The Impact of Outsourcing Schools in Liberia to BRAC after Three Years[1]

Mauricio Romero
(ITAM)

Justin Sandefur
(Center for Global Development)

September 23, 2020

Abstract

Outsourcing the management of 20 randomly-selected government primary schools in Liberia to BRAC led to learning gains of 0.06σ after three years, equivalent to reading roughly -1.3 fewer words per minute. While the effects on test scores are not statistically significant, BRAC reduced sexual abuse by teachers by -3.68 percentage points. In addition, BRAC had no statistically significant effect on dropout or enrollment, and failed to reduced corporal punishment.

[1]We are grateful to the Minister George K. Werner and his team, Minister Prof. Ansu. D Sonii, Sr and his team, the Partnership Schools for Liberia (PSL) team, Susannah Hares, Robin Horn, and Joe Collins from Ark EPG, and the team at Social Finance for their commitment throughout this project to ensuring a rigorous and transparent evaluation of the PSL/LEAP program. We are especially grateful to Wayne A. Sandholtz for his collaboration in the early stages of this project and subsequent discussions. Thanks to Arja Dayal, Dackermue Dolo, and their team at Innovations for Poverty Action who led the data collection. Avi Ahuja, Miguel Ángel Jiménez-Gallardo, Dev Patel, Rony Rodríguez-Ramírez, and Benjamin Tan provided excellent research assistance. We are grateful to Laura Johnson who provided guidance on the sexual violence survey module and protocol. A randomized controlled trials registry entry is available at: https://www.socialscienceregistry.org/trials/1501 as well as the pre-analysis plan. IRB approval was received from IPA (protocol #14227) and the University of Liberia (protocol #17-04-39) prior to any data collection. UCSD IRB approval (protocol #161605S) was received after the first round of data collection but before any other activities were undertaken. The evaluation was supported by the UBS Optimus Foundation, Aestus Trust, and the UK's Economic and Social Research Council (grant number ES/P00604). Romero acknowledges financial support from the Asociación Mexicana de Cultura. The views expressed here are ours, and not those of the Ministry of Education of Liberia or our funders. All errors are our own.

1 Introduction

In 2016, the Liberian government outsourced the management of 93 randomly-selected public schools, comprising 8.6% of public school students, to eight private providers, including for-profit companies and non-profit charities, as well as local and international organizations. The program bundled private management with, in theory, a doubling of education expenditure per child. In this report, we present results after three years for BRAC. Romero, Sandefur, and Sandholtz (2020) and Romero and Sandefur (2019) present results from the full program, across all eight providers.

Two features of the experimental design merit special emphasis: block randomization and intention-to-treat analysis. First, as we randomized treatment within matched pairs, our design amounts to eight experiments (one per provider). Hence, we are able to study heterogeneity across providers. Second, to avoid confounding the treatment effect of the program with sorting of students across schools, we sampled students from pre-treatment enrollment records and followed them for three years (we were able to interview over 96% of the original sample three years later). By assigning each student to their "original" school, regardless of what school (if any) they attend in later years, we are able to provide unbiased estimates of the program's intention-to-treat (ITT) effect.

The intention-to-treat (ITT) treatment effect of BRAC after three academic years is 0.09σ for English (p-value 0.28) and 0.06σ for math (p-value 0.52). The treatment-on-the-treated (ToT) estimate (i.e., the treatment effect for schools that were actually operated by BRAC) is 0.09σ for English (p-value 0.26) and 0.06σ for math (p-value 0.51).

The treatment effect on the likelihood that students enrolled in BRAC schools in 2015/2016 are enrolled in any school three years later is -1.1 percentage points (p-value 0.65) from a base of 86.08%. Turning to school-level enrollment, after three years, BRAC had a treatment effect on enrollment of 57.28 students per school (p-value 0.08).

We measure two aspects of child safety: corporal punishment and sexual abuse. Corporal punishment is widespread across schools: 59.25% in control schools report being hit by their teachers at least occasionally. The treatment effect of BRAC on this margin is 6.35 percentage points (p-value 0.16). Turning to sexual abuse, 6.35% of students in control schools report having sex with a teacher (statutory rape). The treatment effect of BRAC on this margin is -3.68 percentage points (p-value 0.05).

