Online appendix

This appendix provides supplementary material for Romero, Sandefur, and Sandholtz (2017, Center for Global Development working paper #462), "Can outsourcing improve Liberia's schools?" The main text is available online at https://www.cgdev.org/publication/partnership-schools-for-liberia.

A Extra tables and figures

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

	(1) RCT (Treatment and control)	(2) Other public schools	(3) Difference
Students: ECE	142.68	112.71	29.97***
	(5.46)	(1.39)	(5.77)
Students: Primary	151.55	132.38	19.16*
	(9.62)	(2.95)	(10.18)
Students	291.91	236.24	55.67***
	(11.36)	(3.46)	(12.15)
Classrooms per 100 students	1.17	0.80	0.37***
	(0.12)	(0.04)	(0.13)
Teachers per 100 students	3.04	3.62	-0.58**
	(0.10)	(0.26)	(0.28)
Textbooks per 100 students	99.21	102.33	-3.12
	(7.08)	(3.49)	(7.88)
Chairs per 100 students	20.71	14.13	6.58***
	(2.08)	(1.05)	(2.38)
Food from Gov or NGO	0.36	0.30	0.06
	(0.04)	(0.01)	(0.04)
Solid building	0.36	0.28	0.08*
	(0.04)	(0.01)	(0.04)
Water pump	0.62	0.45	0.17***
	(0.04)	(0.01)	(0.04)
Latrine/toilet	0.85	0.71	0.14***
	(0.02)	(0.01)	(0.03)
Distance to MoE (in KM)	153.25	186.99	-33.74***
	(7.32)	(2.71)	(10.41)
Observations	185	2,420	2,605

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. ECE = Early childhood education. MOE= Ministry of Education. Authors' calculations based on 2015/2016 EMIS data.

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

Figure A.1: Timeline

Research Activities	Year	Month	Intervention Activities
	2016	Jun	Operator selection
Randomization		Jul	
		Aug	
Baseline		Sep	School year begins
		Oct	
		Nov	
		Dec	
	2017	Jan	
		Feb	
		Mar	
		Apr	
Midline		May	Year 2 decisions
		Jun	
		Jul	
		Aug	
		Sep	
		Oct	
		Nov	
		Dec	
	2019	Jan	
		Feb	
Endline		Mar	
		Apr	

Note: Bridge signed its MOU with the Government of Liberia in March 2016, and thus started preparing for the program earlier than other contractors.

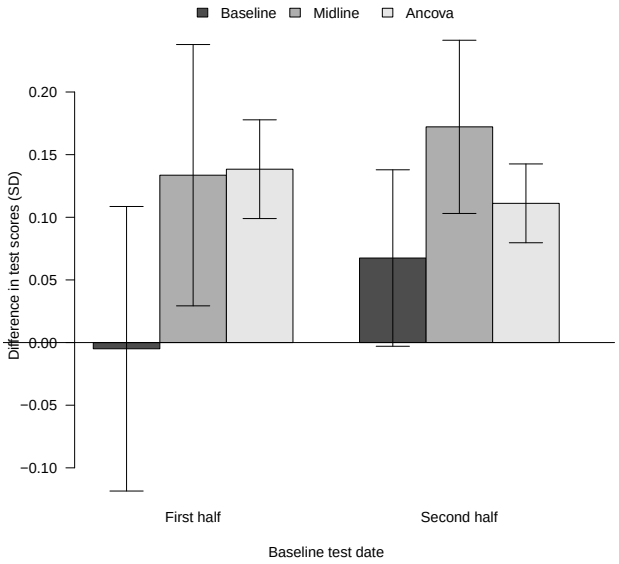
Table A.2: Balance table: Difference in characteristics (EMIS data) between treatment and control schools, pre-treatment year (2015/2016)

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Students: ECE	136.72	148.51	11.79	11.03
	(7.40)	(8.01)	(10.91)	(9.74)
Students: Primary	143.96	159.05	15.10	15.68
	(9.03)	(16.94)	(19.19)	(16.12)
Students	277.71	305.97	28.26	27.56
	(13.03)	(18.51)	(22.64)	(19.46)
Classrooms per 100 students	1.13	1.21	0.09	0.08
	(0.17)	(0.17)	(0.24)	(0.23)
Teachers per 100 students	2.99	3.08	0.09	0.09
	(0.14)	(0.16)	(0.21)	(0.18)
Textbooks per 100 students	95.69	102.69	7.00	7.45
	(9.95)	(10.13)	(14.19)	(13.74)
Chairs per 100 students	22.70	18.74	-3.96	-4.12
	(3.42)	(2.39)	(4.17)	(3.82)
Food from Gov or NGO	0.36	0.36	-0.01	-0.01
	(0.06)	(0.05)	(0.08)	(0.05)
Solid building	0.33	0.39	0.06	0.06
	(0.05)	(0.05)	(0.07)	(0.06)
Water pump	0.67	0.56	-0.11	-0.12*
	(0.05)	(0.05)	(0.07)	(0.06)
Latrine/toilet	0.86	0.85	-0.01	-0.01
	(0.03)	(0.04)	(0.05)	(0.05)
Distance to MoE (in KM)	153.87	152.64	-1.23	-1.00
	(10.39)	(10.38)	(14.69)	(3.06)
Observations	92	93	185	185

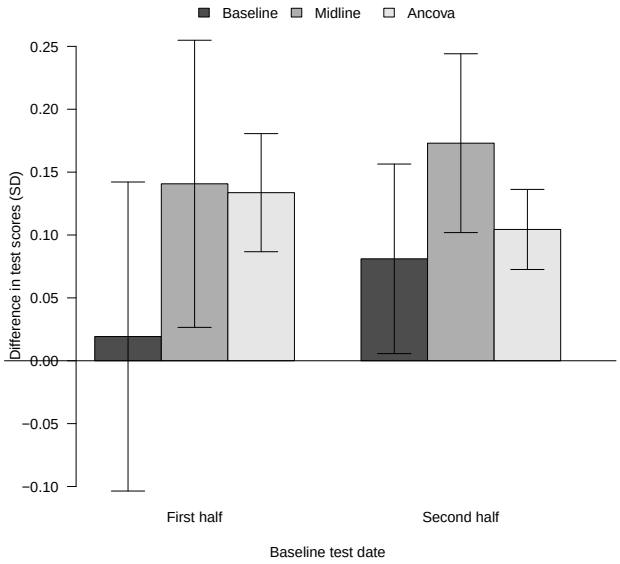
This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2), as well as the difference between treatment and control (Column 3), and the difference taking into account the randomization design (i.e., including "'pair"' fixed effects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

Figure A.2: Treatment effects by month tested at baseline

(a) Math



(b) English



Note: The panel on the left shows results for math test scores at baseline, while the panel on the right shows English test score results at baseline.

