Partnership Schools for Liberia: Pre-Analysis Plan

Mauricio Romero* Justin Sandefur† Wayne Aaron Sandholtz*

May 10, 2017

Abstract

This document presents the pre-analysis plan (PAP) for the external evaluation of the Partnership Schools for Liberia (PSL) program, and is being prepared and registered prior to the principal investigators' receipt of any post-treatment data. The PAP will guide the analysis of year-1 impacts to be measured using the survey data collected in May-June 2017, as well as possible extensions up until 2019 conditional on continuation of the project and preservation of the control group. The primary outcomes of interest are math and English tests scores, school enrollment, pupil attendance, and selection of pupils by operators. Sub-group analysis will vary by outcome, as specified below, on dimensions including student wealth, school type (for-profit vs. non-profit), concentration of competing schools, and oversubscription of enrollment.

1 Introduction

While the global education agenda has pivoted from access to quality on the heels of widespread progress toward universal primary education, Liberia has been left behind: Net primary enrollment stood at only 38% in 2014 (The World Bank, 2014). Learning levels are also low: Among adult women who attended secondary school (or higher), only 35% can read a complete sentence (Liberia Institute of Statistics and Geo-Information Services, 2014). Civil war and the Ebola epidemic contributed to the weakening of the country's educational institutions, where capacity was already limited. Faced with these dire statistics, the Liberian Ministry of Education announced in early 2016 that it would contract the operation of some public primary schools to a group of private entities.

This program, known as the Partnership Schools for Liberia (PSL), delegated management of 93 randomly assigned public schools to a variety of private, for-profit companies and non-profit organizations. The government (via philanthropic entities) provides these operators with funding on a per-pupil basis. In exchange, operators are responsible for the daily management of the schools and are held responsible for academic results. These schools are to remain free and non-selective public schools (i.e., operators are not allowed to charge fees or screen students based on ability or other characteristics). The government retains ownership of PSL school buildings and teachers in PSL schools continue to be government employees (i.e., public servants).

Previous studies on PPPs in education have revealed mixed results. Most of these studies focus on charter schools in the United States, with some exceptions. A recent review of the literature on

*University of California, San Diego
†Center for Global Development

educational PPPs in the U.S. concludes that "charter school effects are highly variable, which likely reflects variations in the quality of education provided both at charter schools and at comparison schools" (Betts & Tang, 2016). Quasi-experimental methods have been used to study the effects on academic outcomes of charter schools in Michigan (Bettinger, 2005), North Carolina (Bifulco & Ladd, 2006), California (Zimmer & Buddin, 2006), Florida (Sass, 2006), Texas (Booker, Gilpatric, Gronberg, & Jansen, 2007; Hanushek, Kain, Rivkin, & Branch, 2007; Booker, Gilpatric, Gronberg, & Jansen, 2008), Wisconsin (Witte, Weimer, Shober, & Schlomer, 2007), New York City (Winters, 2012), an anonymous "large urban school district in the southwest of the U.S." (Imberman, 2011), Venezuela (Allcott & Ortega, 2009), and Colombia (Barrera-Osorio, 2007; Bonilla, 2010; Termes et al., 2015). These studies are limited in their ability to distinguish the effects of student selection and sorting from the causal effect of the PPP itself on test scores.

An alternative to overcome endogeneity issues that arise from student selection or sorting is to use admission lotteries that essentially randomize who enrolls in charter schools.[1] Admission lotteries have been used to study the effect of charter schools in Chicago (Hoxby & Rockoff, 2004), New York City (Hoxby, Murarka, & Kang, 2009), Boston (Abdulkadiroglu et al., 2009; Angrist, Cohodes, Dynarski, Pathak, & Walters, 2016), and schools across 13 states in the U.S. (Gleason, Clark, Tuttle, & Dwoyer, 2010; Clark, Gleason, Tuttle, & Silverberg, 2015). Although these studies are internally valid, external validity remains an issue as charter schools that are oversubscribed are likely different from those that are not (Tuttle, Gleason, & Clark, 2012).

Barrera-Osorio, Galbert, Gaspard, Habyarimana, and Sabarwal (2016) address the external validity concerns raised by Tuttle et al. (2012) by randomly assigning which schools become PPP schools in Uganda. However, internal validity concerns re-emerge here, as student selection or sorting is an issue when treatment is randomized at the school level (as opposed to the student level, as in studies using admission lotteries). Treating a school may cause changes in enrollment that change the observable and un-observable characteristics of students in treatment schools. Thus, differences in test scores between PPP and control schools may be driven by variation in the underlying population of students. Along the same lines, Barrera-Osorio et al. (2013) study a program in which 161 villages randomly selected in Pakistan are provided with schools through a PPP. The program led to an increase in enrollment of over 30% and the authors find that students in PPP schools scored higher on math and English tests. However, the population of children enrolled in school is different in treatment and control villages (due to the treatment itself).

Unlike previous studies, we are able to cleanly identify the effectiveness of PPP schools under an intention-to-treat framework that yields both internally and externally valid estimates. We do so by randomly assigning treatment status at the school level and sampling students from enrollment records for the year prior to the implementation of the program (before the public was aware of the program). Each student will be evaluated as part of his or her "original" school, regardless of what school (if any) he or she attended in subsequent years. Therefore, the population of students we use to test the treatment effect is the same across treatment and control, and not affected by changes in enrollment or selection into treatment schools.[2]

[1] For a review of the literature see Chabrier, Cohodes, and Oreopoulos (2016).

[2] Although we partially overcome the external validity issues of estimates from admission-based lotteries, we randomize treatment within a set of schools that meet certain eligibility criteria. Therefore, our estimates only apply to schools in Liberia that meet those criteria. See Section 2.3 for more details.

2 Experimental design

2.1 Context

Over the past decade, Liberia's basic education budget has been roughly $40 million per year (the majority of which is spent on salaries), while external donors spend about $30 million. Almost all of that aid bypasses the Ministry of Education, financing a diverse and often uncoordinated array of donor contractors and NGO programs in both government and non-state schools. This distinguishes Liberia from most other low-income countries in Africa, which finance the vast bulk of education spending through domestic tax revenue (UNESCO, 2016).

Aid dependence and the limited reach of the state call into question the relevance of some basic analytical assumptions in education systems analysis. The *World Development Report* (2004) framework that underlies much recent work on education systems research posits that parents hold schools accountable, in part, through their elected representatives who finance and delegate responsibilities to ministries and ultimately schools and teachers. In Liberia, more than half of children in preschool and primary school attend non-state schools (Ministry of Education, 2016). Even when parents send their children to government schools, a combination of user fees and donor projects may be more important to school accountability than "citizen power".

The second key feature of Liberia's education system relates to its performance: Not only are learning levels low, but simple access to basic education and progression through school remains woefully inadequate. While the global education agenda has pivoted from 'access' to 'quality' on the heels of widespread progress toward universal primary education, Liberia has been left behind. Net primary enrollment stood at only 38% in 2014 (The World Bank, 2014). Liberia's schools have an extraordinary backlog of over-age children, particularly in early childhood education: The median age of early childhood education enrollees in Liberia is eight years old(Liberia Institute of Statistics and Geo-Information Services, 2016). Learning levels are also low: Among adult women who finish elementary school, only 35% can read a complete sentence (Liberia Institute of Statistics and Geo-Information Services, 2014).

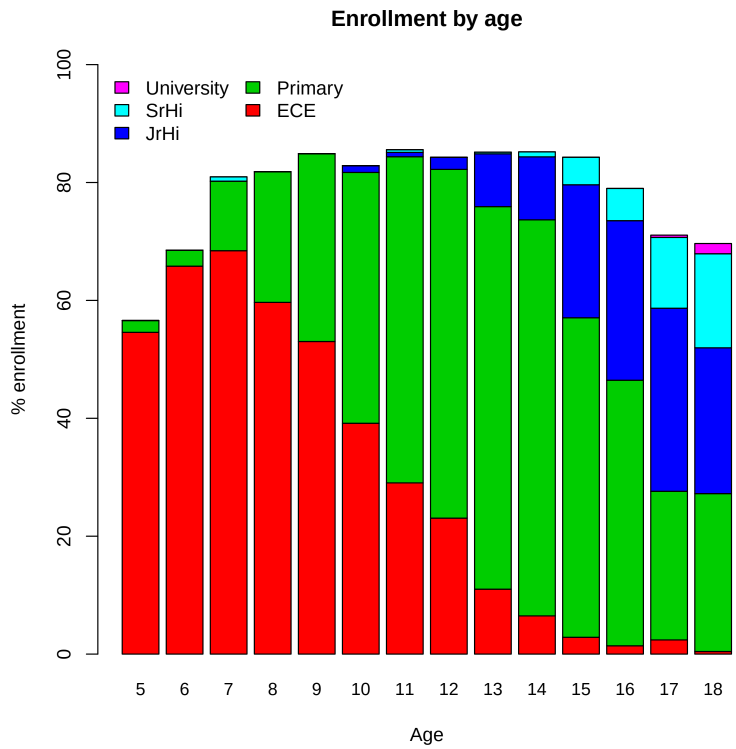

Figure 1: Authors' calculations based on 2014 Household Income and Expenditures Survey

2.2 Intervention

The Partnership Schools for Liberia (PSL) program is a public-private partnership (PPP) for school *management*. Specifically, the Government of Liberia contracted multiple non-state operators to run existing public primary schools (PSL schools). Operators receive funding on a per-pupil basis. In exchange, operators are responsible for the daily management of the schools.