In the first year, the expenditure per pupil for BRAC was 72 USD per pupil. After three years, the average (self-reported) expenditure has changed to 30 USD per pupil. The government's spending target is 50 USD per pupil.

2 Experimental design

Below we summarize the most important features of the program and the experimental design. Further details are provided in Romero et al. (2020).

2.1 The program

2.1.1 Context

The government's primary motivation for the outsourcing program was the low levels of learning in public schools. At baseline $\sim$ 25% of pupils enrolled in fifth grade could not read a single word.

In addition to low learning levels, access remains an unresolved issue. The last nationally representative household survey prior to the experiment reported net primary enrollment at 38%, partially explained by high levels of over-age enrollment (Liberia Institute of Statistics and Geo-Information Services, 2016).

The experimental sample analyzed below is not intended to be representative of the country. Circa 2016, Liberia had 2,619 public primary schools across fifteen counties. To take part in the pilot, schools were required to meet minimum infrastructure standards. While thirteen counties were included in the pilot, only 299 schools satisfied all the criteria. Finally, providers were allowed to filter the list of potential pilot schools before random assignment based on proximity to roads and availability of 3G service, leaving a final sample of 185 eligible schools.[2]

Public primary school is nominally free in Liberia, though informal fees are common. In contrast, fees are permitted for pre-primary classes. At baseline, government spending on public primary schools was $\sim$ 50 USD per pupil, almost entirely devoted to teacher salaries.

2.1.2 Intervention

The Liberian Education Advancement Program (LEAP) — formerly, the Partnership Schools for Liberia (PSL) program — is a public-private partnership for school *management*. Under the program, the government delegated the management of 93 public schools, covering 8.6% of all public school students, to eight different private organizations. Providers were paid on a per-pupil basis and forbidden from charging fees or screening students based on ability.

[2]Schools in the RCT have better facilities and infrastructure than most schools in the country, which limits the external validity of the results (Romero et al., 2020).

Of the eight private organizations, three are for-profit: Bridge International Academies (allocated 23 schools), Omega Schools (allocated 19 schools), and Rising Academies (allocated 5 schools). The other five non-profit providers include BRAC (allocated 20 schools), Street Child (allocated 12 schools), More than Me (currently known as Hilltop Schools, allocated 6 schools), the Liberian Youth Network (currently known as the Youth Movement for Collective Action, allocated 4 schools), and Stella Maris (allocated 4 schools). Liberian Youth Network and Stella Maris are Liberian organizations, the other six are international. While Stella Maris never took control of their assigned schools, the government still considers them part of the program (e.g., they were allocated more schools in an expansion of the program not studied in this paper (Ministry of Education - Republic of Liberia, 2017)).

In contrast to some other public-private partnerships in education (e.g., U.S. charter schools), the teachers in the Liberian public schools that were outsourced to providers were to remain civil servants and were still paid by the government.

There are three noteworthy features of the evolution of the program since it started in 2016. In 2017, the program expanded to an additional 98 schools. These schools were not experimentally assigned and are not included in our analysis. Second, the program changed some of its operating rules. All providers were given uniform contracts (unlike the first year, when Bridge had a different contract) and the Ministry of Education did not allow capping class sizes.[3] Finally, the country had a presidential election in late 2017. The new administration, which took office in early 2018, claims it stopped prioritizing treatment schools in the assignment of teachers or in the process of bringing existing teachers onto the payroll.

Providers must teach the Liberian national curriculum, but beyond that they have flexibility in defining the intervention. They may choose to use school resources in different ways (e.g., providing remedial programs, prioritizing subjects, having longer school days, or other non-academic activities). They can also provide more inputs such as extra teachers, books, or uniforms, as long as they pay for them.

