Partnership Schools for Liberia:

Baseline Report on the Experimental Evaluation

DRAFT

Mauricio Romero* Justin Sandefur† Wayne Aaron Sandholtz*

March 6, 2017‡

Abstract

We summarize the baseline survey data from the randomized evaluation of the Partnership Schools for Liberia (PSL) program. This public-private partnership delegates management of 93 randomly assigned public schools to a variety of private, for-profit companies and non-profit organizations to be operated free of charge to students with no selective admissions. Baseline data was collected in October and November of 2016, shortly after the start of the school year in both treatment and control schools, allowing us to perform three distinct tasks: (*i*) we test and confirm balance on slow-moving school characteristics, suggesting randomization produced a valid control group; (*ii*) we find small, beneficial, short-term impacts on principal and teacher behavior, including reduced absenteeism, greater time teaching, and reduced corporal punishment; and (*iii*) we find that PSL has increased enrollment relative to control schools and no evidence that PSL operators have selectively admitted brighter or wealthier students. Two caveats apply: we cannot (yet) measure what share of this is due to enrolling previously out-of-school children, and some discrepancies with school rosters exist which we hope to resolve during mid-line data collection.

*University of California, San Diego

†Center for Global Development

‡This draft report was prepared for the Liberian Ministry of Education for internal use. We are grateful to Arja Dayal, Dackermue Dolo, and their team at Innovations for Poverty Action who led the data collection. Avi Ahuja and Dev Patel provided excellent research assistance. The evaluation was supported by the UBS Optimus Foundation and Aestus Trust. The views expressed here are ours, and not those of the Ministry or our funders. All errors are ours.

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). The civil war and the Ebola epidemic contributed to the weakening of institutions with already limited capacity. 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 companies. We present results from a large-scale field experiment to evaluate the effect of this public-private partnership (PPP).

The Partnership Schools for Liberia (PSL) program delegated management of 93 randomly assigned public schools to a variety of private, for-profit companies and non-profit organizations. The government (and donors) provide these operators with funding on a per-pupil level. In exchange, operators are responsible for the daily management of the schools and held accountable for results. These schools are to remain free and non-selective public schools (i.e., operators are not allowed to charge fees or choose which students to enroll). PSL school buildings remain under the ownership of the government and teachers in PSL schools are existing government teachers (i.e., public servants). Private providers are accountable to the government for performance.

The evaluation design features two key design features that allows us to estimate the effect of the PSL program on student learning outcomes. First, the randomization at the school level. This guarantees that the schools, and the communities near them, are similar across control and treatment schools. However, treating a school may cause changes in enrollment that alter the observable and un-observable characteristics of students in treatment schools. Thus, differences in test scores and other student-level variables in PPP and control schools may be driven by differences in the underlying population of students. We are able to overcome this by randomly assigning treatment status at the school level and sampling students from the enrollment log of the year prior to the implementation of the program (before awareness of the program). Each student is 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.[1]

[1] Although we partially overcome the external validity issues of estimates from admission-based lotteries, we randomize

A key objective of the PSL program is to give operators flexibility in creating innovative ways to improve school management, teacher accountability, pedagogy, practices, and ultimately student learning. However, the ability of PSL to improve education quality hinges on whether three key accountability relationships in the education system are maintained or improved.

The first is managerial accountability. A central hypothesis underlying Liberia's PSL program is that non-state operators have the capacity to implement routine performance management systems, monitor teacher attendance, track student performance, and provide teachers with feedback and support. Teachers in PSL schools will be drawn from the existing pool of unionized civil servants with lifetime appointments and be paid directly by the Liberian government. Private operators will have limited authority to request teacher reassignments and no authority to promote or dismiss civil service teachers. The hypothesis is that accountability can be generated through monitoring and support, rather than rewards and threats, and there is evidence that management practices can have significant effects on workers' performance (Bloom, Liang, Roberts, & Ying, 2014; Bloom, Eifert, Mahajan, McKenzie, & Roberts, 2013; Bennedsen, Nielsen, Pérez-González, & Wolfenzon, 2007).

But this hypothesis stands in stark contrast to standard labor economics theories of accountability in the workplace, which have dominated the economics of education literature in developing countries. These theories stress civil service protections and labor unions as impediments to accountability (Mbiti, 2016). In response, the experimental literature has focused on solutions such as payment for performance (Muralidharan & Sundararaman, 2011) and flexible labor contracts with credible threat of dismissal (Banerjee, Cole, Duflo, & Linden, 2007; Duflo, Dupas, & Kremer, 2011, 2012; Duflo, Hanna, & Ryan, 2012).

The second relationship is bottom-up accountability. In the framework of the *World Development Report* (2004) on public service delivery, there is a "short route" to accountability if parents are able to exercise "client power" in their interactions with teachers and schools. Client power emerges from the freedom to choose another provider.

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 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

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 2.3 for more details.

schools.

The third relationship is top-down accountability. The 93 pilot schools remain public, subject to regulation and with quality assured by the government. Charter school operators' contracts can be terminated if they do not achieve certain pre-established standards. PSL schools are ultimately public schools and oversight from the government is crucial. In the absence of government accountability, policies that leverage the private sector to improve coverage and quality may fail to deliver better outcomes. For example, Abdulkadiroglu, Pathak, and Walters (2015) show that a voucher program in New Orleans actually resulted in lower test scores for students attending voucher schools. The authors note that "the program's accountability rules do not identify the low-quality schools that drive its negative achievement effects". In other words, the program lacks effective oversight. Crawfurd (2016) compares the management practices of public, private, and PPP schools in Uganda. Overall, there is no difference in management across school types, reflecting a lack of oversight across the board. However, a subset of PPP schools run by a foreign owned, non-profit chain of schools with a robust accountability framework, performs significantly better than other schools. This suggests that PPP have the potential to improve school management, but that strong oversight–from the government or donors– is necessary.