PSL schools will continue to be free and non-selective public schools (i.e., operators are not allowed to charge fees or choose which students to enroll). Note that this is not true for normal government schools: Although public primary education is free[3], tuition for early childhood education is is LBD 3,500 per year (about USD 38).

PSL school buildings will remain under the ownership of the government. Teachers in PSL schools will be employed by the government, and drawn primarily from existing government teachers (i.e., civil servants). In other words, The Ministry of Education's financial obligation to PSL schools is the same as all government-run schools: provide teachers and maintenance. A noteworthy feature of PSL is that operators receive additional funding of USD 50 per student from outside donors (with

[3]Officially, public schools are free, but in reality most schools charge informal fees. See Section 2.4 for statistics on these fees.

a maximum of USD 3,250 per grade or 65 students per grade). On top of that, operators may raise more funds on their own to provide their schools with extra inputs.

Each operator is free to manage schools as they see fit. Operators are required to teach the Liberian national curriculum, but may supplement it with remedial programs, prioritization of subjects, longer school days, and non-academic activities. They are also welcome to provide additional inputs such as extra teachers, books or uniforms, as long as they pay for them.

Eight operators have contracted with the government to manage public schools under the PSL program. The operators, ordered by the number of schools they manage, are: Bridge International Academies is managing 23 schools, BRAC is managing 20 schools, Omega Academies 19, the Liberia Youth Network 4, More than Me 6, Rising Academies 5, Stella Maris 4, and Street Child is managing 12 schools.[4] Table 1 summarizes the differences between treated (PSL) and control (traditional public) schools.

	Control schools	Treatment schools
Management	MoE	Private operator
Teachers	Unionized civil servants	Unionized civil servants [Promised 1 teacher per grade]
Extra funding	None	USD 50 per student
Primary Fees	None	None
ECE fees	$\sim$ USD 40 per year	None
Selection	First come, first served	First come, first served [Class size caps of 45-60 pupils]
Building	Government of Liberia	Government of Liberia
Curriculum	National	National+Supplement

2.3 Sampling

Liberia has 2,619 public schools. Between the operators and the government it was agreed that potential PSL schools should: have 6 classrooms and six teachers, have good road access (as defined by the Education Management Information System [EMIS] data), be single shift, and not have a secondary school within the same compound. Only 299 schools satisfied all the criteria. Note that some of these are "soft" constraints that can be addressed if the program expands. For example, classrooms can be built and teachers added to the school staff. Figure 2 shows all public schools in Liberia and those within our sample. Table 1 shows the difference between schools in the RCT and other public schools. In general, schools in the RCT have more students and better infrastructure, and are closer to Monrovia (Ministry of Education).

[4] Bridge International Academies is managing two additional demonstration schools that were not randomized and are therefore not part of our sample. Omega Academies opted not to operate two of their assigned schools, which we treat as non-compliance. Rising Academies opted for not operating one of their assigned schools (which we treat as non-compliance), but was given one non-randomly assigned school in exchange (which is outside the RCT). Therefore, the set of schools in our analysis is not perfectly aligned with the set of schools actually being managed by PSL operators.

Table 1: External validity: Difference in characteristics between schools in the RCT (both treatment and control) and other public schools (based on EMIS data)

	RCT	Other Public	Difference
Students ECE	142.91	112.71	30.20***
	(73.99)	(66.46)	(5.82)
Students Primary	151.42	132.38	19.04*
	(131.48)	(143.57)	(10.28)
Students	291.99	236.24	55.75***
	(155.29)	(170.34)	(12.26)
Classrooms per 100 students	1.17	0.80	0.37***
	(1.63)	(1.80)	(0.13)
Teachers per 100 students	3.04	3.62	-0.58**
	(1.41)	(12.79)	(0.28)
Textbooks per 100 students	99.28	102.33	-3.05
	(96.80)	(168.91)	(7.95)
Chairs per 100 students	20.58	14.13	6.45***
	(28.45)	(51.09)	(2.40)
Food from Gov or NGO	0.36	0.30	0.07*
	(0.48)	(0.46)	(0.04)
Solid building	0.36	0.28	0.08*
	(0.48)	(0.45)	(0.04)
Water Pump	0.61	0.45	0.17***
	(0.49)	(0.50)	(0.04)
Latrine/Toilet	0.85	0.71	0.14***
	(0.34)	(0.45)	(0.03)
Distance to MoE (in KM)	153.99	186.99	-33.00***
	(100.63)	(106.81)	(10.49)

This table presents the mean and standard error of the mean (in parentheses) for schools in the RCT (Column 1) and other public schools (Column 2), as well as the difference in means across both groups (Column 3). The sample of RCT schools is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

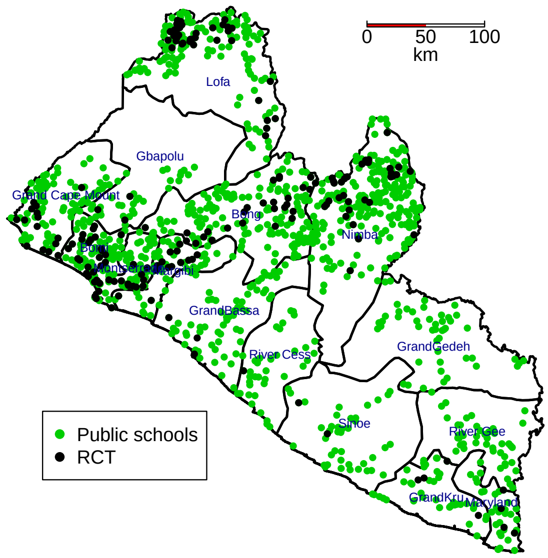

Figure 2: Public schools in Liberia and those within our sample.

Within the list of schools in the RCT, schools were paired within districts according to a PCA index of school quality[5]. A list of "pairs" was given to each operator based on their location preferences, so that each list had twice the number of schools they were assigned to operate. Once each operator "approved" of this list, we randomized the treatment assignment within each pair.[6]

Below, we present balance using our independently collected data. Table 2 shows the difference between treatment and control schools, according to administrative data from the EMIS.

[5]The index was calculated using the first eigenvector of a principal component analysis that included the following variables: students per teacher; students per classroom; students per chair; students per desk; students per bench; students per chalkboard; students per book; whether the school has a solid building; whether the school has piped water, a pump or a well; whether the school has a toilet; whether the school has a staff room; whether the school has a generator; and the number of enrolled students.

[6]There is one threesome due to logistical constraints in the assignment of schools across counties. Therefore, there is one extra treatment school.

Table 2: Balance in administrative (EMIS) data from pre-treatment year: Difference in characteristics between treatment and control schools

	Treatment	Control	Difference	Difference (F.E)
Students ECE	136.72	148.51	11.79	11.03
	(70.24)	(76.83)	(10.91)	(9.74)
Students Primary	143.96	159.05	15.10	15.68
	(86.57)	(163.34)	(19.19)	(16.12)
Students	277.71	305.97	28.26	27.56
	(124.98)	(178.49)	(22.64)	(19.46)
Classrooms per 100 students	1.13	1.21	0.09	0.08
	(1.65)	(1.62)	(0.24)	(0.23)
Teachers per 100 students	2.99	3.08	0.09	0.09
	(1.30)	(1.49)	(0.21)	(0.18)
Textbooks per 100 students	95.69	102.69	7.00	7.45
	(95.40)	(97.66)	(14.19)	(13.74)
Chairs per 100 students	22.70	18.74	-3.96	-4.12
	(32.81)	(23.06)	(4.17)	(3.82)
Food from Gov or NGO	0.36	0.36	-0.01	-0.01
	(0.48)	(0.48)	(0.08)	(0.05)
Solid building	0.33	0.39	0.06	0.06
	(0.47)	(0.49)	(0.07)	(0.06)
Water Pump	0.67	0.56	-0.11	-0.12*
	(0.47)	(0.50)	(0.07)	(0.06)
Latrine/Toilet	0.86	0.85	-0.01	-0.01
	(0.32)	(0.35)	(0.05)	(0.05)
Distance to MoE (in KM)	155.49	152.64	-2.85	-2.62
	(101.07)	(100.07)	(14.79)	(3.78)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2), as well as the difference between treatment and control (Column 3), and the difference taking into account the randomization design (i.e., including "pair" fixed effects (Column 4). The sample is the final treatment and control allocation. Source: EMIS data. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Due to errors in the EMIS data, not all schools originally assigned to treatment were able to be managed by PSL operators. After operators visited their assigned schools to start preparing for the upcoming school year, two treatment schools turned out to be private schools that were incorrectly labeled in the EMIS data as public schools. Two other schools had only two classrooms each. We treat these schools as non-compliant and present results in an ITT framework. This original sample with non-compliance is our main sample. However, we gave these operators new "pairs" of schools and informed them, as before, that they would operate one of these schools (but not which one). Operators approved of the list before they were given the actual assignment from randomization.

In an appendix we will present results for this "final" list of treatment and control schools. We expect the results of the final treatment and control school list and the original list to be almost identical, given that they only differ in four pairs of schools. Figure 3 shows the original treatment assignment.