On paper, the Ministry of Education's financial obligation to treatment schools is the same as to any other government-run school: it provides teachers and maintenance, valued at about $\sim$ 50 USD per student. In addition, providers receive *extra* funding (of 50 USD per student), coordinated by the Ministry of Education but paid by third-party philanthropies. Providers have complete autonomy over the use of these funds. In addition, providers may raise more funds on their own. In the first year, the expenditure

[3] In the experimental sample this had little effect as student expulsions in the first year meant few classes remained above the cap.

per pupil for BRAC was 72 USD per pupil. After three years, the average (self-reported) expenditure has changed to 30 USD per pupil. The government's spending target is 50 USD per pupil.

2.2 Sampling and random assignment

Two key features of the sampling and randomization process are that (a) providers agreed to a list of schools they would be willing to serve before random assignment, and (b) pupils were sampled from lists made before the program began and tracked regardless of where they went.

Based on providers' preferences and requirements, the list of 185 eligible schools was non-randomly partitioned across providers. The schools allocated to each provider were then paired based on their infrastructure quality. Finally, within each pair schools were randomly assigned to treatment or control. Providers did not manage all the schools originally assigned to treatment and we treat these schools as non-compliant, presenting results in an intention-to-treat framework.

Treatment assignment may change the student composition across schools. To prevent differences in the composition of students from driving differences in outcomes, we sampled 20 students per school (from K1 to grade 5) from enrollment logs from the 2015/2016 school year, before the treatment was introduced. We associate each student with his or her "original" school, regardless of what school (if any) he or she attended in subsequent years. The combination of random treatment assignment at the school level with measuring outcomes of a fixed and comparable pool of students allows us to provide unbiased estimates of the program's intention-to-treat (ITT) effect within the student population originally attending study schools.

2.3 Timeline of research and intervention activities

We collected data in schools three times: at the beginning of the school year in September/October 2016, at the end of the school year in May/June 2017, and in March/April of 2019.

2.4 Test design

In all three rounds of data collection, we conducted one-on-one tests in which an enumerator sits with the student, asks questions, and records the answers since literacy

cannot be assumed at any grade level. All students took the same adaptive test, regardless of the grade. Stop rules instructed enumerators to skip higher-order skills if the student is not able to answer questions related to more basic skills. We estimate an item response theory (IRT) model for each round of data collection. Following standard practice, we normalize the IRT scores with respect to the control group.

2.5 Additional data

We surveyed all the teachers in each school and conducted in-depth surveys with those teaching math and English. For a randomly selected class within each school, we conducted a classroom observation using the Stallings Classroom Observation Tool (World Bank, 2015). Furthermore, we conducted school-level surveys to collect information about school facilities, the teacher roster, input availability (e.g., textbooks), and expenditures.

Given the concerns about child safety in program schools raised by the sexual abuse scandals involving two of the providers (Baysah, 2016; F. Young, 2018), we added a sexual violence module to the student survey. Sexual abuse is difficult to measure and rarely reported through official channels in Liberian schools. We collected data via an anonymous survey to students twelve years old and above, where enumerators asked the student questions regarding sexual abuse at school (by teachers and peers) and at home. The student filled in an anonymous answer sheet (pre-filled with the school id and the gender of the child) and placed it in a closed ballot box.

2.6 Balance and attrition

Given our study design, we put considerable effort and resources into minimizing attrition (extra training, generous tracking time, and specialized tracking teams). Students were tracked to their homes and tested there. Attrition in the third wave of data collection is balanced between treatment and control and is below 0.26% (see Table 1). The total attrition rate, for both treatment and control, is 2.62 %)

Table 1: % Interviewed

	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)
% interviewed	97.38	0.26	0.26
	(16)	[0.81]	[0.81]
Observations	389	780	780

Notes: This table presents the attrition rate (proportion of students interviewed in Year 3). Column 1 shows the mean and standard deviation (in parentheses) for the control in Year 3. The differences between treatment and control for Year 3 are presented in Column 2 and p-value that comes from 5,000 randomization inference iterations (in brackets). The differences take into account the randomization design (i.e., including "pair" fixed effects).

