGO	-0.6708***	-0.7703***	-0.3888***	-0.6646***	-0.3697**
	(0.1057)	(0.1116)	(0.1065)	(0.1413)	(0.1539)
p-value		0.0000	0.0003	0.0000	0.0163
bandwidth		0.21	0.21	0.21	0.21
BRAC vs Other NGOs	-0.2861***	-0.2525**	-0.0562	-0.1717	-0.0653
	(0.0998)	(0.1009)	(0.1039)	(0.1160)	(0.1275)
p-value		0.0123	0.5887	0.1388	0.6086
bandwidth		0.13	0.12	0.13	0.12

Notes: (a) Bootstrapped standard errors in parentheses are clustered at the street level; (b) We report Treatment Effects at the 1 percent, 5 percent, and 10 percent significance level denoted by *** , ** , and *, respectively; (c) Family background matching covariates consist of child's age, gender, family size, father absence dummy, father's schooling, and mother's schooling; (d) Pre-trim sample sizes are given in Table 2.

Table 5: Estimating ATET and ATU using Matching to Control for Selection (1:5 Trim, Ex-Ante Fixed Bandwidth Using Epanechnikov Kernel)

	Dependent Variable: Achievement Test Z-Score				
		ATET		ATU	
	(1)	(2)	(3)	(4)	(5)
	Mean Difference (no controls)	**Only Family Background**	**Family Background & K-BIT**	**Only Family Background**	**Family Background & K-BIT**
BRAC vs GOV	-0.4876***	-0.5015***	-0.2566**	-0.3657**	-0.1330
	(0.1024)	(0.1026)	(0.1080)	(0.1729)	(0.1694)
p-value		0.0000	0.0175	0.0345	0.4323
bandwidth		0.23	0.27	0.23	0.27
BRAC vs JAAGO	-0.6708***	-0.7696***	-0.3915***	-0.6608***	-0.3716**
	(0.1057)	(0.1113)	(0.1036)	(0.1425)	(0.1475)
p-value		0.0000	0.0002	0.0000	0.0117
bandwidth		0.18	0.36	0.18	0.36
BRAC vs Other NGOs	-0.2861***	-0.2495**	-0.0565	-0.1690	-0.0554
	(0.0998)	(0.1007)	(0.1033)	(0.1165)	(0.1247)
p-value		0.0132	0.5846	0.1468	0.6569
bandwidth		0.11	0.16	0.11	0.16

Notes: (a) Ex-ante bandwidth taken from *psmatch2* adjusted for choice-based sampling; (b) See notes to Table 4.

Table 6: Estimating ATET and ATU using Matching to Control for Selection (1:5 Trim, Data-Driven Bandwidth Using Normal Kernel)

	Dependent Variable: Achievement Test Z-Score				
		ATET		ATU	
	(1)	(2)	(3)	(4)	(5)
	Mean Difference (no controls)	**Only Family Background**	**Family Background & K-BIT**	**Only Family Background**	**Family Background & K-BIT**
BRAC vs GOV	-0.4876***	-0.4953***	-0.2733***	-0.3479**	-0.1390
	(0.1024)	(0.0982)	(0.0904)	(0.1700)	(0.1666)
p-value		0.0000	0.0025	0.0407	0.4042
bandwidth		0.16	0.23	0.16	0.23
BRAC vs JAAGO	-0.6708***	-0.7428***	-0.4011***	-0.663***	-0.3372**
	(0.1057)	(0.1065)	(0.0989)	(0.1326)	(0.1456)
p-value		0.0000	0.0000	0.0000	0.0206
bandwidth		0.21	0.21	0.21	0.21
BRAC vs Other NGOs	-0.2861***	-0.2554**	-0.0852	-0.1845*	-0.0521
	(0.0998)	(0.1035)	(0.1006)	(0.1072)	(0.1217)
p-value		0.0136	0.3969	0.0853	0.6687
bandwidth		0.13	0.12	0.13	0.12

Notes: See notes to Table 4.