Our baseline survey data show that teachers in PPP schools are more likely to show up to school, less likely to be off-task while in school, and receive more training. Students also report changes in teacher behavior. Specifically, students are less likely to report that English and math teachers miss school, less likely to report that teachers leave the classroom in the middle of a lecture, and less likely to report that teachers hit them. Principals in PSL schools report spending less time teaching and more time supporting other teachers and providing feedback. Principals also seem more engaged with parents (they are more likely to have ready access to the PTA head's phone number). Finally, our data shows that enrollment in PSL schools is higher, and we find no evidence that PSL schools have selectively admitted wealthier students.

2 Experimental Design

2.1 Context

Liberia is highly dependent on external aid to both finance and implement education, even compared to other low-income countries in Africa, who finance the vast bulk of education spending through domestic tax revenue. Over the past decade, Liberia's basic education budget has been roughly $40 million per year (spent almost entirely 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 the African region, who 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 pre-school and primary school attend non-state schools (Ministry of Education, 2016). And 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 children enrolled in "early childhood" education 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).

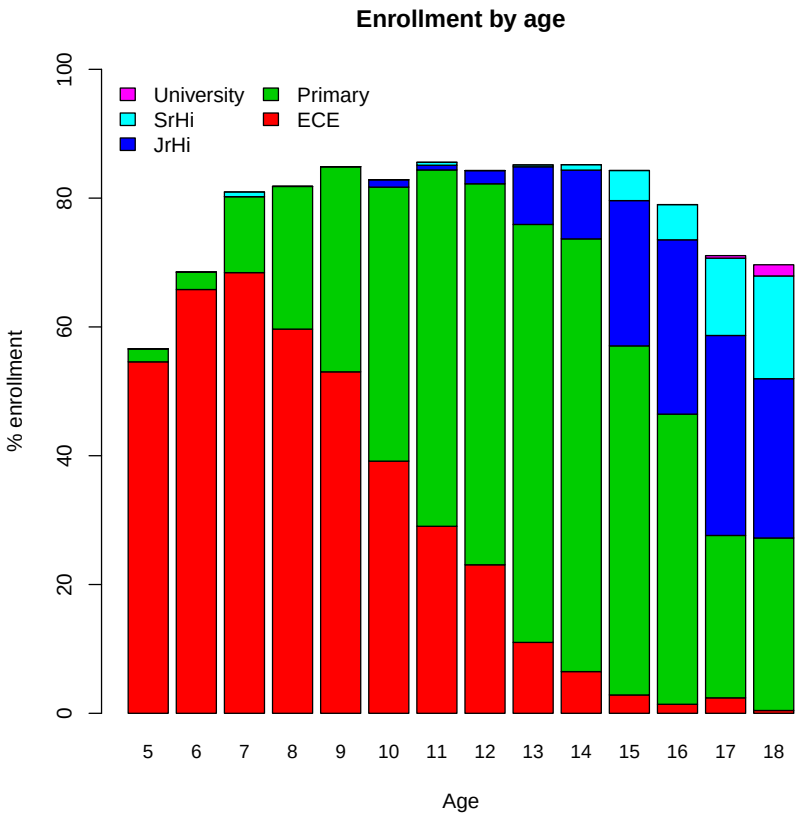

Figure 1: Authors' calculations based on 2014 Household Income and Expenditures Survey

2.2 Intervention

The Partnership Schools for Liberia (PSL) program is a contract-management public-private partnership (PPP). 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'[2], tuition for early-childhood education is LBD $3,500 LBD per year or around US $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

[2]Technically, public schools are free, but in reality most schools charge informal fees. See section 2.4 for statistics on these fees.

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 an additional funding of US $50 per student from outside donors (for a maximum of US $3,250 per grade or 65 students per grade).[3] 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 deliver the Liberian national curriculum, but may supplement it with remedial programs, prioritizing subjects, longer school days, and non-academic activities. They are also welcome to provide additional inputs like extra teachers, books or uniforms, as long as they pay for them.

There are eight operators running public schools under PSL. 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]

2.3 Sampling and design

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 do not have a secondary school within the same compound. Only 299 schools satisfied all the criteria. However, it is important to 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.

[3] One of the operators, Bridge International Academies, has a different arrangement with donors and receive a total of XXX for running their schools in Liberia, regardless of enrollment.

[4] Bridge International Academies is managing two extra "demonstration" schools that were not randomized and are therefore not part of our sample. Omega Academies opted for not operating 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 does not perfectly align with the set of schools actually being managed by PSL operators.