3 Results

We estimate the impact of BRAC's management public schools on four margins: 1) *access*, defined as impacts on enrollment and grade attainment for a fixed sample of pupils; 2) *learning*, as measured by test scores; 3) *sustainability*, which hinges, in part, on whether the program effects come from increases in material inputs or staffing versus improvements in school management; and 4) *child safety*, as measured by pupil surveys on corporal punishment and sexual abuse. An important caveat with these results is that there is limited power given the sample size. Given the small number of schools per provider, we use randomization inference which provides exact tests of sharp hypotheses no matter the sample size (A. Young, 2018).

While we focus on the intent-to-treat (ITT) effect as the key policy-relevant parameter, we also report treatment-on-the-treated (ToT) estimates. There are two sources of non-compliance in our experiment: school-level non-compliance when providers failed to take control of all of the schools assigned to them, and student-level non-compliance when students left their original school, either voluntarily or because providers excluded them. Since non-compliance is unlikely to be random, we use the random assignment as an instrument for compliance to estimate the ToT. While the assumptions required for ToT estimates — monotonicity and the stable unit treatment value assumption (SUTVA) — appear reasonable at the school level, they may not be at the student level. As shown above and in Romero et al. (2020), treatment caused some students to leave (voluntarily

or not) their assigned schools, violating monotonicity. In addition, a student's peers (and related peer-effects) may have changed as some peers were forced out of school by the treatment, violating SUTVA at the student level. Relatedly, the likelihood that two peers stay in the same school is negatively correlated. Thus, we below present also ToT estimates at the school-level, which are estimated using the assigned treatment as an instrument for whether the school was is in fact operated by BRAC.[4]

3.1 Access

First, we focus on school-level enrollment. After three years, BRAC had a treatment effect on enrollment of 57.28 students per school (p-value 0.08). Provider compensation is based on the number of students enrolled rather than the number of students actively attending school. The treatment effect on student attendance is 20.32 percentage points (p-value 0) from a base of 29.48%. See Panel A, Table 2 for more details.

It is important to remember that BRAC was not allowed to charge fees and their schools should be free at all levels, including early-childhood education (ECE). In contrast, control schools are officially permitted to charge fees to ECE students and charge informal fees to primary students. The treatment effect on the likelihood that principals report charging fees in primary in BRAC schools is -15 percentage points (p-value 0.38) from a base of 70%.

Turning to student-level enrollment, the treatment effect on the likelihood that students enrolled in BRAC schools in 2015/2016 are enrolled in any school three years later is -1.1 percentage points (p-value 0.65) from a base of 86.08%. See Panel B, Table 2.

We classify the reasons why students are no longer attending school into four broad categories: 1) Left school to work, 2) pregnancy, 3) could not afford school fees, and 4) others. Students originally enrolled in BRAC schools are 1.38 percentage points (p-value 0.25) more likely to drop out of schools because of pregnancy (from a base of 3.68%).[5] In addition, the treatment effect on the likelihood that students are enrolled in some form of secondary school is -2.57 percentage points (p-value 0.2) from a base of 21.32%.

[4] Apart from the identification concerns, a student-level ToT estimate may also be less policy-relevant, as the ethics of forcing students/parents to enroll in treatment schools or forbidding them to move are questionable.

[5] This result need not imply providers changed the rate of teen pregnancy; alternatively, they may have changed the enforcement of the national policy requiring pregnant girls to drop out of school until after childbirth (Martinez & Odhiambo, 2018).

Table 2: ITT treatment effects on enrollment, attendance, and selection

	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)
Panel A: School level data			
Enrollment change	-89.65	57.28	57.28
	(85.66)	[0.08]	[0.08]
Attendance % (spot check)	29.48	20.32	20.32
	(25.94)	[0]	[0]
% of students with disabilities	0.67	0.26	0.26
	(1.04)	[0.6]	[0.6]
Observations	20	40	40
Panel B: Student level data			
% enrolled in the same school	47.1	3.5	3.5
	(49.98)	[0.48]	[0.48]
% enrolled in school	86.08	-1.1	-1.1
	(34.66)	[0.65]	[0.65]
Days missed, previous week	0.54	0.23	0.23
	(1.12)	[0.11]	[0.11]
Observations	380	758	758