Table A.3: Heterogeneity by student characteristics

	Male (1)	Top wealth quartile (2)	Bottom wealth quartile (3)
Treatment	0.20***	0.18***	0.17***
	(0.047)	(0.035)	(0.035)
Treatment × covariate	-0.030	0.031	0.060
	(0.067)	(0.067)	(0.050)
No. of obs.	3,498	3,498	3,498

Each column shows the interaction of a different covariate with treatment. 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 A.4: ITT and ToT effect

	Difference (Controls)			ANCOVA		
	Math (1)	English (2)	Abstract (3)	Math (4)	English (5)	Abstract (6)
Panel A: ITT						
Treatment	0.18***	0.18***	0.045	0.14***	0.13***	0.031
	(0.034)	(0.031)	(0.038)	(0.023)	(0.021)	(0.036)
No. of obs.	3,498	3,498	3,498	3,498	3,498	3,498
Panel B: ToT						
Treatment	0.23***	0.22***	0.058	0.18***	0.17***	0.040
	(0.042)	(0.038)	(0.047)	(0.029)	(0.026)	(0.046)
No. of obs.	3,498	3,498	3,498	3,498	3,498	3,498

The treatment-on-the-treated treatment effect is estimated using the assigned treatment as an instrument for whether the student is in fact enrolled in a PSL school during the 2016/2017 academic year.
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 A.5: Different measures of student ability

	Difference	Difference (F.E.)	Difference (F.E. + Controls)	Difference (ANCOVA)
	(1)	(2)	(3)	(4)
Panel A: Base IRT model				
English	0.18**	0.18***	0.17***	0.13***
	(0.08)	(0.04)	(0.03)	(0.02)
Math	0.19***	0.19***	0.18***	0.14***
	(0.06)	(0.04)	(0.03)	(0.02)
Panel B: Base IRT model standarized by grade				
English	0.17*	0.19***	0.23***	0.17***
	(0.09)	(0.06)	(0.04)	(0.03)
Math	0.18**	0.21***	0.22***	0.17***
	(0.08)	(0.04)	(0.04)	(0.03)
Panel C: IRT model per grade				
English	0.19**	0.21***	0.24***	0.19***
	(0.09)	(0.06)	(0.04)	(0.03)
Math	0.20**	0.24***	0.25***	0.20***
	(0.08)	(0.05)	(0.04)	(0.03)
Panel D: Base PCA				
English	0.18**	0.17***	0.16***	0.12***
	(0.08)	(0.04)	(0.03)	(0.02)
Math	0.20***	0.20***	0.20***	0.16***
	(0.06)	(0.04)	(0.04)	(0.02)
Panel E: Base PCA standarized by grade				
English	0.16*	0.17***	0.19***	0.14***
	(0.09)	(0.05)	(0.04)	(0.03)
Math	0.19**	0.23***	0.25***	0.20***
	(0.08)	(0.05)	(0.05)	(0.03)
Panel F: PCA per grade				
English	0.16*	0.17***	0.19***	0.14***
	(0.09)	(0.05)	(0.04)	(0.03)
Math	0.19**	0.23***	0.25***	0.20***
	(0.08)	(0.05)	(0.05)	(0.03)
Panel G: % correct answers				
English	3.30**	3.05***	2.90***	2.24***
	(1.41)	(0.75)	(0.55)	(0.37)
Math	3.69***	3.78***	3.68***	2.96***
	(1.14)	(0.73)	(0.64)	(0.42)
Observations	3,498	3,498	3,498	3,498

Column 1 shows the simple difference between treatment and control, Column 2 the difference taking into account the randomization design–i.e., including "'pair"' fixed effects–, Colum 3 the difference taking into account other student and school controls, and the difference using an ANCOVA style specification that controls for baseline test scores is shown in Column 4. Panel A uses our default IRT model and normalizes test scores using the same mean and standard deviation across all grades. Panel B uses the same IRT model as panel B, but normalizes test scores using a different mean and standard deviation for each grade. Panel C estimates a different IRT model for each grade. Panel D estimates students' ability as the first component from a principal component analysis (PCA), and normalizes test scores using a common mean and standard deviation across al grades. Panel E uses the same model as panel D but normalizes test scores using a different mean and standard deviation per grade. Panel F performs a different principal component analysis for each grade. Panel G calculates the percentage of correct responses.

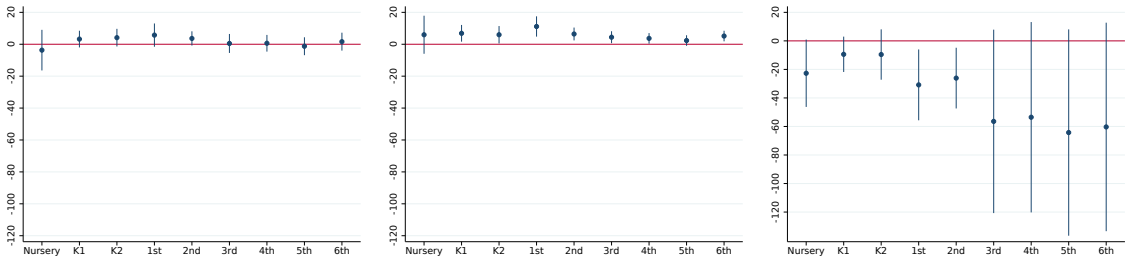
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Student selection

	(1) Same school	(2) Same school	(3) Same school
Treatment	0.053	0.014	0.022
	(0.081)	(0.026)	(0.019)
Treatment × Age	-0.0037		
	(0.0064)		
Treatment × Male		-0.017	
		(0.029)	
Treatment × Asset Index (PCA)			-0.0094
			(0.011)
No. of obs.	3,493	3,493	3,291

Standard errors are clustered at the school level. The sample is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.3: Treatment effect on enrolment by grade



(a) All schools (b) Non-constrained schools (c) Constrained schools

Note: These figures show the difference in enrollment (2016/2017 compared to the 2015/2016 academic year) by grade. The dots represent point estimates, while the bars represent 95% confidence intervals. Panel A.3a shows the effect across all schools. Panel A.3b shows in non-constrained school-grades, and Panel A.3c shows in constrained school-grades.