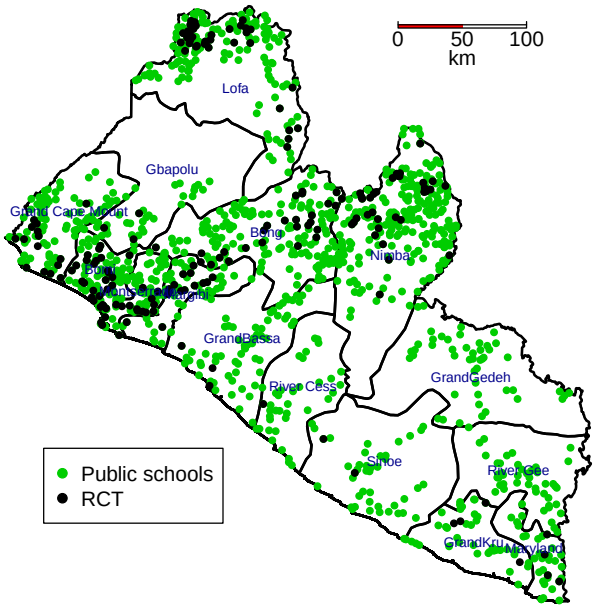

Figure 2: Public schools in Liberia and those within our sample.

Within this list, 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 would operate. Once each operator "approved" of this list, we randomized within each pair the treatment assignment.[6]

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 2 classrooms each. Therefore, we gave these operators new

[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 waters, 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. Therefor, there is one extra treatment school.

"pairs" of schools and told them, as before, they would be operate one of these schools (but not which one), and therefore should be approve of the list before given the actual assignment from randomization. We present results from this final list of treatment and control schools in the main text, and results from the "original" assignment in an the extra tables in the appendix. As expected, since the only difference between one list and the other is four pairs of schools, the results between the final treatment and control school list and the "original" list are almost identical. Figure 3 shows the final treatment assignment.

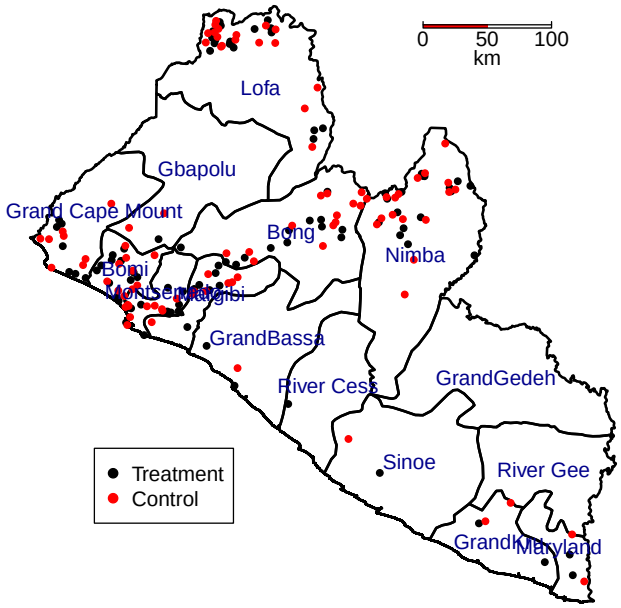

Figure 3: Geographical distribution of treatment and control schools.

2.4 Data and balance

Since the composition of students may change across PSL and control schools in response to treatment assignment, from each school we sampled 20 students from each school's 2015/2016 enrollment log, which was created 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 teachers in the school, and conducted in-depth surveys for those teaching Math or 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. We further conducted school level surveys, collecting information on school facilities, the teaching roster, input availability and expenditures. Unlike most school inputs, textbooks are easily assignable to grades and subjects. We therefore collected information on textbook purchases at the grade subject level.

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. Prior to the collection of the midline data, we submitted a pre-analysis plan to the AEA registry.

Figure 4

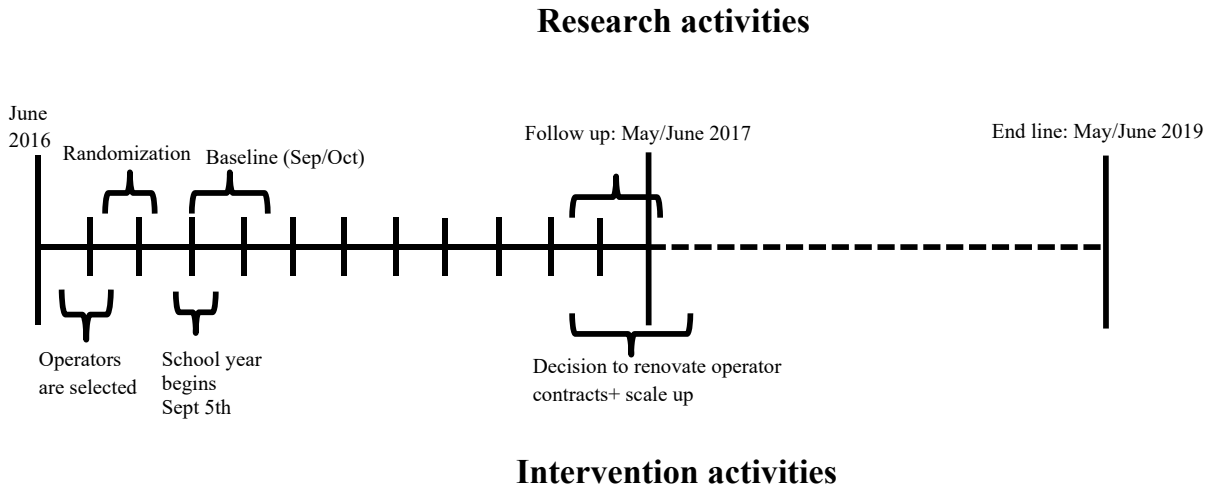

Table 1 shows that observable, time-invariant characteristics of students and schools are balanced across treatment and control. 80% of schools in our sample are in rural areas, over an hour away from the nearest bank (usually located in the nearest urban center); over 10% need to hold some classes outside due to insufficient classrooms. Nearly 55% of our students are male, and they have an average age of twelve. There seems to be a small difference in test scores (English and Math), with students in treatment schools outperforming students in control schools. We attribute this difference to an early treatment effect that is mostly concentrated in students that we survey four weeks or more after school started (see Figure 5 in Appendix B).