Notes: This table presents the mean and standard deviation (in parentheses) for the control (Column 1 in Year 3), as well as the the difference between the two taking into account the randomization design (i.e., including "pair" fixed effects) in Column 2 in Year 3. Panel A presents school level data including enrollment (taken from enrollment logs) and student attendance measured by our enumerators during a spot check in the middle of a school day. If the school was not in session during a regular school day we mark all students as absent. The fraction of students identified as disabled in our sample is an order of magnitude lower than estimates for the percentage of disabled students in the U.S and worldwide using roughly the same criteria (both about 5%) (Brault, 2011; UNICEF, 2013). Panel B presents student-level data including whether the student is still enrolled in the same schools, whether he/she is enrolled in school at all, and whether he/she missed school in the previous week (conditional on being enrolled in school). p-values from 5,000 randomization inference iterations are presented in brackets.

3.2 Learning

The intention-to-treat (ITT) treatment effect of BRAC after three academic years is 0.09σ for English (p-value 0.28) and 0.06σ for math (p-value 0.52), as shown in Column 1 of Table 3.

While we focus on the ITT effect, we also report treatment-on-the-treated (ToT) estimates (i.e., the treatment effect for schools that were actually operated by BRAC). The ToT effect is 0.09σ for English (p-value 0.26) and 0.06σ for math (p-value 0.51), as shown in Column 2 of Table 3.

An important concern when interpreting these results expressed in standard deviations is how much learning they represent. We use correct words per minute as a benchmark. Students enrolled in Grade 1 in 2015/2016 in control schools are able to read 9.6 words per minute on average in 2019. Their counterparts in treatment schools can read 11.4 words per minute. For students enrolled in Grade 5 in 2015/2016, those in control schools can read 22.4 words per minute and those in treatment schools can read 23.2 words per minute in 2019. As a benchmark, to understand a simple passage students should read 45-60 words per minute (Abadzi, 2011).

Table 3: ITT treatment effects on learning

	ITT	TOT
	(1)	(2)
English	0.09	0.09
	[0.28]	[0.26]
Math	0.06	0.06
	[0.52]	[0.51]
Composite	0.06	0.06
	[0.47]	[0.46]
Controls	Yes	Yes
Observations	758	758

Notes: Columns 1 and 2 are based on the third wave of data and show the difference between treatment and control taking into account the randomization design — i.e., including "pair" fixed effects and student and school controls (Column 1), and the treatment-on-the-treated (ToT) estimates (Column 2). The treatment-on-the-treated effects are estimated using the assigned treatment as an instrument for whether the student is in fact enrolled in a partnership school at the time of data collection. p-values from 5,000 randomization inference iterations are presented in brackets.

3.3 Sustainability

The outsourcing program changed the management of treated schools, while also increasing the total resources available to them (this was true even before it surpassed the original budget targets). The sustainability of the program depends in part on the relative importance of these two channels. Furthermore, some of these changes may have imposed negative externalities on the broader school system, by shifting students (see Section 3.1) and under-performing teachers to non-program schools. While we do not attempt any formal mediation analysis to quantify the role of competing mechanisms, this section estimates the effect of BRAC on school resources and management, and explores some of the potential negative externalities.

3.3.1 Inputs and resources

First, we focus on a key input in the education production function: teachers. In the first year, the Ministry of Education agreed to release some underperforming teachers from program schools, replace those teachers, and provide additional new teachers. After three years, BRAC had 3.65 more teachers on average (p-value 0) than control schools (from a base of 7.1 teachers). The treatment effect on the pupil-teacher ratio was -4.6 (p-value 0.08) from a base of 26.82.

Panel B in Table 4 compares the composition of teachers in BRAC schools compared to control schools. Finally, we compare the materials available to students during classroom observations (see Panels C - Table 4). The treatment effect on the likelihood that students have a pen is 8.89 percentage points (p-value 0.36) from a base of 86.3%.