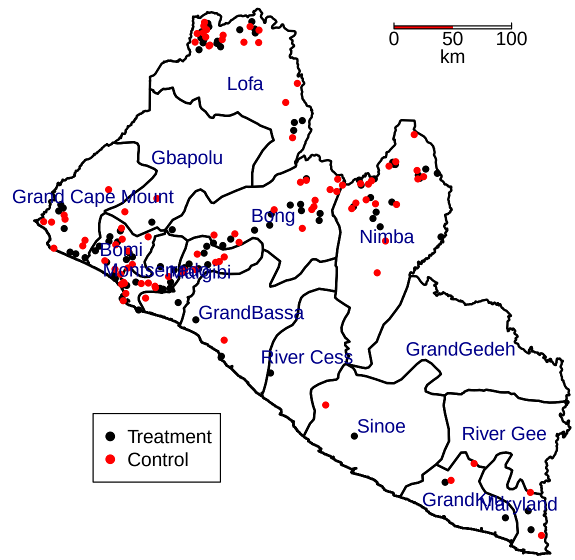

Figure 3: Geographical distribution of treatment and control schools. Original treatment assignment.

Appendix B contains a complete list of the schools included in the PSL program evaluation. The list denotes which schools are part of the original and final RCT lists and which ones ultimately became PSL schools as well as an identification number (groupID) that identifies pairs.

2.4 Data

Since the composition of students may change across PSL and control schools in response to treatment assignment, we sampled 20 students from each school's enrollment log of 2015/2016, the last year before the treatment was introduced. These samples are taken from students enrolled in school prior to any awareness of the PSL program. Each student will be evaluated as part of her/his "original" school, regardless of what school (if any) s/he attended in subsequent years.

We surveyed all of the teachers in each school and conducted in-depth surveys with those teaching math and English. We asked teachers about their time use and teaching strategies. We also obtained teacher opinions on the PSL program. For a randomly selected class within each school, we conducted a classroom observation using the Stallings Classroom Observation tool. Furthermore, we conducted school-level surveys to collect information about school facilities, the teacher roster, input availability and expenditures.

At midline, we survey a sample of households from our student sample. Additionally, we gather data on school enrollment and learning levels for all children 4-8 years old living in these households.

The baseline survey was conducted in September 2016, to be followed by a midline survey in May 2017 and an endline survey in May 2019. See Figure 4 for a timeline of intervention and research activities.

Figure 4:

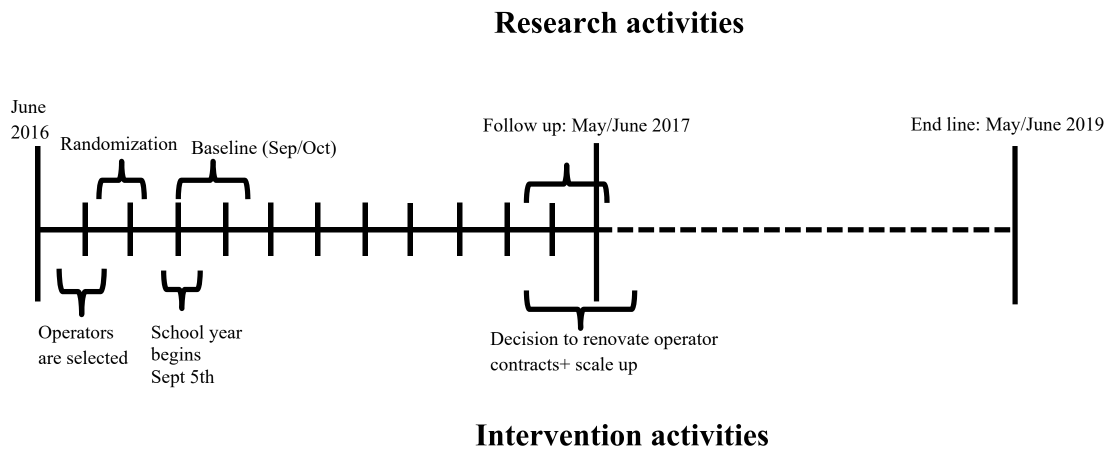

Table 3 shows that observable, time-invariant characteristics of students and schools are balanced across treatment and control. Eighty percent of schools in our sample are in rural areas, over an hour away from the nearest bank (usually located in the nearest urban center); over 10% need to hold some classes outside due to insufficient classrooms. Nearly 55% of our students are boys, and they have an average age of 12.

Table 3: Balance in baseline survey data in observable, time invariant, school and student characteristics

Panel A: School characteristics				
	Control	Treatment	Difference	Difference (F.E)
Facilities (PCA)	0.021	-0.073	-0.094	-0.082
	(1.653)	(1.586)	(0.238)	(0.235)
Hold some classes outside	0.141	0.140	-0.002	0.000
	(0.350)	(0.349)	(0.051)	(0.051)
Rural	0.804	0.796	-0.009	-0.004
	(0.399)	(0.405)	(0.059)	(0.047)
Time to nearest bank	68.043	75.129	7.086	7.079
	(60.509)	(69.099)	(9.547)	(8.774)
Panel B: Student characteristics				
	Control	Treatment	Difference	Difference (F.E)
Age	12.327	12.414	0.087	0.039
	(2.942)	(2.842)	(0.171)	(0.117)
Male	0.562	0.550	-0.012	-0.017
	(0.496)	(0.498)	(0.020)	(0.013)
Wealth Index	-0.028	-0.046	-0.018	0.001
	(1.492)	(1.483)	(0.133)	(0.056)
ECE before grade 1	0.820	0.834	0.014	0.016
	(0.384)	(0.373)	(0.025)	(0.016)

Baseline data was collected 2 to 8 weeks after the beginning of treatment; hence we focus here on immutable characteristics.
This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2), as well as the difference between treatment and control (Column 3), and the difference taking into account the randomization design (i.e., including "pair" fixed effects (Column 4). The school infrastructure index is the first component from a Principal Component Analysis of indicator variables for: classrooms, staff room, student and adult latrines, library, playground, and an improved water source. Standard errors are clustered at the school level. The sample is the final treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3 Hypotheses

In this section we present the hypotheses to be tested with our data, as well as the variables used to construct the outcome of interest. Note that in some cases we also present a conjecture for whether the program has a negative or positive impact. In these cases we intend to conduct one-sided tests.

There are at least three important dimensions of heterogeneity we want to highlight before describing the hypotheses to be tested. The first is enrollment in 2015/2016. Given that operators are allowed (but not required) to cap class sizes, this could lead to students being excluded from their previous school (and either transferred to another school or to no school at all). This "selection" effect is an important mechanism. Since the constraints may be grade-school specific, for each student we will estimate how constrained the school is as the number of students enrolled in the student's 2016/2017 "expected grade" (as given by normal progression based on its 2015/2016 grade) and adjacent grades, divided by the "maximum capacity" in those three grades in 2016/2017. In short,

$$c_{igso} = \frac{Enrollment_{is,g-1} + Enrollment_{is,g} + Enrollment_{is,g+1}}{3 * Maximum_o}$$

where c_{igso} is our constrained measure for student i, expected to be in grade g in 2016/2017, at schools s in a "pair" assigned to operator o. $Enrollment_{is,g-1}$ is the enrollment in a grade below the student's expected grade, $Enrollment_{is,g}$ is enrollment in the student's expected grade, and $Enrollment_{is,g+1}$ is enrollment in grade above the student's expected grade. $Maximum_o$ is the cap size approved for operator o. We include the "adjacent grades" since many operators have placement exams and therefore students are likely to be placed in the grade above or below the "expected grade".

The second important dimension of heterogeneity is school competition. Internationally, the charter school movement is closely tied to policy reforms bestowing parents with freedom of school choice. The standard argument is that charter schools will be more reactive to parents' demands than traditional public schools because their funding is linked directly to enrollment numbers. However, there is limited empirical evidence that parents' choices respond to learning quality in low-income settings (Andrabi, Das, & Khwaja, 2008) and this mechanism may be more relevant for schools in high-density urban locations like Monrovia than remote rural areas where choice is *de facto* limited to one or two schools. We will use the information of every school in the country to measure school competition.

Note that we do not have experimental variation in school competition. Since school competition is highly correlated with population density, road access, and other observable school characteristics, these estimates may be biased. We will define school competition as the number of schools within a 5 km radius of the school (and check for robustness using 2.5 km and 10 km radius).

Finally, we will examine heterogeneity by baseline student test scores. To do this, we will estimate both parametric and non-parametric treatment effects by baseline student test scores.

Below are the hypotheses to be tested at the student, teacher, school and household level. **Bolded** hypotheses are our main outcomes of interest (access to schooling and learning outcomes), while *italics* denote mechanisms for the main outcomes.

Table 4: Hypotheses to be tested: Student outcomes

Outcomes	Question[b]	Hypothesis	Heterogeneity[d]
English test scores[a]	**SS:** reading_test group	+	c,s,o
Math test scores[a]	**SS:** math_test group	+	c,s,o
Performance in new modules[a]	**SS:** daynight, digitspangame, more_multiplication, more_addition, pronoun1-pronoun3, letters_dicttest1-letters_dicttes3, word_dict_test1-word_dict_test3, num_dict_test1-num_dict_test3	?	
Executive function[a]	**SS:** daynight digitspangame	+	
Abstract thinking[a]	**SS:** raven_test1-raven_test5	+	
Student satisfaction	**SS:** opinion1	?	
Aspirations	**SS:** opinion2-opinion4, opinion11 **HH:** e.8_edu_level	?	
Attendance	**SS:** misssch, **PP:** nu_headcount/C2.nu_calc-grade6_headcount/C2.6_calc	?	c,s,o
Socialization	**SS:** opinion6-opinion10, voter_register	?	o

[a] Student "test scores" we will estimate using IRT models based on all questions within the group of questions mentioned above.

[b] Question names as shown in the attached survey instruments. SS refers to the student survey, HH refers to the household survey, PP refers to the principal survey, TT refers to the teacher survey, CO refers to the classroom observations.