Table 1: Balance in observable, time invariant, school and student characteristics

Panel A: School characteristics				
	Control	Treatment	Difference	Difference (F.E)
Facilities (PCA)	0.078	-0.041	-0.119	-0.107
	(1.635)	(1.548)	(0.234)	(0.235)
Hold some classes outside	0.130	0.118	-0.012	-0.011
	(0.339)	(0.325)	(0.049)	(0.050)
Rural	0.815	0.796	-0.020	-0.014
	(0.390)	(0.405)	(0.059)	(0.048)
Time to nearest bank	68.663	74.172	5.509	5.498
	(60.752)	(66.493)	(9.363)	(8.703)
Panel B: Student characteristics				
	Control	Treatment	Difference	Difference (F.E)
Male	0.567	0.549	-0.018	-0.021*
	(0.496)	(0.498)	(0.020)	(0.013)
Wealth Index	-0.064	-0.066	-0.002	-0.013
	(1.475)	(1.457)	(0.126)	(0.055)
ECE before grade 1	0.814	0.834	0.020	0.019
	(0.389)	(0.372)	(0.024)	(0.017)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has school characteristics, while Panel B has student characteristics. 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, play ground, 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 Results

3.1 Principals and school management

Table 2 shows the results from comparing principals in PSL and control schools. The main difference across principal in PSL and control schools is not in demographic characteristics, experience, qualifications, or ability, but in how they use their time, and how organized they are. Principals in PSL schools spend less time teaching, and more time supporting other teachers. Additionally, they are more likely to know the PTA head's phone number or where to find it (a sign they are more organized and/or engaged with parents). Despite the 'free' primary school policy, over 40% of control schools charge some type of fee. While fees are prohibited under PSL, 17% of principals report charging some fees. This may be the result of some operators taking over 3-4 weeks to take charge of some of their schools.

Table 2: Principal's characteristics and time use

Panel A: Principal's characteristics				
	Control	Treatment	Difference	Difference (F.E)
Male	0.902	0.882	-0.020	-0.022
	(0.299)	(0.325)	(0.046)	(0.041)
Age	50.641	49.720	-0.921	-0.751
	(9.377)	(9.401)	(1.381)	(1.347)
Experience (Yrs)	18.424	18.968	0.544	0.606
	(9.953)	(10.279)	(1.488)	(1.547)
Experience (Yrs) as principal	7.304	9.473	2.169*	2.253*
	(7.271)	(8.344)	(1.150)	(1.170)
Experience in a private school	0.446	0.398	-0.048	-0.054
	(0.500)	(0.492)	(0.073)	(0.066)
Teaching certificate or university degree	0.750	0.796	0.046	0.043
	(0.435)	(0.405)	(0.062)	(0.067)
Math (IRT)	0.085	0.121	0.035	0.031
	(0.632)	(0.613)	(0.092)	(0.096)
Memory (IRT)	0.039	-0.014	-0.053	-0.066
	(0.959)	(0.946)	(0.140)	(0.135)
Word-Association (IRT)	0.149	0.256	0.107	0.108
	(0.546)	(0.425)	(0.072)	(0.071)
Principal could not find PTA's head phone No.	0.554	0.387	-0.167**	-0.166**
	(0.500)	(0.490)	(0.073)	(0.070)
Charge any fees	0.467	0.172	-0.295***	-0.296***
	(0.502)	(0.379)	(0.065)	(0.066)
Panel B: Principal's time use				
	Control	Treatment	Difference	Difference (F.E)
Discipline	3.668	4.965	1.297	1.261
	(5.009)	(6.529)	(0.855)	(0.834)
Teaching	7.404	3.971	-3.434***	-3.378***
	(8.539)	(6.569)	(1.121)	(1.040)
Planning	2.333	1.573	-0.760**	-0.740*
	(2.867)	(2.277)	(0.381)	(0.382)
Teacher Support	2.311	3.735	1.424**	1.405**
	(3.518)	(5.371)	(0.667)	(0.655)
PTA/Community	2.803	2.729	-0.074	-0.099
	(3.059)	(2.752)	(0.428)	(0.399)
Off ground	2.015	2.485	0.470	0.495
	(3.132)	(4.677)	(0.585)	(0.586)
Admin task	2.080	2.330	0.250	0.241
	(3.098)	(3.184)	(0.462)	(0.451)
Total time (Hrs)	22.612	21.787	-0.826	-0.815
	(14.421)	(14.363)	(2.116)	(1.986)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has principal characteristics, while Panel B has Principal's time use. Standard errors are clustered at the school level. The sample is the final treatment and control allocation.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2 Classroom observations

Table 3 shows the results from the Stallings Classroom Observation. Importantly, we were not able to conduct classroom observations in all schools since some of them were not in session during a normal school day. Even just a few weeks after school started, noticeable differences emerge in how teachers conduct classes. The most remarkable difference is that teachers in PSL schools are 9% less likely to be off-task (i.e., outside the classroom), and they use this time mostly to do passive instruction. As expected, given the involvement of some operators that rely heavily on technology to deliver lesson plans (e.g., Bridge and Omega) PSL teachers are 4% more likely to use ICT.