Table 7: Estimating ATET and ATU using Matching to Control for Selection (1:5 Trim, Ex-Ante Fixed Bandwidth Using Normal Kernel)

	Dependent Variable: Achievement Test Z-Score				
		ATET		ATU	
	(1)	(2)	(3)	(4)	(5)
	Mean Difference (no controls)	**Only Family Background**	**Family Background & K-BIT**	**Only Family Background**	**Family Background & K-BIT**
BRAC vs GOV	-0.4876***	-0.4943***	-0.2707***	-0.3540**	-0.1350
	(0.1024)	(0.0989)	(0.0888)	(0.1691)	(0.1656)
p-value		0.0000	0.0023	0.0364	0.4150
bandwidth		0.23	0.27	0.23	0.27
BRAC vs JAAGO	-0.6708***	-0.7455***	-0.4240***	-0.6653***	-0.3317**
	(0.1057)	(0.1064)	(0.0972)	(0.1335)	(0.1403)
p-value		0.0000	0.0000	0.0000	0.0181
bandwidth		0.18	0.36	0.18	0.36
BRAC vs Other NGOs	-0.2861***	-0.2576**	-0.0828	-0.1823*	-0.0560
	(0.0998)	(0.1034)	(0.1017)	(0.1073)	(0.1225)
p-value		0.0127	0.4153	0.0892	0.6478
bandwidth		0.11	0.16	0.11	0.16

Notes: (a) Ex-ante bandwidth taken from *psmatch2* adjusted for choice-based sampling; (b) See notes to Table 4.

Table 8: Balancing Tests at the 5% Level for Family Characteristics and IQ (Using Epanechnikov Kernel)

	(A) Matching Using 1:5 Trim and Data-Driven Bandwidth				(B) Matching Using 1:5 Trim and Ex-Ante Fixed Bandwidth			
	LOR estimated using correctly-specified model (including K-BIT in the propensity score)		LOR estimated using misspecified model (excluding K-BIT from the propensity score)		LOR estimated using correctly-specified model (including K-BIT in the propensity score)		LOR estimated using misspecified model (excluding K-BIT from the propensity score)	
	(1) **ATET**	(2) **ATU**	(3) **ATET**	(4) **ATU**	(5) **ATET**	(6) **ATU**	(7) **ATET**	(8) **ATU**
BRAC vs GOV	0	0	1	1	0	0	1	1
BRAC vs JAAGO	1	0	2	1	2	1	1	1
BRAC vs Other NGOs	0	0	1	1	0	0	1	1

Notes: (a) The entry in each cell denotes the number of unbalanced conditioning variables (out of seven) at the 5% level.; (b) Log Odds Ratio (LOR) estimated using correctly-specified model employs all seven matching covariates: child's age, gender, family size, father absence dummy, father's schooling, mother's schooling, and K-BIT Z-scores. (c) Log Odds Ratio (LOR) estimated using the misspecified model employs six conditioning variables, i.e., it excludes K-BIT Z-scores.

Table 9: Balancing Tests at the 5% Level for Family Characteristics and IQ (Using Normal Kernel)

	(A) Matching Using 1:5 Trim and Data-Driven Bandwidth				(B) Matching Using 1:5 Trim and Ex-Ante Fixed Bandwidth			
	LOR estimated using correctly-specified model (including K-BIT in the propensity score)		LOR estimated using misspecified model (excluding K-BIT from the propensity score)		LOR estimated using correctly-specified model (including K-BIT in the propensity score)		LOR estimated using misspecified model (excluding K-BIT from the propensity score)	
	(1) **ATET**	(2) **ATU**	(3) **ATET**	(4) **ATU**	(5) **ATET**	(6) **ATU**	(7) **ATET**	(8) **ATU**
BRAC vs GOV	0	0	1	1	0	0	1	1
BRAC vs JAAGO	1	1	1	2	0	0	1	2
BRAC vs Other NGOs	0	0	1	1	0	0	1	1