[d] Heterogeneity key. w: Wealth - student asset index; s: school competition, rurality; o: Operator - for-profit vs nonprofit, cost, c: constrained at baseline

Table 4 shows the outcomes of interest for students. English and math test scores, executive function scores and Raven's matrices scores (abstract thinking) are all measures of student learning levels and cognitive ability. We plan to test learning outcomes and selection on two samples of students: First, the 20 students per school we sampled from the 2015/2016 enrollment log. This is our main sample and is the core of our ITT estimates. Second, from a sample of children between 4 and 8 years old that allows us to get at any effect in younger children and children who were not previously enrolled in school.

Early childhood education (ECE)–Nursery, K1, and K2–students have to pay fees in traditional public schools, but not in PSL schools. As a result, we expect enrollment to increase dramatically in ECE grades in PSL schools. Since we do not have students enrolled in Nursery in 2016/2017 in our main sample[7], we test children 4-8 years old living in the same household as our sampled students. By doing this we still recover a clean ITT effect of the PSL program on younger students. However, the sample may not be representative of 4-8 year olds in communities near PSL schools.

Student satisfaction, aspirations, and socialization will only be tested in the main sample of

[7] Our data contains students sampled from K1-Grade 5 in 2015/2016. Most students in in Grade 6 would no longer be in primary school in 2016/2017.. Most students enrolled in Nursery School in 2016/2017 are not enrolled in school in 2015/2016.

students.

Table 5: Hypotheses to be tested: Teacher outcomes

Outcomes	Question[c]	Hypothesis	Heterogeneity[d]
Time on task[a]	**TT:** A.1e, **SS:** math_teacher_leave, english_teacher_leave, english_teacher_help, math_teacher_help, **CO:** snap_tch_activity, snap_tch_materials, snap_tch_code	+	o,s
Absenteeism	**TT:** A.1d, **SS:** math_teacher_abs, english_teacher_abs	-	o,s
Teacher satisfaction[b]	**TT:** D.2a_T-D.2i_T, D.1a_T, D.1_T, D.2_T	+	o,s
Use of corporal punishment	**SS:** math_teacher_hit, english_teacher_hit	-	c,o,s

[a] Classroom observations will be coded according to the "Stallings snapshot observation manual. January 2007. Modified for use in The World Bank projects" published by Texas A&M University.
[b] An index based on the first eigenvector of a PCA will be used as the outcome variable.
[c] Question names as shown in the attached survey instruments. SS refers to the student survey, HH refers to the household survey, PP refers to the principal survey, TT refers to the teacher survey, CO refers to the classroom observations.
[d] Heterogeneity key. w: Wealth - student asset index; s: school competition, rurality; o: Operator - for-profit vs nonprofit, cost, c: constrained at baseline

Table 4 provides the outcomes of interest for teachers. Note that these refer to changes in teacher behavior and since the treatment impacts the composition of the pool of teachers, any variation in teacher behavior between treatment and control schools can be considered a composite of effects on the extensive and the intensive margin. We will employ non-experimental techniques to tease out the intensive-margin treatment effect on teacher behavior due to the program.

Table 6: Hypotheses to be tested: School outcomes

Outcomes	Question[b]	Hypothesis	Heterogeneity[c]
Enrollment	**PP:** C1.*, C2.*	?	s,o,c
Selection	**SS:** school_current, school_current2, dateenroll	?	c,s,o,w
Pupil-teacher ratio	**PP:** C2.*/E.1a	?	s,o
Teacher rotation (hiring & firing)	**PP:** TeachersGone, **TT:** A.7a	+	s,o
Teacher characteristics	**TT:** Demographic: A.3, A.4, A.6b, A.9, A.10a. Memory:[a] G.1A_T-G.1E_T, Personality (Big 5): H.1A_T-H.1J_T, Ravens:[a] I2.a-I2.e, English skills:[a] J.1_T-J.5_T, Math Skills:[a] J.6a_T, J.7a_T	+	s,o
Input availability for students[a]	**CO:** total_students/numlongdesk, total_students/numshordest, total_students/numofarmchair, numstwithtextbk/total_students, numstwithpen/total_students	+	s,o
Hours of instructional time	**PP:** G.3A_a-G.3A_e	+	
School management[a]	**PP:** B.2a-B.2g, C2.12, neat, R.1, R.1a, F.1a, G.2, I.4b	+	s,o
PTA strength[a]	**PP:** I.2a-I.2f I.5, I.6	?	
Crowding out/in of resources	M.1a-M.4e_11c	+	s,o
Fees	**PP:** N.0a-N.1b_66c, **HH:** a_sch_fees	?	

[a] An index based on the first eigenvector of a PCA will be used as the outcome variable.
[b] Question names as shown in the attached survey instruments. SS refers to the student survey, HH refers to the household survey, PP refers to the principal survey, TT refers to the teacher survey, CO refers to the classroom observations.
[c] Heterogeneity key. w: Wealth - student asset index; s: school competition, rurality; o: Operator - for-profit vs nonprofit, cost, c: constrained at baseline

Table 6 shows the outcomes of interest for schools; several of these outcomes rely on administrative data. For example, teacher rotation will rely heavily on the EMIS data to study the evolution of teachers who work in schools. For enrollment, although we are interested in seeing whether there are differences in enrollment changes between treatment and control schools, we also want to study whether these are net gains in enrollment (enrollment of students previously out of school), or if the increases in enrollment are due to students changing schools. To do this, we will rely on the EMIS data to study total enrollment numbers in the catchment area for each school.

Table 7: Hypotheses to be tested: Household outcomes

Outcomes	Question[b]	Hypothesis	Heterogeneity[c]
Education expenditure	**HH:** expenditure_group questions, donate_money questions	?	
Parental engagement[a]	**SS:** opinion5 **HH:** e.1a_meeting, f.1_homework, f.2_breakfast, f.5_report	?	
Parental satisfaction	**HH:** i.0_satisfied	+	c,s,o

[a] An index based on the first eigenvector of a PCA will be used as the outcome variable
[b] Question names as shown in the attached survey instruments. SS refers to the student survey, HH refers to the household survey, PP refers to the principal survey, TT refers to the teacher survey, CO refers to the classroom observations.
[c] Heterogeneity key. w: Wealth - student asset index; s: school competition, rurality; o: Operator - for-profit vs nonprofit, cost, c: constrained at baseline

Table 7 provides the outcomes of interest for households. Note that children between ages 4-8 will be tested in these households, and used as an additional sample to measure student outcomes for ECE children. We gather data on enrollment status (2016/2017), previous enrollment status (2015/2016), and date on which each household member moves into the home. We will use this data to study how PSL affects enrollment decisions for all household members and whether there is any sorting in the form of child migration across family members (e.g., cousins moving to a household near a PSL school).

4 Specification

All regressions use sample weights to recover statistics for the average student. For the analysis on the full sample and for sub-group analyses where we have 40 or more schools, we will cluster the standard errors at the school level. For sub-group analysis where we have fewer than 40 schools per group, we will collapse individual results to the mean at the school level before performing the analysis.

For the hypotheses that relate to individual, pupil-level outcomes, we will report treatment-effect estimates based on three specifications. The first specification amounts to a simple comparison of post-treatment outcomes for treatment and control individuals, where Y_{isg} is any one of the aforementioned outcomes for student i in school s and group g (denoting the matched pairs used for randomization); α_g is a matched-pair fixed effect; $treat_{sg}$ is an indicator for whether school s was randomly chosen for treatment; and ε_{isg} is an individual error term.

$$Y_{isg} = \alpha_g + \beta_1 treat_{sg} + \varepsilon_{isg} \tag{1}$$
$$Y_{isg} = \alpha_g + \beta_2 treat_{sg} + \varepsilon_{isg} + \gamma_2 X_{isg} + \delta_2 Z_{sg} \tag{2}$$
$$Y_{isg} = \alpha_g + \beta_3 treat_{sg} + \varepsilon_{isg} + \gamma_3 X_{isg} + \delta_3 Z_{sg} + \zeta_3 Y_{isg,-1} \tag{3}$$

The second specification adds controls for pupil demographic and other pre-determined baseline characteristics measured at the individual level (X_{isg}) and school level (Z_{sg}). These controls are enumerated in Table 9. In the third line, equation (3), we use an ANCOVA specification that controls for baseline individual outcomes.

Adding controls, as in equation (2), should increase the precision of our results. However, controlling for baseline outcomes, as in equation (3), may also risk attenuation bias in the treatment effect estimates if the baseline outcomes are imbalanced. This is, in fact, what we observe in our baseline data, as shown in Table 8. For both English and math scores, students in treatment schools score higher than control schools (significant at the 5% level), by roughly 0.07 standard deviations in each case.

There is some evidence that this imbalance is not simply due to "chance bias" in randomization, but rather a treatment effect that materialized in the weeks between the beginning of the school year and the application of the baseline survey. First, there is no significant effect on abstract reasoning, which is arguably less amenable to short-term improvements through teaching (although the difference between a significant English/math effect and an insignificant abstract reasoning effect here is not itself significant). Second, the effects on English and math appear to materialize in the later weeks of the fieldwork, as shown in Figure 5, consistent with a treatment effect rather than imbalance.

Thus we may face a trade-off between precision and attenuation bias in choosing between the three specifications above. Our preferred specificaiton is equation (2), though we will report all three results.