Table 4: ITT treatment effects on inputs and resources

	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)
Panel A: School-level data			
Number of teachers	7.1	3.65	3.65
	(1.92)	[0]	[0]
Pupil-teacher ratio (PTR)	26.82	-4.6	-4.6
	(9.53)	[0.08]	[0.08]
New teachers	2.75	0.95	0.95
	(1.07)	[0.05]	[0.05]
Teachers dismissed	1.7	-0.55	-0.55
	(1.08)	[0.18]	[0.18]
Observations	20	40	40
Panel B: Teacher-level data			
Age in years	40.69	-4.08	-4.08
	(6.84)	[0]	[0]
Experience in years	11.06	-2.67	-2.67
	(5.58)	[0.05]	[0.05]
% has worked at a private school	25.3	3.34	3.34
	(22.43)	[0.64]	[0.64]
Test score in standard deviations	-0.39	0.25	0.25
	(0.63)	[0.05]	[0.05]
% certified (or tertiary education)	48.22	11.88	11.88
	(23.27)	[0.02]	[0.02]
Salary (USD / month)–Conditional on salary > 0	63.18	43.5	43.5
	(37.45)	[0]	[0]
Observations	20	40	40
Panel C: Classroom observation			
Number of seats	18.54	2.77	2.77
	(10.52)	[0.29]	[0.29]
% of students with pens / pencils	86.3	8.89	8.89
	(25.18)	[0.36]	[0.36]
Observations	13	26	26

Notes: This table presents the mean and standard deviation (in parentheses) for the control (Column 1 in Year 3), as well as the the difference taking into account the randomization design (i.e., including "pair" fixed effects) in Column 2 in Year 3. Panel A has school level outcomes. Panel B presents teacher-level outcomes including their score in tests conducted by our survey teams. Panel C presents data on inputs measured during classroom observations. p-values from 5,000 randomization inference iterations are presented in brackets.

3.3.2 School management

The treatment effect on the likelihood that schools are open during a regular school day (i.e., the school is open, students and teachers are on campus, and classes are taking place) is 10 percentage points (p-value 0.5) from a base of 75%. The impact on the length of the school day is 4.61 hours per week (p-value 0) from a base of 21.49 hours.

The treatment effect on management practices (as measured by a "good practices" PCA index normalized to a mean of zero and standard deviation of one in the control group) is 0.4σ (p-value 0.21). We also measure management practices using a a D-WMS style survey (Lemos & Scur, 2016). According to this index, the treatment effect on management practices was 0.68σ (p-value 0.17).

Table 5: ITT treatment effects on school management

	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)
% school in session at spot check	75	10	10
	(44.43)	[0.5]	[0.5]
Instruction time (hrs/week)	21.49	4.61	4.61
	4.4	[0]	[0]
Principal's working time (hrs/week)	24.82	-4.71	-4.71
	(10.08)	[0.06]	[0.06]
% of principal's time spent on management	39.27	20.87	20.87
	(17.96)	[0.01]	[0.01]
Index of good practices (PCA)	-0.02	0.4	0.4
	(1.1)	[0.21]	[0.21]
Management index (DWMS-style)	-0.4	0.68	0.68
	(1.06)	[0.17]	[0.17]
Observations	15	30	30

Notes: This table presents the mean and standard deviation (in parentheses) for the control (Column 1 in Year 3), as well as the the difference taking into account the randomization design (i.e., including "pair" fixed effects) in Column 2 in Year 3. The index of good practices is normalized to have mean zero and standard deviation of one in the control group. The management index is based on Development World Management Survey (DWMS) style-questions. p-values from 5,000 randomization inference iterations are presented in brackets.

3.3.3 Teacher behavior

The treatment effect on the likelihood that teachers are in schools and in a classroom during a spot check is -3.5 percentage points (p-value 0.63) and 2.28 percentage points (p-value 0.77) (from a base of 51.5% and 34.44%).

Classroom observations allow us to see further changes in teacher behavior and pedagogical practices. The treatment effect on the likelihood that teachers are off-task is -13 percentage points (p-value 0.21) from a base of 41.5%. The treatment effect on the likelihood that teachers are on-task (either active or passive instruction) is 13 percentage points (p-value 0.17) from a base of 46.5%.