Table A.7: Intensive margin effect on teacher attendance and classroom observation with Lee bounds

	(1) Control	(2) Treatment	(3) Difference (F.E)	(4) Difference	(5) 90% CI (bounds)
Panel A: Spot check (N = 930)					
% On schools campus	52.29	68.15	15.87***	14.23***	2.81
	(2.33)	(2.15)	(4.44)	(3.75)	28.09
% In classroom	40.96	50.96	10.00**	10.02**	-1.10
	(2.30)	(2.31)	(4.77)	(3.86)	24.36
B: Classroom observation (N = 133)					
Active Instruction (% class time)	28.73	37.86	9.12*	8.79*	-5.66
	(3.71)	(3.16)	(4.88)	(4.94)	23.26
Passive Instruction (% class time)	13.10	16.19	3.09	4.92	-6.86
	(2.40)	(1.89)	(3.05)	(3.43)	10.32
Classroom management (% class time)	10.70	21.07	10.37***	9.30***	-0.13
	(1.80)	(2.30)	(2.92)	(3.44)	18.58
Teacher off-task (% class time)	47.46	24.88	-22.58***	-23.02***	-43.48
	(4.87)	(3.58)	(6.05)	(6.64)	-9.76
Student off-task (% class time)	58.45	55.24	-3.21	-4.82	-18.80
	(4.13)	(3.42)	(5.36)	(5.12)	14.37
Panel C: Inputs (N = 133)					
Number of seats	20.38	20.46	0.07	0.51	-8.12
	(1.71)	(1.50)	(2.27)	(2.02)	6.12
% with students sitting on the floor	4.48	2.44	-2.04	-3.26	-7.92
	(2.55)	(1.71)	(3.07)	(2.28)	2.92
% with chalk	77.61	96.34	18.73***	17.93***	10.08
	(5.13)	(2.09)	(5.54)	(5.91)	29.45
% of students with textbooks	17.16	36.32	19.15***	24.56***	-5.96
	(4.23)	(4.74)	(6.35)	(6.40)	36.83
% of students with pens/pencils	78.46	88.41	9.95**	8.70*	1.48
	(3.74)	(2.20)	(4.34)	(4.42)	23.30

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Column 5 has the 90% confidence interval using Lee (2009) bounds. Panel A has the spot check using the EMIS data (2015/2016) information on teachers as a baseline, and treating teachers that no longer teach at school as attriters. Panel B has the classroom observation information without imputing values for schools not in session during our visit, and treating the missing information as attrition. 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 A.8: Treatment effect on school's good practices

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Has a physical enrollment log	0.80	0.90	0.10*	0.10*
	(0.04)	(0.03)	(0.05)	(0.05)
Enrollment log has name	0.82	0.89	0.08	0.08
	(0.04)	(0.03)	(0.05)	(0.05)
Enrollment log has grade	0.84	0.94	0.10**	0.10**
	(0.04)	(0.03)	(0.05)	(0.05)
Enrollment log has age	0.64	0.65	0.00	-0.00
	(0.05)	(0.05)	(0.07)	(0.07)
Enrollment log has gender	0.83	0.89	0.07	0.06
	(0.04)	(0.03)	(0.05)	(0.05)
Enrollment log has contact information	0.13	0.26	0.13**	0.13**
	(0.04)	(0.05)	(0.06)	(0.06)
Enrollment log is clean and neat	0.26	0.39	0.13*	0.13*
	(0.05)	(0.05)	(0.07)	(0.07)
Has official time table	0.89	0.98	0.09**	0.09***
	(0.03)	(0.02)	(0.04)	(0.03)
Official time table is posted	0.70	0.84	0.14**	0.14**
	(0.05)	(0.04)	(0.06)	(0.06)
Has a PTA	0.98	0.99	0.01	0.01
	(0.02)	(0.01)	(0.02)	(0.02)
Principal has PTA head's number at hand	0.26	0.41	0.15**	0.15**
	(0.05)	(0.05)	(0.07)	(0.06)
Has record of expenditures	0.09	0.14	0.05	0.05
	(0.03)	(0.04)	(0.05)	(0.05)
Has written budget	0.22	0.26	0.04	0.04
	(0.04)	(0.05)	(0.06)	(0.06)
Observations	92	93	185	185

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). 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 A.9: Treatment effect on household expenditure

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Fees (USD/year)	8.03	5.69	-2.34**	-2.93***
	(0.42)	(0.41)	(0.96)	(0.60)
Tutoring (USD/year)	0.39	0.34	-0.04	-0.03
	(0.06)	(0.05)	(0.09)	(0.08)
Textbooks (USD/year)	0.85	0.62	-0.24*	-0.22**
	(0.07)	(0.06)	(0.13)	(0.09)
Copy books (USD/year)	1.09	1.02	-0.07	-0.09
	(0.08)	(0.08)	(0.14)	(0.13)
Pencils (USD/year)	2.95	3.25	0.30	0.21
	(0.12)	(0.12)	(0.30)	(0.16)
Uniform (USD/year)	11.42	9.27	-2.15***	-1.95***
	(0.22)	(0.26)	(0.65)	(0.43)
Food (USD/year)	46.27	43.78	-2.49	-1.49
	(3.27)	(2.89)	(6.97)	(3.91)
Other (USD/year)	3.05	3.42	0.38	0.32
	(0.18)	(0.18)	(0.34)	(0.27)
Observations	520	596	1,116	1,116

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). 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 A.10: Treatment effect on household's engagement