Table 8: Students

	Control	Treatment	Difference	Difference (F.E)
IRT English score	-0.000	0.048	0.048	0.066**
	(1.000)	(1.027)	(0.081)	(0.033)
IRT Math score	0.000	0.080	0.080	0.067**
	(1.000)	(1.021)	(0.067)	(0.031)
IRT Abstract score	0.000	0.054	0.054	0.045
	(1.000)	(0.973)	(0.060)	(0.038)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2), as well as the difference between treatment and control (Column 3), and the difference taking into account the randomization design–i.e., including "pair" fixed effects (Column 4).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

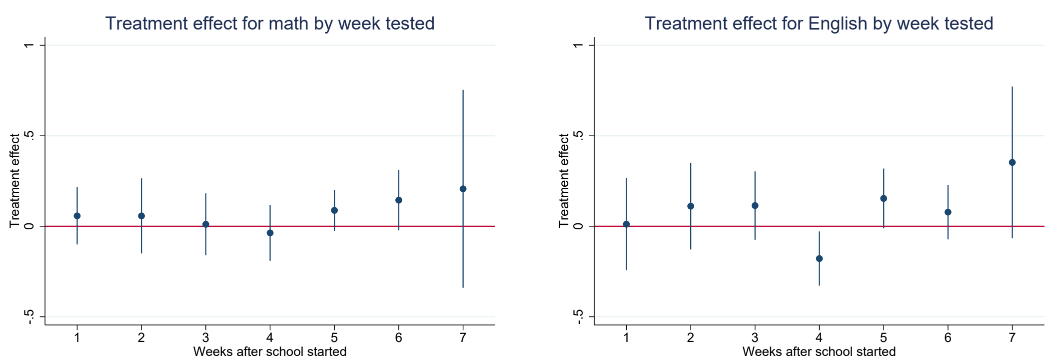

Figure 5: The panel on the left shows results for math test scores, while the panel on the right shows English test score results.

Table 9: Control Variables

Student Controls	Question	Questionnaire	
Wealth index	A1-A7	Student	Baseline
Age	B1	Student	Baseline
Gender	B2	Student	Baseline
Grade 2015/2016	B6a	Student	Baseline
School Controls			
Enrollment last year	C1	Principal	Baseline
Infrastructure quality from last year	L1-L3	Principal	Baseline
Distance to nearest bank	L6	Principal	Baseline
Rurality	L7	Principal	Baseline
NGO programs in 2015/2016	M1-M4	Principal	Baseline
Donations in 2015/2016	N1A-N3b_a_5	Principal	Baseline
Household Controls			
Home language	E1	Student	Baseline
ECE attendance	E2	Student	Baseline
Asset index - Student	E3-E11	Student	Baseline
HH size and composition	hh_number	Household	
Parent education	hh_member_education,hh_member_grade	Household	
Parent employment	b.8a, b.8_occupation, b.8_employment	Household	
Asset index - Household	c.8a_hh_asset-c.8g_hh_asset	Household	
Parent cognitive level	h.1_eng_reading-h.3_math_result2	Household	

4.1 On comparisons across operators

Is important to note that the assignment of operators to schools was not random. Different operators stated different preferences for locations and some volunteered to manage schools in more remote and marginalized areas. Thus, any heterogeneous effects by operator or by operator's characteristics are not experimental. Figure 10 shows the treatment and control schools allocated to each operator. Table 6 shows the difference in school (both treatment and control) characteristics across operators.

Ultimately, the RCT does not allow us to get comparable estimates across operators, and is underpowered at the operator level. We will estimate heterogeneous treatment effects by operator characteristics (for-profit/non-profit). To mitigate the possible bias due to differences in location and school characteristics we will include a comprehensive set of school controls, as well as an interaction of those controls with a treatment dummy to make sure we capture heterogeneous treatment effects that go beyond any differences in location/schools.

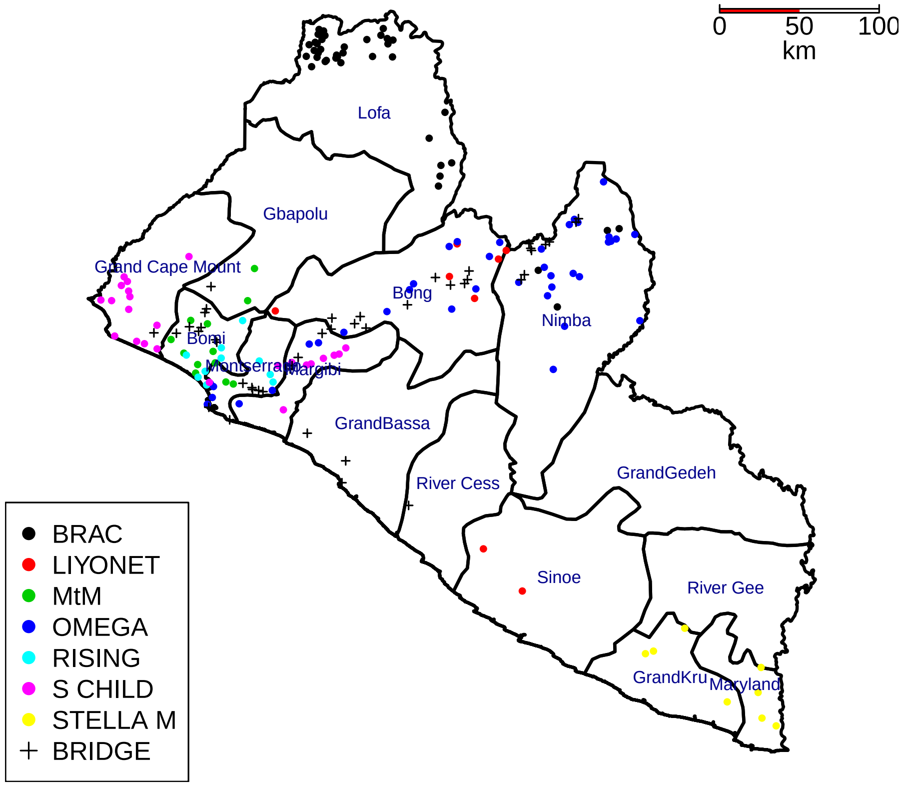

Figure 6: Geographical distribution of operators across the country.

Table 10: Pre-treatment EMIS characteristics of treatment schools by operator

	BRAC Non-profit	BRIDGE For-profit	LIYONET Non-profit	MtM Non-profit	OMEGA For-profit	RISING For-profit	SCHILD Non-profit	STELLAM Non-profit	Total
Students ECE	126.1	178.5	115.4	106.8	158.4	123.7	154.9	115.2	146.9
	(12.18)	(18.27)	(21.66)	(11.04)	(9.546)	(18.21)	(11.62)	(13.80)	(6.036)
Students Primary	152.2	225.1	110.4	140.3	115.1	120	109.4	99	148.3
	(11.72)	(35.58)	(20.35)	(43.47)	(7.958)	(14.47)	(7.575)	(16.13)	(9.679)
Students	278.3	403.6	225.9	247.1	273.5	243.7	264.2	214.2	295.2
	(19.59)	(39.60)	(32.47)	(46.23)	(13.21)	(26.78)	(14.53)	(29.01)	(11.97)
Classrooms per 100 students	0.974	1.276	1.451	2.164	0.561	1.899	1.108	0	1.074
	(0.262)	(0.195)	(0.659)	(0.946)	(0.204)	(0.661)	(0.327)	(0)	(0.123)
Teachers per 100 students	2.965	2.492	3.165	3.953	3.167	3.553	2.760	3.208	2.981
	(0.189)	(0.169)	(0.452)	(1.107)	(0.183)	(0.622)	(0.263)	(0.288)	(0.109)
Textbooks per 100 students	139.1	75.74	75.67	58.67	96.39	120.8	83.64	68.20	96.63
	(16.65)	(11.50)	(24.30)	(23.96)	(22.27)	(42.49)	(19.15)	(15.53)	(7.900)
Chairs per 100 students	6.188	25.42	41.69	38.68	15.56	34.82	23.20	15.49	20.33
	(2.226)	(3.301)	(16.75)	(11.89)	(2.945)	(9.860)	(7.275)	(11.59)	(2.040)
Food from Gov or NGO	0.0286	0.389	0	0.667	0.314	0.778	0.636	0.667	0.358
	(0.0286)	(0.0824)	(0)	(0.167)	(0.0796)	(0.147)	(0.105)	(0.211)	(0.0382)
Solid building	0.257	0.611	0.714	0.333	0.143	0.667	0.409	0	0.371
	(0.0750)	(0.0824)	(0.184)	(0.167)	(0.0600)	(0.167)	(0.107)	(0)	(0.0384)
Water Pump	0.314	0.639	0.714	0.556	0.714	0.889	0.727	0.833	0.616
	(0.0796)	(0.0812)	(0.184)	(0.176)	(0.0775)	(0.111)	(0.0972)	(0.167)	(0.0387)
Latrine/Toilet	0.784	0.870	0.857	0.807	0.881	0.889	0.909	0.932	0.858
	(0.0666)	(0.0552)	(0.143)	(0.130)	(0.0520)	(0.0774)	(0.0627)	(0.0683)	(0.0261)
Distance to MoE (in KM)	239.7	111.2	180.2	35.07	180.2	35.00	75.80	379.1	154.3
	(2.753)	(13.11)	(19.03)	(6.860)	(15.88)	(4.506)	(4.438)	(11.26)	(7.990)

This table presents the mean and standard error of the mean (in parenthesis) for several school characteristics across operators. The sample is the final treatment and control allocation. Source: EMIS data.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 On missing data

After the baseline had concluded we realized the sampling protocol (by which students were to be sampled from the 2015/2016 enrollment log) had not been followed at two schools: 110142 and 20284. Therefore, for these schools we sampled and tested a new batch of students from the 2015/2016 enrollment log during midline. For these students we do not have baseline data.