Table 3: Stallings Clasroom Observation

Panel A: Teacher time use				
	Control	Treatment	Difference	Difference (F.E)
Observation	0.90	0.96	0.05	0.05
	(0.30)	(0.20)	(0.04)	(0.04)
Active Instruction	0.43	0.45	0.02	0.01
	(0.28)	(0.24)	(0.04)	(0.04)
Passive Instruction	0.19	0.26	0.07**	0.06**
	(0.18)	(0.21)	(0.03)	(0.03)
Classroom Management	0.14	0.16	0.03	0.02
	(0.16)	(0.16)	(0.02)	(0.02)
Teacher Off-Task	0.24	0.12	-0.11**	-0.09**
	(0.33)	(0.23)	(0.04)	(0.04)
Student Off-Task	0.27	0.32	0.04	0.04
	(0.28)	(0.28)	(0.04)	(0.05)
Non-involved students	0.20	0.24	0.04	0.04
	(0.23)	(0.24)	(0.04)	(0.04)
Panel B: Materials				
	Control	Treatment	Difference	Difference (F.E)
No Material	0.16	0.12	-0.05*	-0.04
	(0.17)	(0.15)	(0.03)	(0.03)
Books	0.11	0.13	0.02	0.00
	(0.22)	(0.20)	(0.03)	(0.03)
Paper/Pencil	0.11	0.13	0.02	0.01
	(0.17)	(0.19)	(0.03)	(0.03)
Blackboard/Projection	0.59	0.55	-0.04	-0.02
	(0.26)	(0.29)	(0.04)	(0.05)
Learning Aides	0.02	0.03	0.01	0.01
	(0.09)	(0.11)	(0.02)	(0.01)
ICT	0.00	0.04	0.04***	0.04***
	(0.00)	(0.12)	(0.01)	(0.01)
Panel C: Student interaction				
	Control	Treatment	Difference	Difference (F.E)
Cooperative	0.00	0.01	0.00	0.00
	(0.02)	(0.04)	(0.00)	(0.01)
Everyone	0.38	0.39	0.01	-0.01
	(0.31)	(0.31)	(0.05)	(0.05)
Small Group	0.38	0.37	-0.02	0.01
	(0.27)	(0.28)	(0.04)	(0.05)
Big Group	0.48	0.46	-0.02	0.00
	(0.26)	(0.28)	(0.04)	(0.05)
Individual Student	0.31	0.30	-0.01	-0.00
	(0.28)	(0.26)	(0.04)	(0.04)
Social interaction	0.00	0.00	-0.00	0.00
	(0.02)	(0.01)	(0.00)	(0.00)
Non-involved students	0.20	0.24	0.04	0.04
	(0.23)	(0.24)	(0.04)	(0.04)

This tables presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, as well as the difference between treatment and control (Column 3), and the difference taking into account the randomization designn–i.e., including 'pair' fixed effects– (Column 4). Panel A has teacher's time use as defined by the Stalling Classroom Observation, Panel B has usage of material by the teacher, and Panel C has student interactions. Standard errors are clustered at the school level. The sample is the final treatment and control allocation. The sample for the first row of panel A is all the classroom observations we attempted. The sample for every other row is only the schools on which we were able to conduct

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3 Teachers

Table 4 shows the results from the teacher interviews. There are several noteworthy findings. First, there are more teachers in PSL schools than in Control schools. Importantly, these teachers are less likely to be on the government's payroll, and more likely to be paid by some NGO (this includes the operator). Teachers in PSL have less experience teaching, but have higher scores on a simple math and memory test.

There are also some perplexing results regarding PSL and job satisfaction: Teachers in PSL schools are less satisfied with their job, but of the teachers who have heard of PSL, more than half, both in treatment and control schools, would prefer to work in a PSL school. Over 80% would support the expansion of the PSL program. (The number of teachers who have heard of PSL, would prefer to work in a PSL school, and support the expansion of the program is higher in PSL schools).

Table 4: Teachers

	Control	Treatment	Difference	Difference (F.E)
Avg. number teachers interviewed per school	5.522	7.333	1.812***	1.805***
	(2.509)	(2.082)	(0.339)	(0.314)
Fraction of teachers getting salary from gov't	0.711	0.578	-0.132***	-0.132***
	(0.306)	(0.345)	(0.048)	(0.046)
Fraction of teachers getting salary from an NGO	0.000	0.034	0.034**	0.034**
	(0.000)	(0.148)	(0.015)	(0.015)
Year the avg. teacher started teaching	2002.632	2006.260	3.628***	3.609***
	(5.813)	(4.746)	(0.807)	(0.753)
Fraction of teachers who heard of PSL	0.348	0.572	0.223***	0.222***
	(0.287)	(0.348)	(0.047)	(0.040)
Fraction of teachers who prefer to work at a PSL school	0.740	0.895	0.154***	0.192***
	(0.399)	(0.223)	(0.053)	(0.052)
Fraction of teachers who support PSL expansion	0.960	0.986	0.027	0.028
	(0.176)	(0.058)	(0.022)	(0.024)
Fraction that has recieved training since August 2016	0.119	0.221	0.102***	0.098***
	(0.194)	(0.260)	(0.034)	(0.029)
Job satisfaction index	0.100	-0.054	-0.154	-0.144
	(0.944)	(0.814)	(0.130)	(0.121)
(mean) SecondJob	0.298	0.223	-0.075*	-0.073*
	(0.296)	(0.243)	(0.040)	(0.038)
(mean) TeachingCareer	0.897	0.933	0.035	0.035
	(0.182)	(0.113)	(0.022)	(0.023)
Math score IRT	-0.059	0.019	0.078**	0.076**
	(0.246)	(0.261)	(0.037)	(0.038)
Memory score IRT	-0.046	0.038	0.084	0.085
	(0.342)	(0.348)	(0.051)	(0.054)
Word Association score IRT	-0.030	0.007	0.037	0.037
	(0.357)	(0.300)	(0.048)	(0.045)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4 Students