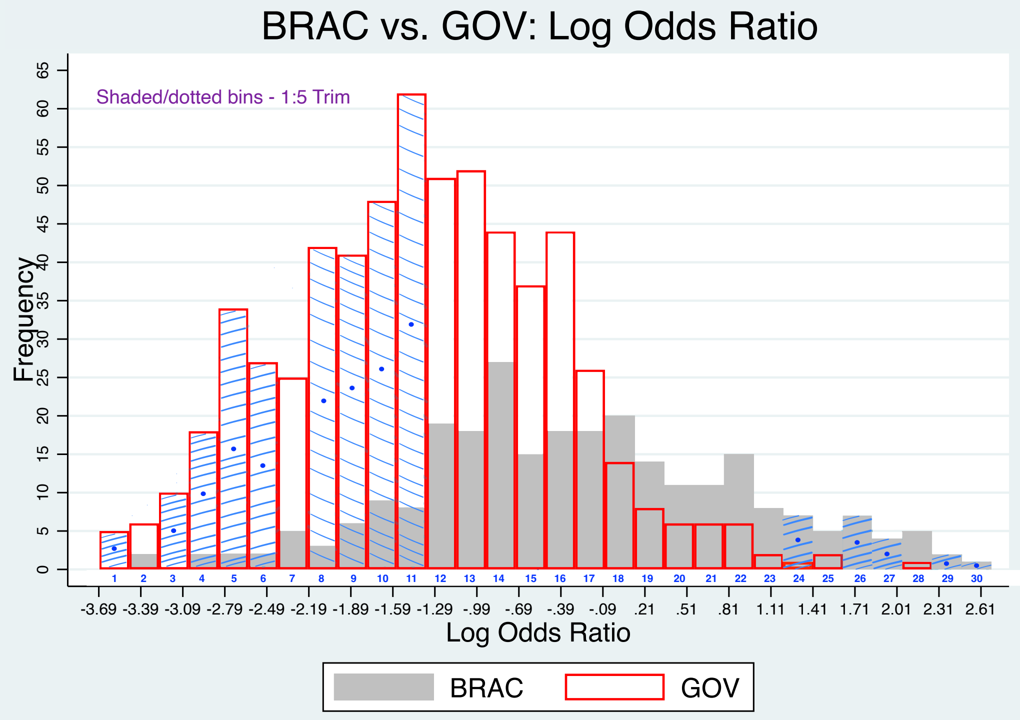
Notes: See notes in Table 8.

Appendix

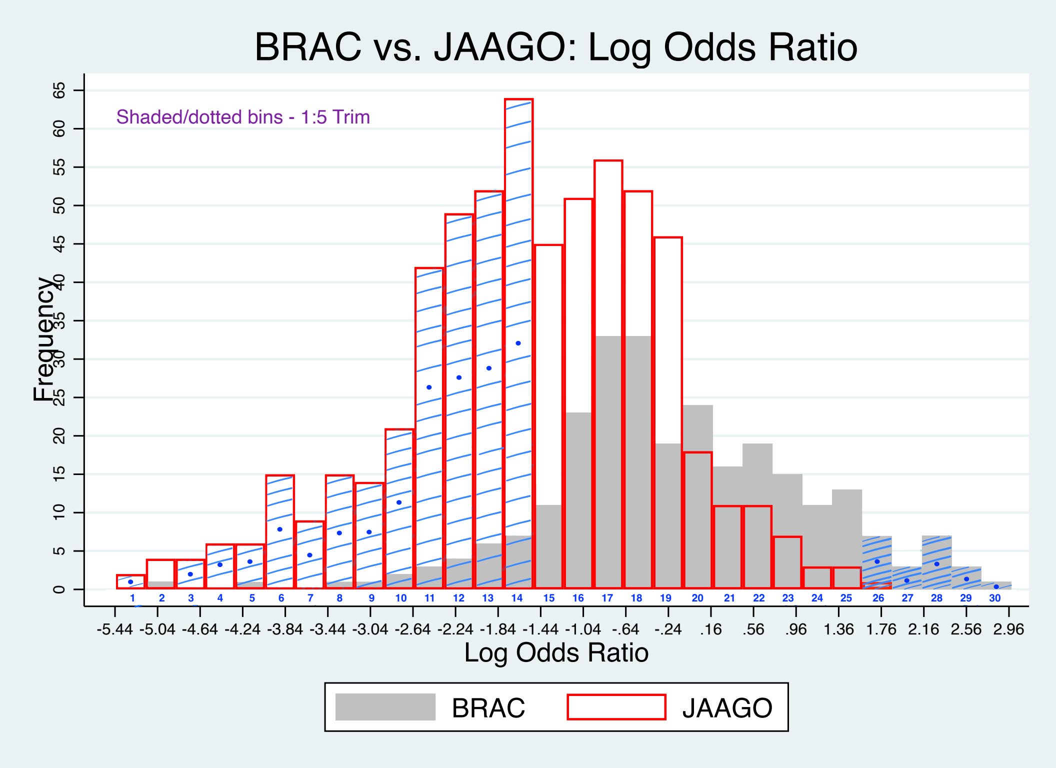
Table A.1: OLS Regression of Achievement on Conditioning Variables