Since we could not verify this was not an issue in other schools without physically visiting each one, we created a protocol at midline to detect any other sampling errors. Enumerators had to verify that the students sampled at baseline were in the 2015/2016 enrollment log. If more than 10% of the students sampled at baseline were not found in the 2015/16 log (suggesting that sampling may have been erroneously done from the 2016/17 logs), then a new batch of pupils from the 2015/2016 enrollment log was sampled and tested at midline.

To include all of our observations in the regressions and include student level controls, we will replace any missing values for the controls with zeroes. To account for this in our regression specification we will include an indicator for whether the baseline data is missing and interact this indicator with all the baseline controls. In short, we will estimate the following modified version of the three specifications presented in equations (4)-(6):

$$
\begin{aligned}
Y_{isg} &= \alpha_g + \beta_4 treat_{sg} + \varepsilon_{isg} && (4)\\
Y_{isg} &= \alpha_g + \beta_5 treat_{sg} + \varepsilon_{isg} + \gamma_5 X_{isg} \times Base_{isg} + \beta_5 Base_{isg} + \delta_5 Z_{sg} && (5)\\
Y_{isg} &= \alpha_g + \beta_3 treat_{sg} + \varepsilon_{isg} + \gamma_6 X_{isg} \times Base_{isg} + \beta_2 Base_{isg} + \delta_6 Z_{sg} && (6)\\
&\quad + \zeta_6 Y_{isg,-1} \times Base_{isg}
\end{aligned}
$$

where $Base_s$ is equal to 1 if the data for student i in school s is not missing.

4.3 On multiple testing

Since we plan to examine various outcomes and dimensions of heterogeneity, we must correct for multiple comparisons. We will present adjusted p-values using a step-down procedure (e.g., See Romano and Wolf (2005), Romano and Wolf (2010) or List, Shaikh, and Xu (2016)) that controls the family-wise error rate (i.e., the probability of one or more false rejections).

There are two types of multiple hypothesis testing taking place in our setting. First, when testing the effect of treatment on several outcomes we will adjust p-values using pre-specified outcome groups. The adjusted p-values will control the family-wise error within those groups.[8]

[8] Note that the "correct" procedure is doing a full family-wise correction of all tested outcomes, which would dramatically undermine the power of the experiment.

Table 11: Outcome grouping to adjust p-values

Outcome	Group
Student Outcomes	
English	No group
Math	No group
Abstract Thinking	No group
Executive function	No group
Attendance	Other student outcomes
Student satisfaction	Other student outcomes
Aspirations	Other student outcomes
Socialization	Other student outcomes
Teacher Outcomes	
Time on task	Teacher mechanisms
Absenteeism	Teacher mechanisms
Use of corporal punishment	Teacher mechanisms
Teacher satisfaction	No group
School Outcomes	
Enrollment	No group
Selection	No group
Pupil-teacher ratio	School mechanisms
Teacher rotation (hiring & firing)	School mechanisms
Teacher characteristics	School mechanisms
Input availability for students	School mechanisms
Hours of instructional time	School mechanisms
School management	School mechanisms
PTA strength	School mechanisms
Crowding out/in of resources	School mechanisms
Fees	No group
Household Outcomes	
Education expenditure	Household mechanisms
Parental engagement	Household mechanisms
Parental satisfaction	No group

The second instance is testing for heterogeneous treatment effects. In these cases, we will present adjusted p-values that control the family-wise error rate of testing the same outcome in multiple subgroups.

Appendices

A Civic attitudes

A related question, ancillary to whether PSL improves student outcomes, is whether the program impacts the civic and political attitudes of those affected by it, including support for the incumbent government and support for private provision of public goods. This question is especially timely as presidental and legislative elections will be held in October 2017, shortly after our results are slated to be made public.

Arguments about how political competition leads to growth often depend upon the fact that democracy allows voters to hold governments accountable for good policy ((Stigler, 1972; Besley, Persson, & Sturm, 2010; Acemoglu, Naidu, Restrepo, & Robinson, 2014)). But it's not clear how direct this accountability is, nor what constitutes good policy in the minds of voters. A wealth of evidence shows that voters reward politicians for redistribution which benefits them directly, in developing countries (Manacorda, Miguel, & Vigorito, 2011; Pop-Eleches & Pop-Eleches, n.d.; Bursztyn, 2016) and developed countries (Levitt & Snyder, 1997). In clientelistic democracies, rents captured by those in public office are shared with supporters in exchange for political backing, to the detriment of investment in public goods which produce more broad-based growth (Wantchekon, 2003; Kramon, Posner, et al., 2016). The immediate, targeted benefits voters gain from a patronage-based political environment may often be more salient than the benefits of better public goods provision, which are often diffuse, difficult to measure, and realized over time. There is scant evidence that voters reward any given political entity for investing in good policy as a growth-producing public good (a notable exception is Larreguy, Marshall, and Trucco (2015), which finds that voters may reward federal politicians for good policy even if it hampers the effectiveness of clientelism at the local level).

One reason for this dearth of evidence is that government policies are rarely subjected to rigorous evaluation, making it difficult to know what constitutes "good" policy. Another reason is that policy often aims to maximize parties' perceived incumbency advantage, e.g. by redistributing based on factors such as co-ethnicity or party affinity (Kramon et al., 2016). Exogenous policy shocks delivered at an electorally meaningful geographic level, and at an electorally salient time, are relatively rare. PSL overcomes both problems by commissioning an independent third-party randomized evaluation of the program to measure impacts on student learning.

We aim to measure the effect of PSL on the civic and political attitudes of voting-age household members whose children are in our student sample, as well as on the attitudes of teachers. Our outcomes of interest in this section fall into four main categories:

- **Political Behavior:** While self-reported, these questions measure more concrete political actions compared to the attitudes we measure with other questions in this section. Did respondents register to vote in the upcoming election? (Voter registration was only possible from 1 February to 7 March 2017). Do respondents plan to vote in the October 2017 presidential and legislative elections?

- **Incumbent Support:** This is a broad category, encompassing the following: How do respondents perceive and rate the performance of various levels of government in selected areas? More broadly, do they perceive the country as moving forward? Do respondents perceive their own ethnic group to suffer unfair treatment at the hands of the government?

- **Policy Priorities:** Which areas do respondents think the government should spend more money on?

A.1 Hypotheses

Table 12: Hypotheses to be tested: Civic and Political Attitudes (Teachers and Household)

Outcomes	Question, HH[a]	Question, TT	Heterogeneity[b]
Political Behavior	**HH:** Voter registration (J.3a), Plans to vote (J.3c)	I.10, I.10a	
Incumbent Support	**HH:** Government performance (J.2b-d, J.6a-e); Party support (J.9b,J.10a)	I.9b-d, I.12b-3, extra2, extra3 + Union support (I.9a)	
Private provision of public goods	**HH:** Support for (J.1b,c); Perception of (J.7a,b); Discrimination (J.8b)	Support only (I.8b-c)	
Policy priorities	**HH:** J.5b	I.11	c,s,o

[a] Question names as shown in the attached survey instrument. HH refers to the household survey, TT refers to the teacher survey.

[b] Heterogeneity key. w: Wealth - household asset index; i: informed or interested voters; o: Operator - for-profit vs nonprofit, cost, c: constrained at baseline

Table 12 lists the hypotheses we plan to test, with separate columns for which questions appear in the Household and Teacher surveys, respectively. The political questions we included in the teacher survey constitute a subset of those we included in the household survey (plus one extra question about attitudes toward teacher unions). We acknowledge that because teacher composition is an outcome of the program, treatment effects on teachers' attitudes are a composite of this selection effect and the effect on teaches' attitudes conditional on being in a given school.

We are interested in some of the same dimensions of heterogeneity here as in the learning outcomes above, plus one extra dimension (enumerated in the footnote of table 12. Schools' class size constraints could play an important role, as parents might view the program differently based on whether or not they perceive that their child was excluded from school because of it. Reactions to for-profit vs nonprofit school operators could similarly provoke divergent parent responses. The heterogeneity we add in this section is for how informed respondents are, based on whether they can name either of the two current senators from their county, and whether they say that they ever get news about politics. While these measures of respondents' political informedness are crude, we think this heterogeneity is important given low average levels of policy awareness in many parts of the country.

A.2 Specification

Our specification here follows the same broad pattern as the learning outcomes analysis. We follow the same convention as above for clustering: for the analysis on the full sample and for sub-group analyses where we have 40 or more schools, we will cluster the standard errors at the school level. For sub-group analysis where we have fewer than 40 schools per group, we will collapse individual results to the mean at the school level before performing the analysis.

$$Y_{isg} = \alpha_g + \beta_7 treat_{sg} + \varepsilon_{isg} \quad (7)$$
$$Y_{isg} = \alpha_g + \beta_8 treat_{sg} + \varepsilon_{isg} + \gamma_2 X_{isg} + \delta_2 Z_{sg} \quad (8)$$
$$(9)$$

Equation 7 is the simple comparison of means across treatment and control groups. Equation 8 adds controls at the school level and the individual level (here, the household controls), both listed in Table 9 above. Y_{isg} is any one of the aforementioned outcomes for the adult respondent from the household of student i in school s and group g (denoting the matched pairs used for randomization); α_g is a matched-pair fixed effect; $treat_{sg}$ is an indicator for whether school s was randomly chosen for treatment; and ε_{isg} is an individual error term. X_{isg} and (Z_{sg}) are individual- and school-level controls, respectively.