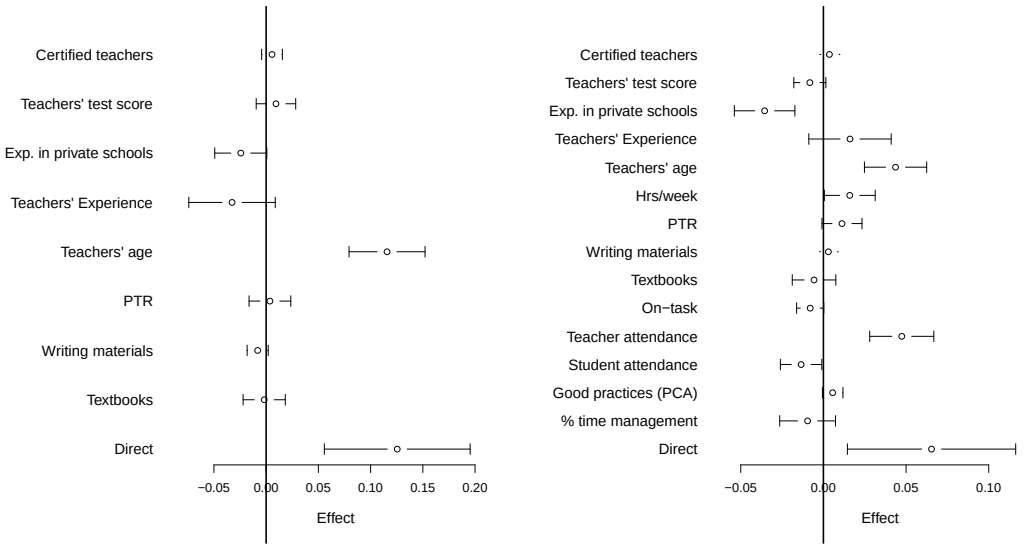
	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Attended school meeting	0.77	0.76	-0.01	0.03
	(0.02)	(0.02)	(0.04)	(0.02)
Donated cash	0.11	0.13	0.02	-0.00
	(0.01)	(0.01)	(0.02)	(0.02)
Donated in-kind	0.04	0.03	-0.01	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Donated work	0.14	0.13	-0.01	-0.00
	(0.02)	(0.01)	(0.03)	(0.02)
Helped with homework	0.61	0.58	-0.03	-0.03
	(0.02)	(0.02)	(0.04)	(0.03)
Observations	543	620	1,163	1,163

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). 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 A.11: 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
Travel time 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	

Figure A.4: Direct and causal mediation effects



(a) Inputs (b) Inputs & Management

Note: Direct (β_5) and mediation effects ($\beta_4 \times \theta_5$) for all the possible mediators. Note that the point estimates within the same panel are directly comparable to each other. Point estimates and 90% confidence intervals are plotted. Panel A.4a shows treatment effects allowing only change in inputs as mediators. Panel A.4b shows treatment effects allowing change in inputs and in the use of inputs as mediators.

Table A.12: Raw (fully experimental) treatment effects by contractor

	(1) BRAC	(2) Bridge	(3) YMCA	(4) MtM	(5) Omega	(6) Rising	(7) St. Child	(8) Stella M
Panel A: Student test scores								
English (standard deviations)	0.19**	0.27***	0.57**	0.19	-0.07	0.34	0.24*	-0.22
	(0.10)	(0.09)	(0.27)	(0.23)	(0.11)	(0.24)	(0.13)	(0.23)
Math (standard deviations)	0.09	0.38***	0.26	0.19	-0.05	0.41*	0.29**	-0.16
	(0.10)	(0.09)	(0.26)	(0.22)	(0.11)	(0.24)	(0.13)	(0.23)
Composite (standard deviations)	0.14	0.34***	0.37	0.18	-0.07	0.41*	0.28**	-0.18
	(0.10)	(0.09)	(0.27)	(0.22)	(0.11)	(0.24)	(0.13)	(0.23)
Panel B: Changes to the pool of teachers								
% teachers dismissed	-6.81	50.46***	21.20	14.11	-8.04	-5.80	-3.03	-11.16
	(6.45)	(6.30)	(14.40)	(11.79)	(6.84)	(12.77)	(8.52)	(14.40)
% new teachers	39.63***	63.11***	62.48**	74.05***	24.18*	24.36	41.04**	-20.18
	(12.21)	(11.93)	(27.25)	(22.32)	(12.94)	(24.17)	(16.12)	(27.25)
Age in years (teachers)	-5.04***	-10.86***	3.25	-11.23***	-5.43***	-10.79***	-5.77**	-4.53
	(1.93)	(2.01)	(4.30)	(3.52)	(2.04)	(3.82)	(2.54)	(4.31)
Test score in standard deviations (teachers)	0.03	0.37**	-0.59	0.48	0.19	0.18	0.32	0.17
	(0.17)	(0.16)	(0.38)	(0.31)	(0.17)	(0.33)	(0.22)	(0.38)
Panel C: Enrollment and access								
Δ Enrollment	36.38	-27.91	50.73	-28.65	47.43	16.96	42.08	38.17
	(35.63)	(33.41)	(79.49)	(65.15)	(36.61)	(70.43)	(46.94)	(79.53)
Δ Enrollment (constrained grades)	0.00	-40.04***	0.00	0.00	4.47	0.00	47.66	0.00
	(0.00)	(10.60)	(0.00)	(0.00)	(38.70)	(0.00)	(64.26)	(0.00)
Student attendance (%)	20.12***	6.96	13.36	37.54***	7.16	29.32**	20.23**	5.23
	(6.46)	(6.05)	(14.40)	(11.80)	(6.63)	(12.76)	(8.50)	(14.41)
% of students still in any school	1.22	4.72	4.51	-2.00	4.97	2.48	3.91	6.27
	(4.57)	(4.29)	(12.55)	(10.61)	(5.12)	(11.27)	(6.29)	(10.86)
% of students still in same school	0.83	4.55**	-0.77	1.05	1.68	3.63	-0.71	1.10
	(2.21)	(2.08)	(6.08)	(5.14)	(2.48)	(5.46)	(3.05)	(5.26)
Panel D: Satisfaction								
% satisfied with school (parents)	11.59	13.75*	18.10	0.74	0.33	4.38	-5.23	29.65*
	(7.29)	(7.13)	(16.28)	(13.34)	(7.54)	(14.44)	(9.62)	(16.28)
% students that think school is fun	5.81	2.03	20.92	1.16	4.65	9.66	3.14	-17.61
	(4.87)	(4.58)	(13.38)	(11.31)	(5.45)	(12.02)	(6.71)	(11.58)
Observations	40	45	8	12	38	10	24	8

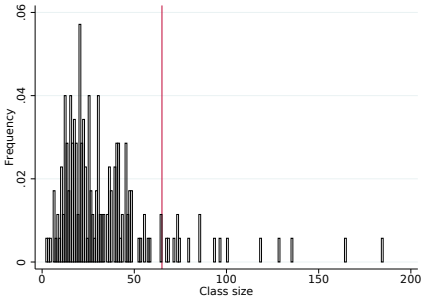
This table presents the raw treatment effect for each contractor on different outcomes. The sample is the original treatment and control allocation. Note that the estimates for each contractor are *not* comparable to each other without further assumptions, and thus we do intentionally do not include a test of equality. 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 A.13: Descriptive statistics by contractor and treatment