Table 6: ITT treatment effects on teacher behavior

	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)
Panel A: Spot checks			
% on schools campus	51.5	-3.5	-3.5
	(29.3)	[0.63]	[0.63]
% in classroom	34.44	2.28	2.28
	(30.6)	[0.77]	[0.77]
Observations	20	40	40
Panel B: Student reports			
Teacher missed school previous week (%)	40.24	-1.39	-1.39
	(16.02)	[0.75]	[0.75]
Teacher never hits students (%)	40.75	6.35	6.35
	(16.08)	[0.16]	[0.16]
Observations	20	40	40
Panel C: Classroom observations			
Instruction (active + passive) (% of class time)	46.5	13	13
	(39.11)	[0.17]	[0.17]
Classroom management (% class time)	12	0	0
	(14.73)	[0.91]	[0.91]
Teacher off-task (% class time)	41.5	-13	-13
	(45.57)	[0.21]	[0.21]
Student off-task (% class time)	31.3	-5.6	-5.6
	(37.84)	[0.5]	[0.5]
Observations	20	40	40

Notes: This table presents the mean and standard deviation (in parentheses) for the control (Column 1 in Year 3), as well as the the difference taking into account the randomization design (i.e., including "pair" fixed effects) in Column 2 in Year 3. Panel A presents data from spot checks conducted by our survey teams in the middle of a school day. Panel B presents data from our panel of students where we asked them about their teachers' behavior. Panel C presents data from classroom observations. If the school was not in session during a regular school day we mark all teachers not on campus as absent and teachers and students as off-task in the classroom observation. p-values from 5,000 randomization inference iterations are presented in brackets.

3.4 Child safety

We measure two aspects of child safety: corporal punishment and sexual abuse. Corporal punishment is widespread across schools: 59.25% in control schools report being hit by their teachers at least occasionally. The treatment effect of BRAC on this margin is

6.35 percentage points (p-value 0.16).

Sexual abuse rates in our data are lower than those reported in previous studies (Postmus et al., 2015; Steiner, Johnson, Postmus, & Davis, 2018): 6.35% of students in control schools report having sex with a teacher (statutory rape). The treatment effect of BRAC on this margin is -3.68 percentage points (p-value 0.05).[6] Is possible that the likelihood of *reporting* an incident may have changed in program schools. Reported cases of forced sexual intercourse at home — where the true rate is unlikely to be affected by the program — changed by -1.25 percentage points (p-value 0.49) from a base of 4.76%.

[6] In a companion paper, Johnson, Romero, Sandefur, and Sandholtz (2019) compare the survey protocol we used with a protocol identical to the one used by Postmus et al. (2015) and Steiner et al. (2018) and find similar rates of sexual abuse across protocols.

Table 7: Gender based violence

	All			Boys			Girls		
	Control	Treatment Effect	Treatment Effect TOT	Control	Treatment Effect	Treatment Effect TOT	Control	Treatment Effect	Treatment Effect TOT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Teacher: Sex	6.35	-3.68	-3.68	7.87	-4.63	-4.63	4.38	-2.59	-2.59
	(24.42)	[0.05]	[0.05]	(27)	[0.07]	[0.07]	(20.54)	[0.46]	[0.07]
Teacher: Touched	5.71	1.26	1.26	7.87	1.81	1.81	2.92	1.26	1.26
	(23.25)	[0.53]	[0.53]	(27)	[0.63]	[0.63]	(16.9)	[0.61]	[0.63]
Teacher: Forced sex	2.54	-0.33	-0.33	2.81	0.15	0.15	2.19	-0.78	-0.78
	(15.76)	[0.77]	[0.77]	(16.57)	[0.84]	[0.85]	(14.69)	[0.63]	[0.85]
Student: Touched	11.78	3.25	3.25	11.86	5.73	5.73	11.68	-1.31	-1.31
	(32.29)	[0.24]	[0.24]	(32.43)	[0.2]	[0.2]	(32.23)	[0.72]	[0.2]
Student: Forced sex	3.17	0.87	0.87	4.49	0.26	0.26	1.46	1.45	1.45
	(17.56)	[0.65]	[0.65]	(20.78)	[0.94]	[0.94]	(12.04)	[0]	[0.94]
Family: Touched	10.48	-2.38	-2.38	11.8	-3.9	-3.9	8.76	-0.7	-0.7
	(30.67)	[0.39]	[0.39]	(32.35)	[0.33]	[0.33]	(28.37)	[0.81]	[0.33]
Family: Forced sex	4.76	-1.25	-1.25	3.37	1.91	1.91	6.57	-4.88	-4.88
	(21.33)	[0.49]	[0.49]	(18.1)	[0.47]	[0.46]	(24.87)	[0.09]	[0.46]
Observations	315	617	617	178	341	341	137	276	276