References

Abdulkadiroglu, A., Angrist, J., Cohodes, S., Dynarski, S., Fullerton, J., Kane, T., & Pathak, P. (2009). Informing the debate: Comparing boston's charter, pilot and traditional schools.

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2014). *Democracy does cause growth* (Tech. Rep.). National Bureau of Economic Research.

Allcott, H., & Ortega, D. E. (2009). The performance of decentralized school systems: Evidence from fe y alegría in república bolivariana de venezuela. *Emerging Evidence on Vouchers and Faith-Based Providers in Education*, 81.

Andrabi, T., Das, J., & Khwaja, A. I. (2008). A dime a day: The possibilities and limits of private schooling in pakistan. *Comparative Education Review*, *52*(3), 329–355.

Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. R. (2016). Stand and deliver: Effects of boston's charter high schools on college preparation, entry, and choice. *Journal of Labor Economics*, *34*(2), 275-318. Retrieved from `http://dx.doi.org/10.1086/683665` doi: 10.1086/683665

Barrera-Osorio, F. (2007). The impact of private provision of public education: empirical evidence from bogota's concession schools.

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L., Raju, D., & Rya, S. (2013). Leveraging the private sector to improve primary school enrolment: Evidence from a randomized controlled trial in pakistan.

Barrera-Osorio, F., Galbert, D., Gaspard, P., Habyarimana, J. P., & Sabarwal, S. (2016). Impact of public-private partnerships on private school performance: evidence from a randomized controlled trial in uganda.

Besley, T., Persson, T., & Sturm, D. M. (2010, oct). Political Competition, Policy and Growth: Theory and Evidence from the US. *Review of Economic Studies*, *77*(4), 1329–1352. Retrieved from `http://restud.oxfordjournals.org/lookup/doi/10.1111/j.1467-937X.2010.00606.x` doi: 10.1111/j.1467-937X.2010.00606.x

Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. *Economics of Education Review*, *24*(2), 133 - 147. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0272775704000779` doi: http://dx.doi.org/10.1016/j.econedurev.2004.04.009

Betts, J. R., & Tang, Y. E. (2016). A meta-analysis of the literature on the effect of charter schools on student achievement. *Society for Research on Educational Effectiveness*.

Bifulco, R., & Ladd, H. F. (2006). The impacts of charter schools on student achievement: Evidence from north carolina. *Education*, *1*(1), 50–90.

Bonilla, J. D. (2010). Contracting out public schools for academic achievement: Evidence from colombia.

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D. (2007). The impact of charter school attendance on student performance. *Journal of Public Economics*, *91*(5-6), 849 - 876. Retrieved from http://www.sciencedirect.com/science/article/pii/S0047272706001393 doi: http://dx.doi.org/10.1016/j.jpubeco.2006.09.011

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D. (2008). The effect of charter schools on traditional public school students in texas: Are children who stay behind left behind? *Journal of Urban Economics*, *64*(1), 123 - 145. Retrieved from http://www.sciencedirect.com/science/article/pii/S0094119007001210 doi: http://dx.doi.org/10.1016/j.jue.2007.10.003

Bursztyn, L. (2016, oct). Poverty and the Political Economy of Public Education Spending: Evidence From Brazil. *Journal of the European Economic Association*, *14*(5), 1101–1128. Retrieved from http://doi.wiley.com/10.1111/jeea.12174 doi: 10.1111/jeea.12174

Chabrier, J., Cohodes, S., & Oreopoulos, P. (2016, August). What can we learn from charter school lotteries? *Journal of Economic Perspectives*, *30*(3), 57-84. Retrieved from http://www.aeaweb.org/articles?id=10.1257/jep.30.3.57 doi: 10.1257/jep.30.3.57

Clark, M. A., Gleason, P. M., Tuttle, C. C., & Silverberg, M. K. (2015). Do charter schools improve student achievement? *Educational Evaluation and Policy Analysis*, *37*(4), 419-436. Retrieved from http://dx.doi.org/10.3102/0162373714558292 doi: 10.3102/0162373714558292

Gleason, P., Clark, M., Tuttle, C. C., & Dwoyer, E. (2010). The evaluation of charter school impacts: Final report. ncee 2010-4029. *National Center for Education Evaluation and Regional Assistance*.

Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. *Journal of Public Economics*, *91*(5-6), 823 - 848. Retrieved from http://www.sciencedirect.com/science/article/pii/S0047272706001496 doi: http://dx.doi.org/10.1016/j.jpubeco.2006.09.014

Hoxby, C. M., Murarka, S., & Kang, J. (2009). How new york city's charter schools affect achievement.

Hoxby, C. M., & Rockoff, J. E. (2004). *The impact of charter schools on student achievement.*

Imberman, S. A. (2011). The effect of charter schools on achievement and behavior of public school students. *Journal of Public Economics*, *95*(7-8), 850 - 863. Retrieved from http://www.sciencedirect.com/science/article/pii/S0047272711000193 doi: http://dx.doi.org/10.1016/j.jpubeco.2011.02.003

Kramon, E., Posner, D. N., et al. (2016). Ethnic favoritism in education in kenya. *Quarterly Journal of Political Science*, *11*(1), 1–58.

Larreguy, H., Marshall, J., & Trucco, L. (2015). Breaking Clientelism or Rewarding Incumbents? Evidence from an Urban Titling Program in Mexico. (September).

Levitt, S. D., & Snyder, J. M. (1997, feb). The Impact of Federal Spending on House Election Outcomes. *Journal of Political Economy*, *105*(1), 30–53. Retrieved from http://www.journals.uchicago.edu/doi/10.1086/262064 doi: 10.1086/262064

Liberia Institute of Statistics and Geo-Information Services. (2014). *Liberia demographic and health survey 2013.* Liberia Institute of Statistics and Geo-Information Services.

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.

List, J. A., Shaikh, A. M., & Xu, Y. (2016). *Multiple hypothesis testing in experimental economics* (Tech. Rep.). National Bureau of Economic Research.

Manacorda, M., Miguel, E., & Vigorito, A. (2011, jul). Government Transfers and Political Support. *American Economic Journal: Applied Economics*, *3*(3), 1–28. Retrieved from `http://pubs.aeaweb.org/doi/10.1257/app.3.3.1` doi: 10.1257/app.3.3.1

Ministry of Education. (2016). Liberia education statistics report 2015-2106.

Pop-Eleches, C., & Pop-Eleches, G. (n.d.). Government Spending and Pocketbook Voting: Quasi-Experimental Evidence from Romania. *princeton.edu*. Retrieved from `https://www.princeton.edu/{~}gpop/Pop-ElechesRDpocketbookvotingJune09.pdf`

Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282. Retrieved from `http://dx.doi.org/10.1111/j.1468-0262.2005.00615.x` doi: 10.1111/j.1468-0262.2005.00615.x

Romano, J. P., & Wolf, M. (2010). Balanced control of generalized error rates. *The Annals of Statistics*, 598–633.

Sass, T. R. (2006). Charter schools and student achievement in florida. *Education*, *1*(1), 91–122.

Stigler, G. J. (1972). Economic competition and political competition. *Public Choice*, *13*(1), 91–106.

Termes, A., Bonal, X., Verger, A., Zancajo, A., López, L., Ramírez, Y. C., & Angélica, S. (2015). Public-private partnerships in colombian education: the equity and quality impications of" colegios en concesión".

The World Bank. (2014). *Life expectancy.* (data retrieved from World Development Indicators, `http://data.worldbank.org/indicator/SE.PRM.NENR?locations=LR`)

Tuttle, C. C., Gleason, P., & Clark, M. (2012). Using lotteries to evaluate schools of choice: Evidence from a national study of charter schools. *Economics of Education Review*, *31*(2), 237–253.

UNESCO. (2016). Global monitoring report 2016.

Wantchekon, L. (2003, apr). Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin. *World Politics*, *55*(03), 399–422. Retrieved from `http://www.journals.cambridge.org/abstract{_}S0043887100003798` doi: 10.1353/wp.2003.0018

Winters, M. A. (2012). Measuring the effect of charter schools on public school student achievement in an urban environment: Evidence from new york city. *Economics of Education Review*, *31*(2), 293 - 301. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0272775711001476` (Special Issue: Charter Schools) doi: http://dx.doi.org/10.1016/j.econedurev.2011.08.014

Witte, J., Weimer, D., Shober, A., & Schlomer, P. (2007). The performance of charter schools in wisconsin. *Journal of Policy Analysis and Management*, *26*(3), 557–573. Retrieved from `http://dx.doi.org/10.1002/pam.20265` doi: 10.1002/pam.20265

World development report. (2004). The World Bank. Retrieved from `http://elibrary.worldbank.org/doi/abs/10.1596/0-8213-5637-2` doi: 10.1596/0-8213-5637-2

Zimmer, R., & Buddin, R. (2006). Charter school performance in two large urban districts. *Journal of Urban Economics*, *60*(2), 307 - 326. Retrieved from `http://www.sciencedirect`

.com/science/article/pii/S0094119006000234 doi: http://dx.doi.org/10.1016/j.jue.2006.03.003

B Full list of schools

Below is a complete list of the schools included in the PSL program evaluation. School ID is the EMIS code for the school, Operator indicates the operator the "pair" was assigned to, and groupID identifies "pairs". Treatment is equal to one if the school was treated under the random assignment (and is missing for schools outside the RCT), Original RCT is equal to one for schools in the original RCT list, and Final RCT is equal to one for schools in the final RCT list after swaps. PSL school indicates whether the school actually became a PSL school or not.