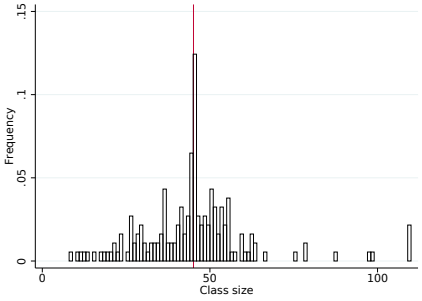
(1) Contractor	(2) Treatment	(3) Schools	(4) Teachers	(5) Teachers	(6) Teachers	(7) Teachers	(8) Enrollment	(9) Enrollment	(10) Enrollment	(11) Enrollment in constrained classes	(12) Enrollment in constrained classes	(13) Enrollment in constrained classes
			2015/2016	2016/2017	Dismissed	New	Classes	2015/2016	2016/2017	Constrained classes	2015/2016	2016/2017
BRAC	0	20	141	148	41	48	180	5,694	5,107	10	780	703
BRAC	1	20	141	209	33	101	180	5,684	5,872	11	1,130	1,138
Bridge	0	22	177	174	38	35	198	7,110	6,610	61	3,969	3,648
Bridge	1	23	236	212	174	150	207	9,788	8,282	72	6,909	3,475
YMCA	0	4	20	22	1	3	36	729	727	2	142	120
YMCA	1	4	27	40	6	19	36	908	1,068	2	217	238
MtM	0	6	52	41	21	10	54	1,140	1,312	2	155	167
MtM	1	6	46	64	20	38	54	1,145	1,223	2	171	159
Omega	0	19	132	130	33	31	171	4,895	5,200	12	1,255	1,232
Omega	1	19	151	196	26	71	171	5,764	6,841	19	1,953	2,446
Rising	0	5	47	43	23	19	45	1,209	1,308	2	202	185
Rising	1	5	36	47	11	22	45	918	1,134	1	87	89
St. Child	0	12	88	68	29	9	108	3,094	2,794	7	738	557
St. Child	1	12	81	100	22	41	108	3,351	3,506	9	877	797
Stella M	0	4	20	20	8	8	36	765	683	1	73	45
Stella M	1	4	31	27	9	5	36	958	978	3	213	192

This tables has total numbers of teachers and students in treatment and control schools for each operator. Note that teachers in 2015/2016 are taken from the EMIS data, while teachers in 2016/2017 are taken from our first year follow up. Dismissed is the number of teachers in the 2015/2016 EMIS data, that are not working in the school at the end of the 2016/2017 academic year. "New" is the number of teachers working in the school at the end of the 2016/2017 academic year that are not in the 2015/2016 EMIS data. "Constrained classes" are those with more students in 2015/2016 than the class size cap.

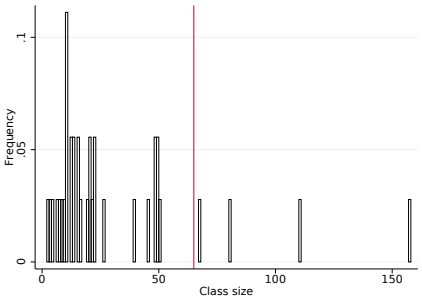
Figure A.5: Class sizes and class caps



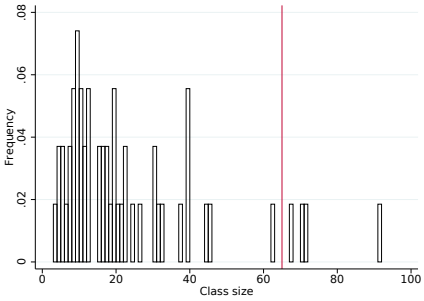
(a) BRAC



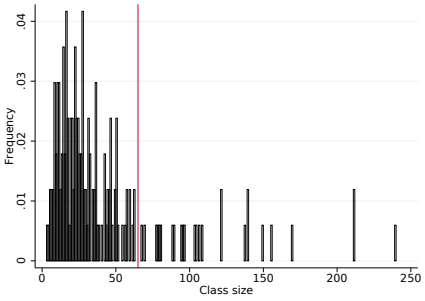
(b) Bridge



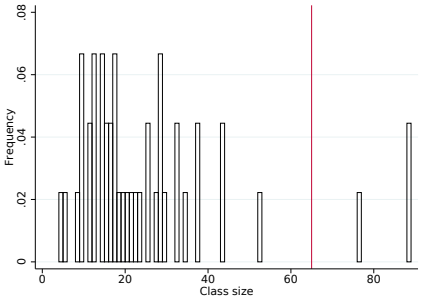
(c) YMCA



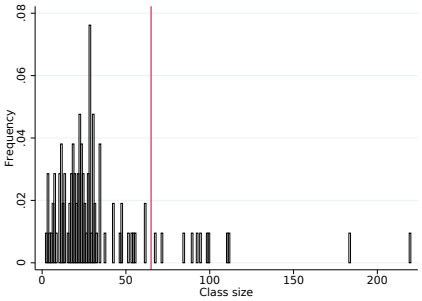
(d) More than me



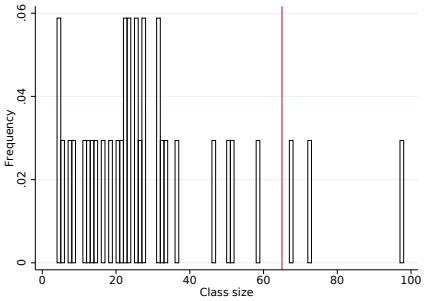
(e) Omega schools



(f) Rising



(g) St. Child



(h) Stella M

Note: These figures show the distribution of class sizes in treatment schools during the 2016/2017 academic year, as well as the class cap for each contractor. Note that the cap for all contractors is 65 students, except for Bridge that has a cap of 45.

B School competition

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 giving parents freedom of school choice. The standard argument is that charter schools will be more responsive 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). Furthermore, 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.

To measure school competition, we calculate the number of schools within a 5 KM radius (as pre-committed to in the pre-analysis plan). Since we do not experimentally vary the level of competition, we rely on sampling variation generated by the randomization assignment and control for baseline school characteristics and their interactions with treatmnet. Table B.1 shows that test scores, enrollment, and attendance figures are statistically indistinguishable from each other in schools facing competition below and above the median.[54] Figure B.1 shows that this is also true if we let the treatment effect vary in a more flexible way. These results suggest that competition is not relevant to the PSL program.