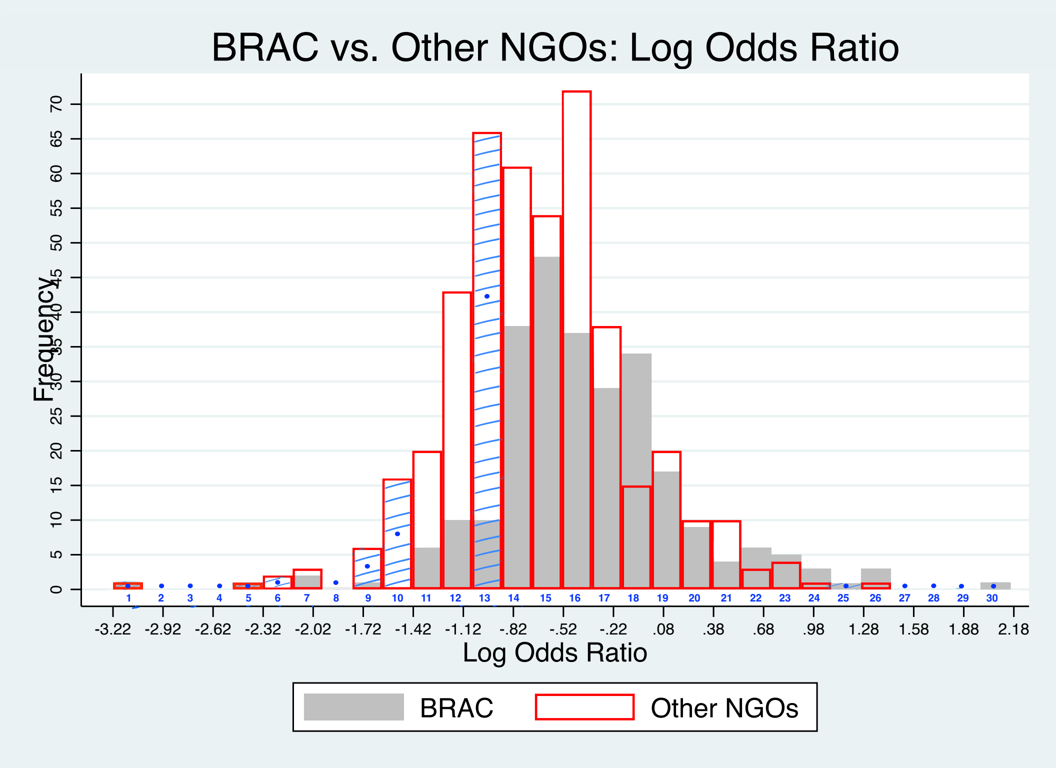
	Dependent Variable: Achievement Test Z-scores	
	(1) Family Background (Exclude K-BIT)	(2) Family Background (Include K-BIT)
Male	0.1439***	0.1418***
	(0.0481)	(0.0410)
Age (years)	0.0107	0.0075
	(0.0171)	(0.0157)
Father absent	-0.2689**	-0.2263**
	(0.1099)	(0.0913)
Family Size	-0.0000	0.0133
	(0.0228)	(0.0211)
Father's schooling	0.0351***	0.0253***
	(0.0064)	(0.0058)
Mother's schooling	0.0378***	0.0211***
	(0.0088)	(0.0062)
K-BIT (IQ)		0.4332***
		(0.0249)
Constant	-0.3753*	-0.3307**
	(0.1920)	(0.1513)
Observations	1936	1936

Notes: (a) Standard errors are in parentheses; (b) We report the regression estimates at the 1 percent, 5 percent, and 10 percent significance level denoted by *** , ** , and *, respectively.



BRAC vs. GOV: Log Odds Ratio
Shaded/dotted bins - 1:5 Trim
Frequency
0
5
10
15
20
25
30
35
40
45
50
55
60
65
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
-3.69 -3.39 -3.09 -2.79 -2.49 -2.19 -1.89 -1.59 -1.29 -.99 -.69 -.39 -.09 .21 .51 .81 1.11 1.41 1.71 2.01 2.31 2.61
Log Odds Ratio
BRAC
GOV


BRAC vs. JAAGO: Log Odds Ratio
Shaded/dotted bins - 1:5 Trim
Frequency
0
5
10
15
20
25
30
35
40
45
50
55
60
65
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
-5.44 -5.04 -4.64 -4.24 -3.84 -3.44 -3.04 -2.64 -2.24 -1.84 -1.44 -1.04 -.64 -.24 .16 .56 .96 1.36 1.76 2.16 2.56 2.96
Log Odds Ratio
BRAC
JAAGO


BRAC vs. Other NGOs: Log Odds Ratio
Frequency
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
-3.22 -2.92 -2.62 -2.32 -2.02 -1.72 -1.42 -1.12 -.82 -.52 -.22 .08 .38 .68 .98 1.28 1.58 1.88 2.18
Log Odds Ratio
BRAC
Other NGOs