Table 13: School list

School ID	Operator	Treatment	GroupID	Original RCT	Final RCT	PSL school
10035	BRIDGE	1	1	1	1	1
110027	BRIDGE	0	1	1	1	0
90031	BRIDGE	0	2	1	1	0
130045	BRIDGE	1	2	1	1	1
30004	BRIDGE	0	3	1	1	0
40279	BRIDGE	1	3	1	1	1
120108	BRIDGE	1	3	1	1	1
120097	BRIDGE	0	4	1	1	0
120446	BRIDGE	1	4	1	1	1
120694	BRIDGE	1	5	1	1	1
120101	BRIDGE	0	5	1	1	0
10100	MtM	0	6	1	1	0
10038	MtM	1	6	1	1	1
20027	BRIDGE	0	7	1	1	0
20057	BRIDGE	1	7	1	1	1
20167	LIYONET	1	8	1	1	1
20182	LIYONET	0	8	1	1	0
20082	OMEGA	0	9	1	1	0
20011	OMEGA	1	9	1	1	1
20176	OMEGA	0	10	1	1	0
20284	OMEGA	1	10	1	1	1
30036	MtM	1	11	0	1	1
30032	MtM	0	11	0	1	0
110355	BRIDGE	0	12	1	1	0
110354	BRIDGE	1	12	1	1	1
110069	BRIDGE	1	13	1	1	1
110072	BRIDGE	0	13	1	1	0
10025	RISING	0	14	1	1	0
10029	RISING	1	14	1	1	1
10107	MtM	1	15	0	1	1
10115	MtM	0	15	0	1	0

Table 13 Continued

School ID	Operator	Treatment	GroupID	Original RCT	Final RCT	PSL school
70009	STELLAM	0	16	1	1	0
70073	STELLAM	1	16	1	1	1
80206	BRAC	1	17	1	1	1
80214	BRAC	0	17	1	1	0
80230	BRAC	1	18	1	1	1
80195	BRAC	0	18	1	1	0
80192	BRAC	1	19	1	1	1
80266	BRAC	0	19	1	1	0
80189	BRAC	0	20	1	1	0
80226	BRAC	1	20	1	1	1
80227	BRAC	0	21	1	1	0
80202	BRAC	1	21	1	1	1
80188	BRAC	0	22	1	1	0
80212	BRAC	1	22	1	1	1
80196	BRAC	0	23	1	1	0
80201	BRAC	1	23	1	1	1
50010	BRIDGE	1	24	1	1	1
50009	BRIDGE	0	24	1	1	0
50012	SCHILD	1	25	1	1	1
50008	SCHILD	0	25	1	1	0
20026	BRIDGE	1	26	1	1	1
20282	BRIDGE	0	26	1	1	0
20038	BRIDGE	1	27	1	1	1
20025	BRIDGE	0	27	1	1	0
120281	BRAC	0	28	1	1	0
120285	BRAC	1	28	1	1	1
120294	OMEGA	0	29	1	1	0
120288	OMEGA	1	29	1	1	1
120280	OMEGA	1	30	1	1	1
120270	OMEGA	0	30	1	1	0
90128	SCHILD	1	31	1	1	1
90127	SCHILD	0	31	1	1	0
90039	SCHILD	0	32	1	1	0
90035	SCHILD	1	32	1	1	1
40077	BRIDGE	1	33	1	1	1
40019	BRIDGE	0	33	1	1	0
50014	SCHILD	0	34	1	1	0
50024	SCHILD	1	34	1	1	1
50147	SCHILD	1	35	0	1	1
50092	SCHILD	0	35	0	1	0
70161	STELLAM	1	36	1	1	1
70097	STELLAM	0	36	1	1	0
110007	MtM	0	37	1	0	0

Table 13 Continued

School ID	Operator	Treatment	GroupID	Original RCT	Final RCT	PSL school
112015	MtM	1	37	1	0	0
110269	OMEGA	0	38	1	1	0
110261	OMEGA	1	38	1	1	0
90155	BRIDGE	1	39	1	1	1
90153	BRIDGE	0	39	1	1	0
90161	SCHILD	0	40	1	0	0
90136	SCHILD	1	40	1	0	0
10068	BRIDGE	0	41	1	1	0
10134	BRIDGE	1	41	1	1	1
10067	BRIDGE	0	42	1	1	0
10053	BRIDGE	1	42	1	1	1
10059	MtM	0	43	1	0	0
10012	MtM	1	43	1	0	0
10052	MtM	1	44	1	1	1
10072	MtM	0	44	1	1	0
10054	MtM	1	45	1	1	1
10051	MtM	0	45	1	1	0
80185	BRAC	0	46	1	1	0
80137	BRAC	1	46	1	1	1
80154	BRAC	1	47	1	1	1
80162	BRAC	0	47	1	1	0
80155	BRAC	1	48	1	1	1
80164	BRAC	0	48	1	1	0
80180	BRAC	1	49	1	1	1
80138	BRAC	0	49	1	1	0
111001	MtM	1	50	1	1	1
111022	MtM	0	50	1	1	0
80096	BRAC	1	51	1	1	1
80061	BRAC	0	51	1	1	0
90037	OMEGA	1	52	1	1	1
90139	OMEGA	0	52	1	1	0
90122	SCHILD	0	53	1	1	0
90130	SCHILD	1	53	1	1	1
90169	SCHILD	0	54	0	1	0
90198	SCHILD	1	54	0	1	1
90008	OMEGA	0	55	1	1	0
90018	OMEGA	1	55	1	1	1
100011	STELLAM	0	56	1	1	0
100061	STELLAM	1	56	1	1	1
110142	BRIDGE	1	57	1	1	1
160011	BRIDGE	0	57	1	1	0
111253	SCHILD	0	58	1	1	0
111276	SCHILD	1	58	1	1	1

Table 13 Continued

School ID	Operator	Treatment	GroupID	Original RCT	Final RCT	PSL school
120305	BRAC	1	59	1	1	1
120242	BRAC	0	59	1	1	0
120271	OMEGA	1	60	1	1	1
120139	OMEGA	0	60	1	1	0
120106	OMEGA	0	61	1	1	0
120064	OMEGA	1	61	1	1	0
20173	LIYONET	0	62	1	1	0
20200	LIYONET	1	62	1	1	1
20178	OMEGA	0	63	1	1	0
20207	OMEGA	1	63	1	1	1
10009	RISING	0	64	1	1	0
111290	RISING	1	64	1	1	0
111212	RISING	0	65	1	1	0
111230	RISING	1	65	1	1	1
110040	OMEGA	1	66	1	1	1
110048	OMEGA	0	66	1	1	0
120328	OMEGA	1	67	1	1	1
120304	OMEGA	0	67	1	1	0
120327	OMEGA	0	68	1	1	0
120320	OMEGA	1	68	1	1	1
120245	BRIDGE	0	69	1	1	0
120257	BRIDGE	1	69	1	1	1
120259	OMEGA	1	70	1	1	1
120252	OMEGA	0	70	1	1	0
20245	BRIDGE	0	71	1	1	0
20003	BRIDGE	1	71	1	1	1
20009	BRIDGE	0	72	1	1	0
20005	BRIDGE	1	72	1	1	1
20021	BRIDGE	1	73	1	1	1
20213	BRIDGE	0	73	1	1	0
80102	BRAC	1	74	1	1	1
80110	BRAC	0	74	1	1	0
120224	BRIDGE	1	75	1	1	1
120226	BRIDGE	0	75	1	1	0
120215	OMEGA	1	76	1	1	1
120228	OMEGA	0	76	1	1	0
120208	OMEGA	0	77	1	1	0
120207	OMEGA	1	77	1	1	1
10089	BRIDGE	1	78	1	1	1
10043	BRIDGE	0	78	1	1	0
150043	LIYONET	0	79	1	1	0
150082	LIYONET	1	79	1	1	1
100111	STELLAM	0	80	1	1	0

Table 13 Continued

School ID	Operator	Treatment	GroupID	Original RCT	Final RCT	PSL school
100022	STELLAM	1	80	1	1	1
20053	OMEGA	0	81	1	1	0
20047	OMEGA	1	81	1	1	1
10007	RISING	0	82	1	1	0
10018	RISING	1	82	1	1	1
50030	SCHILD	1	83	1	1	1
50029	SCHILD	0	83	1	1	0
50070	SCHILD	0	84	1	1	0
50107	SCHILD	1	84	1	1	1
50111	SCHILD	1	85	1	0	0
50064	SCHILD	0	85	1	0	0
50076	SCHILD	0	86	1	1	0
50063	SCHILD	1	86	1	1	1
50067	SCHILD	0	87	1	1	0
50081	SCHILD	1	87	1	1	1
110092	RISING	0	88	1	1	0
110167	RISING	1	88	1	1	1
80023	BRAC	0	89	1	1	0
80014	BRAC	1	89	1	1	1
80051	BRAC	0	90	1	1	0
80056	BRAC	1	90	1	1	1
80027	BRAC	1	91	1	1	1
80022	BRAC	0	91	1	1	0
80047	BRAC	0	92	1	1	0
80001	BRAC	1	92	1	1	1
120361	OMEGA	0	93	1	1	0
120352	OMEGA	1	93	1	1	1
80060	BRAC	1	94	1	1	1
80070	BRAC	0	94	1	1	0
20063	LIYONET	1	95	1	1	1
20239	LIYONET	0	95	1	1	0
20071	OMEGA	1	96	1	1	1
20066	OMEGA	0	96	1	1	0
110022	BRIDGE			0	0	1
20131	BRIDGE			0	0	1
10129	RISING			0	0	1