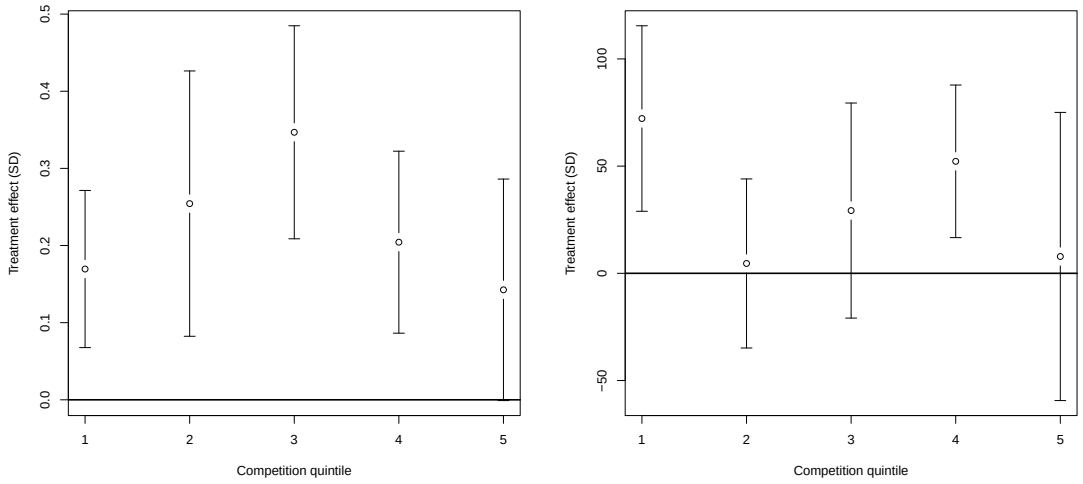
Table B.1: Competition, test scores and enrollment

	Test scores			Access	
	Math	English	Composite	Δ enrollment	Student attendance
Competition=0 × Treatment	0.20***	0.21***	0.21***	36.99**	15.09***
	(0.06)	(0.05)	(0.05)	(16.95)	(4.98)
Competition=1 × Treatment	0.20***	0.23***	0.21***	36.00**	14.24***
	(0.06)	(0.05)	(0.05)	(18.23)	(3.96)
No. of obs.	3,468	3,468	3,468	183	183
C-NC	-0.01	0.02	0.00	-0.99	-0.84
p-value (H_0:C-NC=0)	0.91	0.76	0.99	0.97	0.90

Treatment effect for schools with and without competition. Standard errors are clustered at the school level. The sample is the original treatment and control allocation. C-NC is the difference between the treatment effect for school with competition (C) and without (NC). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[54] To make the effects comparable we estimate the treatment effects for schools with and without competition at the average level of school and student covariates in our sample.

Figure B.1: Treatment effect by deciles of competition (number of schools in in a 5 km radius)



(a) Test scores (composite)

(b) Enrollment

Note: Treatment effect by deciles of competition (number of schools in in a 5 km radius). Bars represent 90% and 95% confidence intervals (thick lines and thin lines, respectively). Panel B.1b shows the treatment effect on test scores. Panel B.1b shows the treatment effect on enrollment. Original treatment assignment.

C Satisfaction and support for the PSL program

For a government program to be politically viable it needs the support of those affected by it. The PSL program has met with resistance from teacher unions and international organizations.[55] In Table C.1 we show data collected independently by us of support and satisfaction of the PSL program from students, parents and teachers.

There are three main messages from the data in this table. First, students are happier in PSL than in traditional public schools (measured by whether they think going to school is fun or not). Second, households with children in PSL schools (enrolled in 2015/2016) are 7.4 percentage points (p-value .022) more likely to be satisfied with the education their children are receiving. Additionally, most households, even in the control group, would prefer that contractors manage more schools the following year (87% of households overall) and would rather send their children to a school managed by an contractor than to a traditional public school (72% of households overall). Third, despite any (statistically significant) difference in the satisfaction of teachers across treatment and control schools, most teachers, even in control schools, would rather work in a school managed by a contractor (64% of teachers overall) and would prefer that contractors managed more schools the following year (85% of teachers overall).

[55]The Liberian government's announcement of the PSL program generated international headlines from the BBC to the New York Times about "outsourcing" and "privatization" (The New York Times, 2016; BBC Africa, 2016; Vox World, 2016; Foreign Policy, 2016; Mail & Guardian Africa, 2016b, 2016a), and even condemnation from a UN Special Rapporteur that Liberia was abrogating its responsibilities under international law (OHCHR, 2016).

Table C.1: Student, household and teacher satisfaction and opinion

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Students (N = 3,498)				
School is fun (%)	52.49	58.22	5.73**	5.68**
	(1.61)	(1.61)	(2.28)	(2.45)
Panel B: Households (N = 185)				
% satisfied with school	67.47	74.89	7.43**	7.45**
	(2.50)	(2.00)	(3.20)	(3.23)
% have heard of PSL	14.34	17.72	3.38	3.36
	(1.68)	(1.61)	(2.33)	(2.21)
% have heard of contractor	23.87	54.45	30.58***	30.64***
	(2.54)	(3.27)	(4.14)	(3.93)
% thinks contractor should manage more schools	81.69	90.66	8.97*	11.55**
	(4.49)	(1.94)	(4.89)	(4.91)
% would prefer to send child to contractor's school	61.95	79.63	17.68***	18.07**
	(5.39)	(2.88)	(6.11)	(7.07)
Panel C: Teachers (N = 185)				
% satisfied with life	79.28	78.87	-0.41	-0.63
	(2.19)	(2.23)	(3.10)	(3.57)
% would choose teaching as a career	88.23	90.74	2.51	1.99
	(1.86)	(1.33)	(2.32)	(2.56)
% work a second job	23.77	16.27	-7.50**	-7.45**
	(2.69)	(2.11)	(3.45)	(3.74)
Job satisfaction index (PCA)	-0.14	0.05	0.18	0.21
	(0.09)	(0.09)	(0.13)	(0.14)
% have heard of PSL	28.43	64.81	36.38***	35.19***
	(2.82)	(3.02)	(4.50)	(4.03)
% have heard of operator	39.76	93.99	54.23***	54.76***
	(3.80)	(1.85)	(4.53)	(4.28)
% would rather work at an operator school	43.12	70.99	27.87***	21.93***
	(5.01)	(2.37)	(6.00)	(5.98)
% thinks operator should manage more schools	81.15	85.80	4.65	1.46
	(4.39)	(1.92)	(4.97)	(5.15)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Standard errors are clustered at the school level. The sample is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D What "managing" a school means in practice