Notes: This table presents the mean and standard deviation (in parentheses) for the control as well as the difference and p-value that comes from 5,000 randomization inference iterations (in brackets) taking into account the randomization design (i.e., including "pair" fixed effects) for all students (Column 1-3), only boys (Column 4-6), and only girls (Columns 7-8).

References

Abadzi, H. (2011). *Reading fluency measurements in EFA FTI partner countries: Outcomes and improvement prospects.* World Bank.

Baysah, A. M., Jr. (2016). *Liberia: Police charge youth activist for sodomy.* Retrieved from https://web.archive.org/web/20161103182507/https://allafrica.com/stories/201611020824.html

Brault, M. (2011). *School-aged children with disabilities in U.S. metropolitan statistical areas: 2010. American community survey briefs* (Tech. Rep.). ACSBR/10-12. US Census Bureau.

Johnson, L., Romero, M., Sandefur, J., & Sandholtz, W. (2019). *Comparing three measures of sexual violence in Liberian schools.* (Mimeo)

Lemos, R., & Scur, D. (2016). *Developing management: An expanded evaluation tool for developing countries.* (mimeo)

Liberia Institute of Statistics and Geo-Information Services. (2016). *Liberia - household income and expenditure survey 2014-2015*. Liberia Institute of Statistics and Geo-Information Services.

Martinez, E., & Odhiambo, A. (2018). *Leave no girl behind in Africa: Discrimination in education against pregnant girls and adolescent mothers.* (Tech. Rep.). Human Rights Watch. Retrieved from https://www.hrw.org/sites/default/files/report_pdf/au0618_insert_webspreads.pdf

Ministry of Education - Republic of Liberia. (2017). *PSL school allocation: Decision points.* Retrieved 28/07/2017, from http://moe.gov.lr/wp-content/uploads/2017/06/Allocation-final.pdf

Postmus, J. L., Hoge, G. L., Davis, R., Johnson, L., Koechlein, E., & Winter, S. (2015). Examining gender based violence and abuse among liberian school students in four counties: An exploratory study. *Child abuse & neglect*, *44*, 76–86.

Romero, M., & Sandefur, J. (2019). Beyond short-term learning gains: The impact of outsourcing schools in Liberia after three years. *Center for Global Development Working Paper*, *521*.

Romero, M., Sandefur, J., & Sandholtz, W. A. (2020). Outsourcing education: Experimental evidence from Liberia. *American Economic Review*, *110*(2), 364-400. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20181478 doi: 10.1257/aer.20181478

Steiner, J. J., Johnson, L., Postmus, J. L., & Davis, R. (2018). Sexual violence of liberian school age students: An investigation of perpetration, gender, and forms of abuse.

Journal of child sexual abuse, 1–20.

UNICEF. (2013). *The state of the world's children: Children with disabilities* (Tech. Rep.). United Nations.

World Bank. (2015). *Conducting classroom observations: analyzing classrooms dynamics and instructional time, using the stallings' classroom snapshot'observation system. user guide* (Tech. Rep.). World Bank Group.

Young, A. (2018). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics*, *134*(2), 557-598. doi: 10.1093/qje/qjy029

Young, F. (2018). *Unprotected.* Retrieved from https://features.propublica.org/liberia/unprotected-more-than-me-katie-meyler-liberia-sexual-exploitation/