Table 5 shows results from student interviews. Students in PSL schools are less likely to be wearing a uniform; this may be positive, as uniform fees may be a barrier in regular public schools for some families. PSL has not generated any "selection" issues: Students enrolled in PSL schools in 2015/2016 are equally likely to still be enrolled in schools (in any school, or in the PSL school). Importantly, PSL students are less likely to miss school (it goes down from 30% in Control schools to 22% in PSL schools), less likely to report that both English and math teachers missed school the previous week (teacher absenteeism goes down by 5 percentage points from a base of17%), more likely to report that English and math teachers never hit them (in each case it goes up by about 10 percentage points from a base of about 55%), and more likely to report

that teachers never leave the classroom (an increase of 7 percentage points from a base of 65%).

Table 5: Students

	Control	Treatment	Difference	Difference (F.E)
Is the student wearing a uniform?	0.618	0.482	-0.136**	-0.122***
	(0.486)	(0.500)	(0.053)	(0.034)
Still goes to school	0.981	0.983	0.002	0.004
	(0.138)	(0.130)	(0.006)	(0.004)
Goes to the same school	0.931	0.931	0.000	0.008
	(0.253)	(0.253)	(0.016)	(0.011)
Missed school last week	0.297	0.229	-0.068***	-0.077***
	(0.457)	(0.420)	(0.024)	(0.018)
English teacher missed school last week	0.149	0.123	-0.026	-0.048***
	(0.356)	(0.328)	(0.019)	(0.013)
Math teacher missed school last week	0.174	0.135	-0.039*	-0.061***
	(0.380)	(0.342)	(0.022)	(0.015)
Math teacher never hits students	0.559	0.641	0.082***	0.090***
	(0.497)	(0.480)	(0.029)	(0.023)
English teacher never hits students	0.575	0.652	0.077***	0.075***
	(0.494)	(0.476)	(0.029)	(0.021)
Math teacher helped outside of class	0.429	0.461	0.032	0.039*
	(0.495)	(0.499)	(0.033)	(0.021)
English teacher helped outside of class	0.441	0.448	0.007	0.022
	(0.497)	(0.497)	(0.031)	(0.019)
Math teacher never leaves classroom	0.663	0.712	0.049*	0.063***
	(0.473)	(0.453)	(0.028)	(0.020)
English teacher never leaves classroom	0.680	0.731	0.051*	0.062***
	(0.466)	(0.444)	(0.027)	(0.018)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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 for the first four rows is all students we interview. For the fifth row and below is the sample of students who are still enrolled in school. The sample of schools is the final treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Conclusion

This report summarizes the baseline survey data from the randomized impact evaluation of the Partnership Schools for Liberia (PSL) program. Because the data was collected shortly after the start of the school year, the baseline information can be used both to test for balance, and to measure very short-term impacts of the program. We find the experiment is balanced on slow-moving school characteristics such as school infrastructure. Turning to short-term impacts, we find small, beneficial, effects on principal and teacher behavior, including reduced absenteeism, greater time teaching, and reduced corporal punishment. We also show that PSL has also increased enrollment relative to control schools, but we cannot (yet) measure what

share of this is due to enrolling previously out-of-school children, rather than attracting pupils from other schools.

A key concern with any PPP program in education is that private operators will screen students based on ability or wealth. We find no evidence that this is happening systematically in PSL. However, a caveat applies: there are a few unexplained anomalies in the school rosters from 2015/16 that were used for sampling which may obscure screening behavior. We hope to resolve these issues during mid-line data collection in June 2017.

References

Abdulkadiroglu, A., Pathak, P. A., & Walters, C. R. (2015). *School vouchers and student achievement: Evidence from the louisiana scholarship program* (Tech. Rep.). National Bureau of Economic Research.

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.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in india. *The Quarterly Journal of Economics*, *122*(3), 1235-1264. Retrieved from http://qje.oxfordjournals.org/content/122/3/1235.abstract doi: 10.1162/qjec.122.3.1235

Bennedsen, M., Nielsen, K. M., Pérez-González, F., & Wolfenzon, D. (2007). Inside the family firm: The role of families in succession decisions and performance. *The Quarterly Journal of Economics*, *122*(2), 647–691.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*, *128*(1), 1–51.

Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2014). Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics*, qju032.

Crawfurd, L. (2016). School management in uganda.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, *101*(5), 1739-74. Retrieved from http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.1739 doi: 10.1257/aer.101.5.1739

Duflo, E., Dupas, P., & Kremer, M. (2012, March). *School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from kenyan primary schools* (Working Paper No. 17939). National

Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17939 doi: 10.3386/w17939

Duflo, E., Hanna, R., & Ryan, S. P. (2012, June). Incentives work: Getting teachers to come to school. *American Economic Review*, *102*(4), 1241-78. Retrieved from http://www.aeaweb.org/articles?id=10.1257/aer.102.4.1241 doi: 10.1257/aer.102.4.1241

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.