In this section we offer two pieces of information that readers may find useful in interpreting the results. First, we ask each contractor for a brief statement of what school management entails for them. Addition-

ally, we show evidence from teacher data data on contractor activities in each school and community. Note that our pair-matched design allowed us to ask contractor-specific questions of control schools. Table D.1 shows teacher reports on contractor activities. First, note that no contractor visited a control school on a a regular basis, nor did they provide control schools with inputs. On the other hand, only 62% of treatment schools received contractor visits on a regular basis (recall that there is non-compliance in our sample). Managing a school does seem to entail a wide range of activities. Teachers report that contractors provided hard inputs (textbooks, copybooks, tablets, and repairs) and soft inputs (training and community meetings). The two most likely activities during the last visit from the contractor entailed either checking attendance and school records and/or observing teaching practices.

Table D.1: Contractor activities, according to teachers

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: General opinion (N = 1,097)				
Heard of PSL	0.28	0.65	0.36***	0.35***
	(0.02)	(0.02)	(0.04)	(0.03)
Heard of operator	0.40	0.94	0.54***	0.55***
	(0.02)	(0.01)	(0.05)	(0.03)
Operator staff visits at least once a week	0.00	0.64	0.64***	0.62***
	(0.00)	(0.02)	(0.04)	(0.04)
Operator support rating (0-100)	15.08	67.30	52.22***	53.48***
	(2.37)	(1.19)	(3.88)	(3.64)
Panel B: What do contractors provide? (N = 803)				
Teacher guides (or teacher manuals)	0.02	0.74	0.72***	0.77***
	(0.01)	(0.02)	(0.03)	(0.03)
Textbooks	0.03	0.88	0.85***	0.87***
	(0.01)	(0.01)	(0.02)	(0.03)
Copybooks	0.01	0.58	0.56***	0.46***
	(0.01)	(0.02)	(0.05)	(0.05)
Paper	0.01	0.69	0.68***	0.69***
	(0.01)	(0.02)	(0.04)	(0.04)
Teacher training	0.02	0.80	0.77***	0.81***
	(0.01)	(0.02)	(0.03)	(0.03)
School repairs	0.01	0.34	0.32***	0.37***
	(0.01)	(0.02)	(0.04)	(0.03)
Organize community meetings	0.02	0.62	0.60***	0.65***
	(0.01)	(0.02)	(0.04)	(0.03)
Food programs	0.02	0.03	0.01	0.01
	(0.01)	(0.01)	(0.02)	(0.01)
Computers, tablets, electronics	0.01	0.45	0.44***	0.58***
	(0.01)	(0.02)	(0.06)	(0.05)
Panel C: What did contractors do during their last visit (N = 715)				
Check attendance and collect records	0.10	0.50	0.40***	0.28***
	(0.03)	(0.02)	(0.06)	(0.06)
Observe teaching practices and give suggestions	0.13	0.63	0.50***	0.45***
	(0.04)	(0.02)	(0.06)	(0.06)
Provide/deliver educational materials	0.01	0.26	0.25***	0.22***
	(0.01)	(0.02)	(0.03)	(0.04)
Ask students questions to test learning	0.09	0.30	0.21***	0.10**
	(0.03)	(0.02)	(0.06)	(0.05)
Monitor other school-based government programs	0.01	0.08	0.07***	0.09***
	(0.01)	(0.01)	(0.02)	(0.03)
Meet with principal	0.30	0.42	0.11	0.08
	(0.05)	(0.02)	(0.08)	(0.08)
Meet with PTA committee	0.01	0.11	0.10***	0.10**
	(0.01)	(0.01)	(0.02)	(0.04)
Monitor health/sanitation issues	0.00	0.07	0.07***	0.06***
	(0.00)	(0.01)	(0.02)	(0.02)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Standard errors are clustered at the school level. The sample is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Tracking and attrition

A potential issue with our sampling strategy is differential attrition at baseline and midline. At baseline, enumerators were instructed to sample 20 students from the 2015/2016 enrollment logs, track them, and test them. However, if a student had moved to another village, had died, or was impossible to track, the enumerators were instructed to sample another student. Thus, even at baseline an endogenous sampling problem arises if treatment makes students easier or harder to track in combination with enumerator shrinkage. To mitigate this issue, enumerators participated in additional training on tracking and its importance and were provided with a generous amount of tracking time both at baseline and midline. Students were tracked to their homes and tested there when not available at school. As Table E.1 shows, we have no reason to believe that this issue arose for either the baseline or midline. Panel A shows that the effort required to track students was different between treatment and control (is easier to track students at the school), yet the total number of students sampled, to obtain a sample of 20 students, is balanced between treatment and control. Panel B shows that attrition from our original sample is also balanced between treatment and control (and is low overall, below 4%).

Table E.1: Tracking

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Number of students sampled	24.6	24.8	0.14	0.047
	(0.54)	(0.61)	(0.82)	(0.81)
Found at the school	16.7	18.2	1.50***	1.566***
	(0.50)	(0.25)	(0.55)	(0.55)
Found at home	2.91	1.73	-1.18**	-1.223**
	(0.42)	(0.23)	(0.48)	(0.47)
Interviewed	19.5	19.8	0.31	0.332
	(0.23)	(0.089)	(0.25)	(0.26)
Observations	90	88	178	171

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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 the average number of students we sampled (and tried to track), the number of students we were able to track at the assigned school or at home, and the total number of students we tracked and found at baseline. Standard errors are clustered at the school level. The sample is the original treatment and control allocation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

F Test design

Most modules follow the Early Grade Reading Assessment (EGRA), Early Grade Mathematics Assessment (EGMA), Uwezo and Trends in International Mathematics and Science Study (TIMSS) assessments. For baseline, the test had a module for each one of the following skills: object identification (like the Peabody Picture Vocabulary Test), letter reading (adapted from EGRA), word reading (adapted from

EGRA), a preposition module, reading comprehension (adapted from Uwezo), listening comprehension (adapted from EGRA), counting (adapted from Uwezo), number discrimination (adapted from Uwezo), number identification (adapted from EGMA), addition (adapted from Uwezo and EGMA), subtraction (adapted from Uwezo and EGMA), multiplication (adapted from Uwezo and EGMA), division (adapted from Uwezo and EGMA), shape identification, fractions, and word problems in mathematics.