Mbiti, I. M. (2016, August). The need for accountability in education in developing countries. *Journal of Economic Perspectives*, *30*(3), 109-32. Retrieved from http://www.aeaweb.org/articles?id=10.1257/jep.30.3.109 doi: 10.1257/jep.30.3.109

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

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from india. *Journal of Political Economy*, *119*(1), pp. 39-77. Retrieved from http://www.jstor.org/stable/10.1086/659655

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

UNESCO. (2016). Global monitoring report 2016.

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

A Extra tables

Table 6: Balance 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)
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)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has school characteristics, while Panel B has student characteristics. 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, play ground, and an improved water source. Standard errors are clustered at the school level. The sample is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Stallings clasroom observation, immediately after treatment

Panel A: Teacher time use				
	Control	Treatment	Difference	Difference (F.E)
Observation	0.89	0.95	0.05	0.05
	(0.31)	(0.23)	(0.04)	(0.04)
Active Instruction	0.43	0.45	0.01	0.01
	(0.29)	(0.24)	(0.04)	(0.04)
Passive Instruction	0.19	0.25	0.07**	0.06*
	(0.18)	(0.21)	(0.03)	(0.03)
Classroom Management	0.14	0.16	0.03	0.02
	(0.16)	(0.16)	(0.02)	(0.02)
Teacher Off-Task	0.24	0.14	-0.11**	-0.09**
	(0.33)	(0.23)	(0.04)	(0.04)
Student Off-Task	0.27	0.32	0.04	0.05
	(0.28)	(0.28)	(0.04)	(0.04)
Non-involved students	0.20	0.24	0.04	0.05
	(0.23)	(0.24)	(0.04)	(0.04)
Panel B: Materials				
	Control	Treatment	Difference	Difference (F.E)
No Material	0.17	0.12	-0.05*	-0.04
	(0.17)	(0.16)	(0.03)	(0.03)
Books	0.11	0.14	0.03	-0.00
	(0.22)	(0.22)	(0.04)	(0.04)
Paper/Pencil	0.11	0.13	0.02	0.00
	(0.17)	(0.19)	(0.03)	(0.03)
Blackboard/Projection	0.58	0.53	-0.05	-0.02
	(0.26)	(0.30)	(0.04)	(0.05)
Learning Aides	0.02	0.03	0.01	0.01
	(0.09)	(0.11)	(0.02)	(0.01)
ICT	0.00	0.04	0.04***	0.04***
	(0.00)	(0.12)	(0.01)	(0.02)
Panel C: Student interaction				
	Control	Treatment	Difference	Difference (F.E)
Cooperative	0.00	0.01	0.00	0.00
	(0.02)	(0.04)	(0.00)	(0.01)
Everyone	0.38	0.38	-0.00	-0.04
	(0.32)	(0.31)	(0.05)	(0.05)
Small Group	0.38	0.37	-0.01	0.03
	(0.27)	(0.28)	(0.04)	(0.05)
Big Group	0.48	0.46	-0.02	0.02
	(0.27)	(0.28)	(0.04)	(0.05)
Individual Student	0.31	0.31	0.00	0.03
	(0.28)	(0.26)	(0.04)	(0.05)
Social interaction	0.00	0.00	-0.00	0.00
	(0.02)	(0.01)	(0.00)	(0.00)
Non-involved students	0.20	0.24	0.04	0.05
	(0.23)	(0.24)	(0.04)	(0.04)

This table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has teacher's time use as defined by the Stalling Classroom Observation, Panel B has usage of material by the teacher, and Panel C has student interactions. Standard errors are clustered at the school level. The sample is the original treatment and control allocation.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Teacher responses immediately after treatment

	Control	Treatment	Difference	Difference (F.E)
Avg. number teachers interviewed per school	5.533	7.323	1.790***	1.783***
	(2.658)	(2.299)	(0.366)	(0.303)
Fraction of teachers getting salary from gov't	0.723	0.588	-0.135***	-0.135***
	(0.296)	(0.343)	(0.047)	(0.046)
Fraction of teachers getting salary from an NGO	0.000	0.034	0.034**	0.034**
	(0.000)	(0.148)	(0.015)	(0.015)
Year the avg. teacher started teaching	2002.572	2006.230	3.658***	3.601***
	(5.720)	(4.710)	(0.798)	(0.770)
Fraction of teachers who heard of PSL	0.346	0.566	0.220***	0.219***
	(0.290)	(0.354)	(0.048)	(0.040)
Fraction of teachers who prefer to work at a PSL school	0.763	0.880	0.117**	0.154***
	(0.383)	(0.245)	(0.053)	(0.054)
Fraction of teachers who support PSL expansion	0.973	0.986	0.013	0.012
	(0.133)	(0.058)	(0.017)	(0.020)
Fraction that has recieved training since August 2016	0.125	0.213	0.087***	0.083***
	(0.200)	(0.249)	(0.033)	(0.028)
Job satisfaction index	0.144	-0.084	-0.227*	-0.218*
	(0.965)	(0.836)	(0.133)	(0.127)
(mean) SecondJob	0.288	0.221	-0.068*	-0.066*
	(0.295)	(0.244)	(0.040)	(0.037)
(mean) TeachingCareer	0.897	0.937	0.040*	0.039*
	(0.187)	(0.112)	(0.023)	(0.024)
Math score IRT	-0.050	0.014	0.064*	0.062*
	(0.237)	(0.261)	(0.037)	(0.037)
Memory score IRT	-0.048	0.060	0.107**	0.108**
	(0.342)	(0.339)	(0.050)	(0.054)
Word Association score IRT	-0.019	0.004	0.023	0.023
	(0.347)	(0.300)	(0.048)	(0.045)

Teachers are sampled from those working in the school after treatment began, interviewed just a few weeks into the school year; thus the table combines both selection through hiring/retention policies and causal effects on given teachers. The tables presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has teacher's time use as defined by the Stalling Classroom Observation, Panel B has usage of material by the teacher, and Panel C has student interactions. Standard errors are clustered at the school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Student responses immediately after treatment

	Control	Treatment	Difference	Difference (F.E)
Is the student wearing a uniform?	0.616	0.484	-0.132**	-0.113***
	(0.487)	(0.500)	(0.053)	(0.034)
Still goes to school	0.980	0.983	0.003	0.006
	(0.140)	(0.128)	(0.006)	(0.004)
Goes to the same school	0.933	0.933	-0.000	0.009
	(0.250)	(0.251)	(0.016)	(0.011)
Missed school last week	0.297	0.230	-0.068***	-0.076***
	(0.457)	(0.421)	(0.024)	(0.018)
English teacher missed school last week	0.157	0.123	-0.034*	-0.052***
	(0.364)	(0.329)	(0.020)	(0.013)
Math teacher missed school last week	0.183	0.138	-0.045**	-0.063***
	(0.387)	(0.345)	(0.022)	(0.015)
Math teacher never hits students	0.547	0.632	0.085***	0.088***
	(0.498)	(0.482)	(0.029)	(0.023)
English teacher never hits students	0.566	0.644	0.077***	0.072***
	(0.496)	(0.479)	(0.029)	(0.021)
Math teacher helped outside of class	0.421	0.464	0.043	0.045**
	(0.494)	(0.499)	(0.032)	(0.021)
English teacher helped outside of class	0.432	0.451	0.018	0.027
	(0.496)	(0.498)	(0.030)	(0.018)
Math teacher never leaves classroom	0.668	0.705	0.036	0.050**
	(0.471)	(0.456)	(0.029)	(0.021)
English teacher never leaves classroom	0.686	0.727	0.041	0.054***
	(0.464)	(0.445)	(0.028)	(0.019)

Students are sampled from enrollment rosters in the year prior to treatment, and the table can be interpreted as intent-to-treat effects comparing the original treatment and control allocation. Data was collected just weeks into the school year. The table presents the mean and standard error of the mean (in parenthesis) for the treatment (Column 1) and control (Column 2) group, 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). Panel A has teacher's time use as defined by the Stalling Classroom Observation, Panel B has usage of material by the teacher, and Panel C has student interactions. Standard errors are clustered at the school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Extra figures

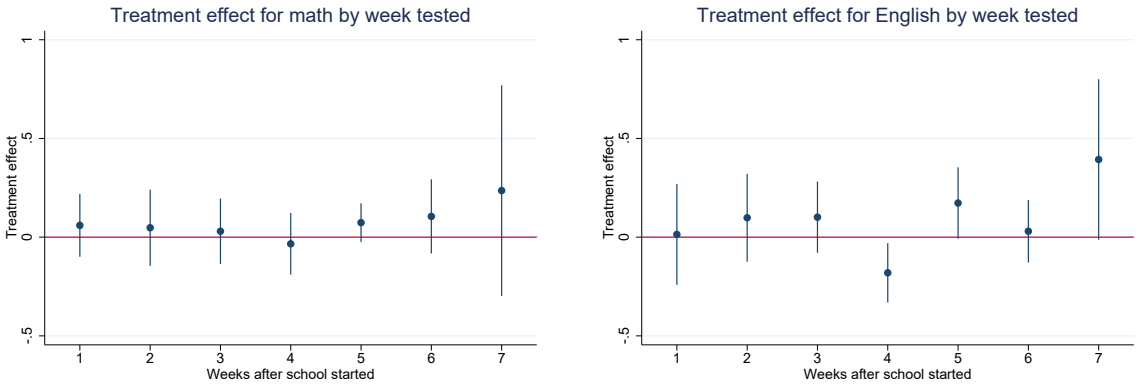

Figure 5: Treatment effect by week of data collection. The sample is the final treatment and control allocation.

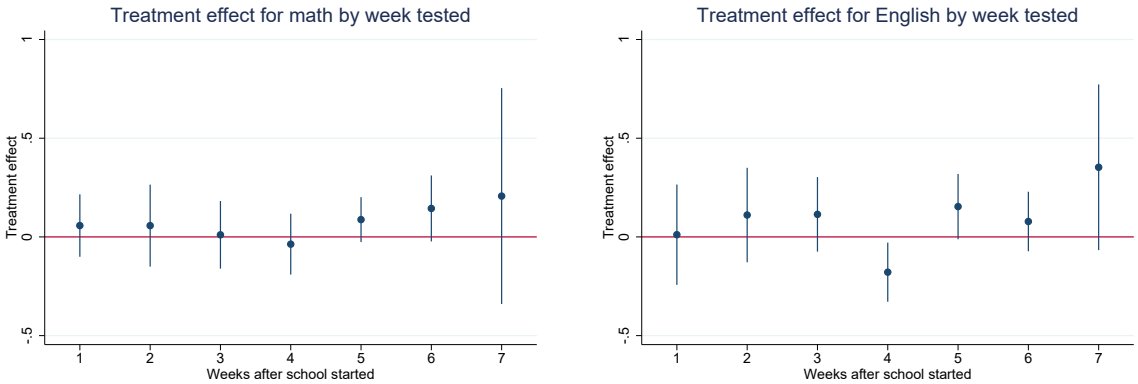

Figure 6: Treatment effect by week of data collection. The sample is the original treatment and control allocation.