At follow-up, the test did not include the preposition, the shape identification, and the fractions modules. These modules were excluded given the low variation in responses at baseline and to make space for new modules. Instead, it included letter, word and number dictation, and a verb and a pronoun module. Additionally, we included some "conceptual" questions from TIMSS released items (items M031317 and M031316) that do not resemble the format of standard textbook exercises but rather test knowledge in an unfamiliar way. The number identification module remained exactly the same between baseline at follow-up (to allow us to have absolute learning curves on these two items), while every other module was different. At follow-up the word and number identification module were identical to the EGRA/EGMA assessments used in Liberia before (for comparability with other impact evaluations taking place in Liberia, most notably USAID's reading program (Piper & Korda, 2011) and the LTTP program (King et al., 2015)), while at baseline they were different. Two of the reading comprehension questions were taken from the Pre-Pirls released items (L11L01C and L11L02M) and one of the word problems was taken from TIMSS released items (M031183) for the follow-up. Finally, we added a Raven's style module to measure the students' abstract thinking abilities.

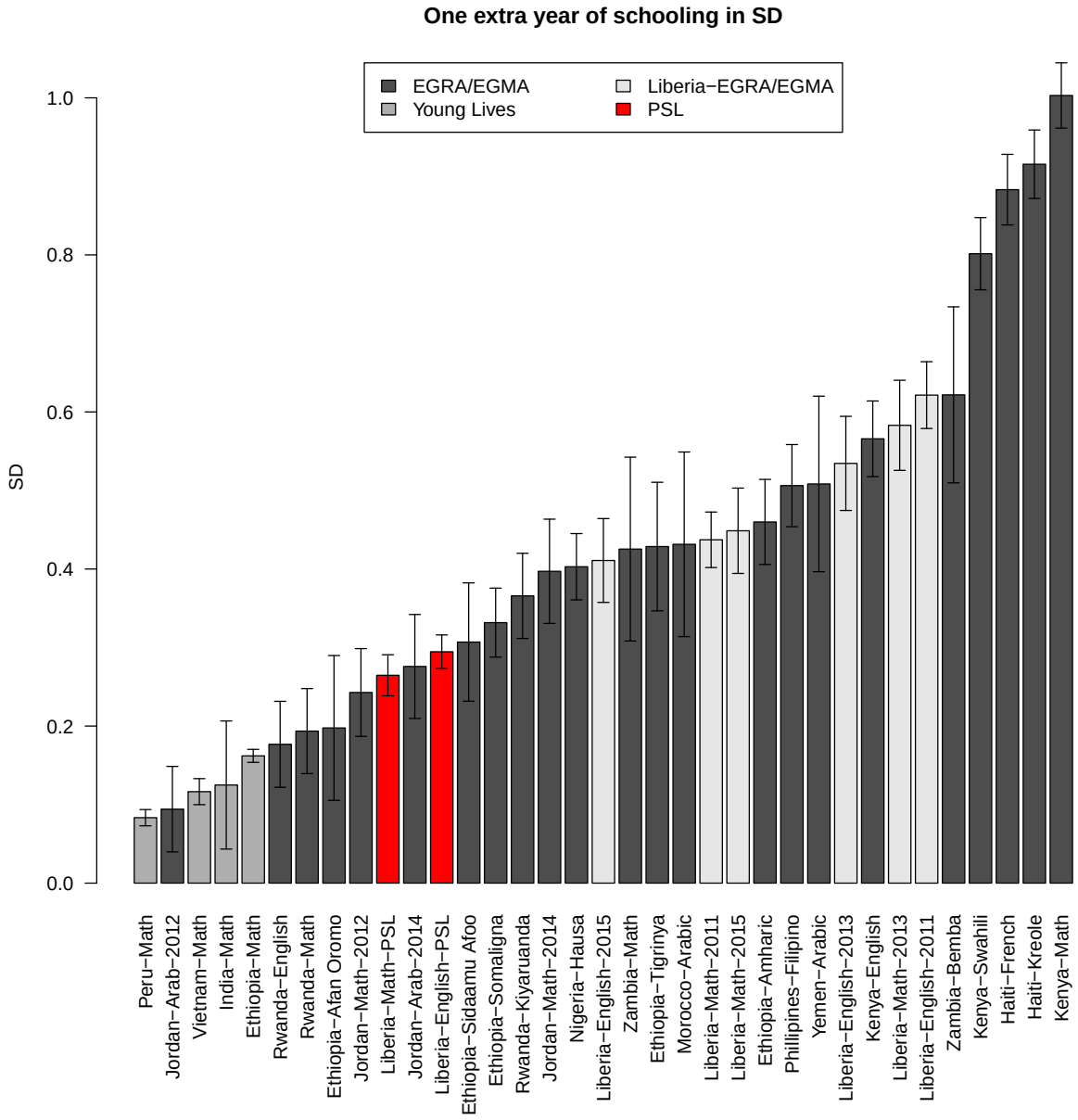
Finally, we added a Raven's style module to measure the students' abstract thinking abilities, and three executive function assessments to

G Standard deviation and equivalent years of schooling

Figure G.1 shows how many standard deviations are equivalent to an additional year of schooling in different countries, with different exams and different underlying populations. Specifically, each bar's height is equal to the estimate of $\beta_1 + \beta_2$ from running the following equation $Z_i = \beta_0 + \beta_1 Grade_i + \beta_2 age_i + \beta_3 male_i + \varepsilon_i$ in each data set. This is slighly different from the methdology used by Evans and Popova (2016). The 90% confidence interval of $\beta_1 + \beta_2$ is also shown. For each data set a vertically linked 2LP IRT model was used to estimate comparable scores across grades.[56] This graph conveys an important message: Reporting results in terms of standard deviations can be misleading. What a standard deviation means in practice (in terms of business as usual) varies a lot depending on what exam is used, what population is tested, and in which country.

[56]The EGRA/EGMA data was provided by the Global Reading Network(https://globalreadingnetwork.net). The Young Live data can be downloaded from the UK Data service webpage. Abhijeet Singh kindly provided the complementary files needed to vertically link the questions for Young Lives.

Figure G.1: International benchmark: how much do children learn per year?



H Absolute learning levels

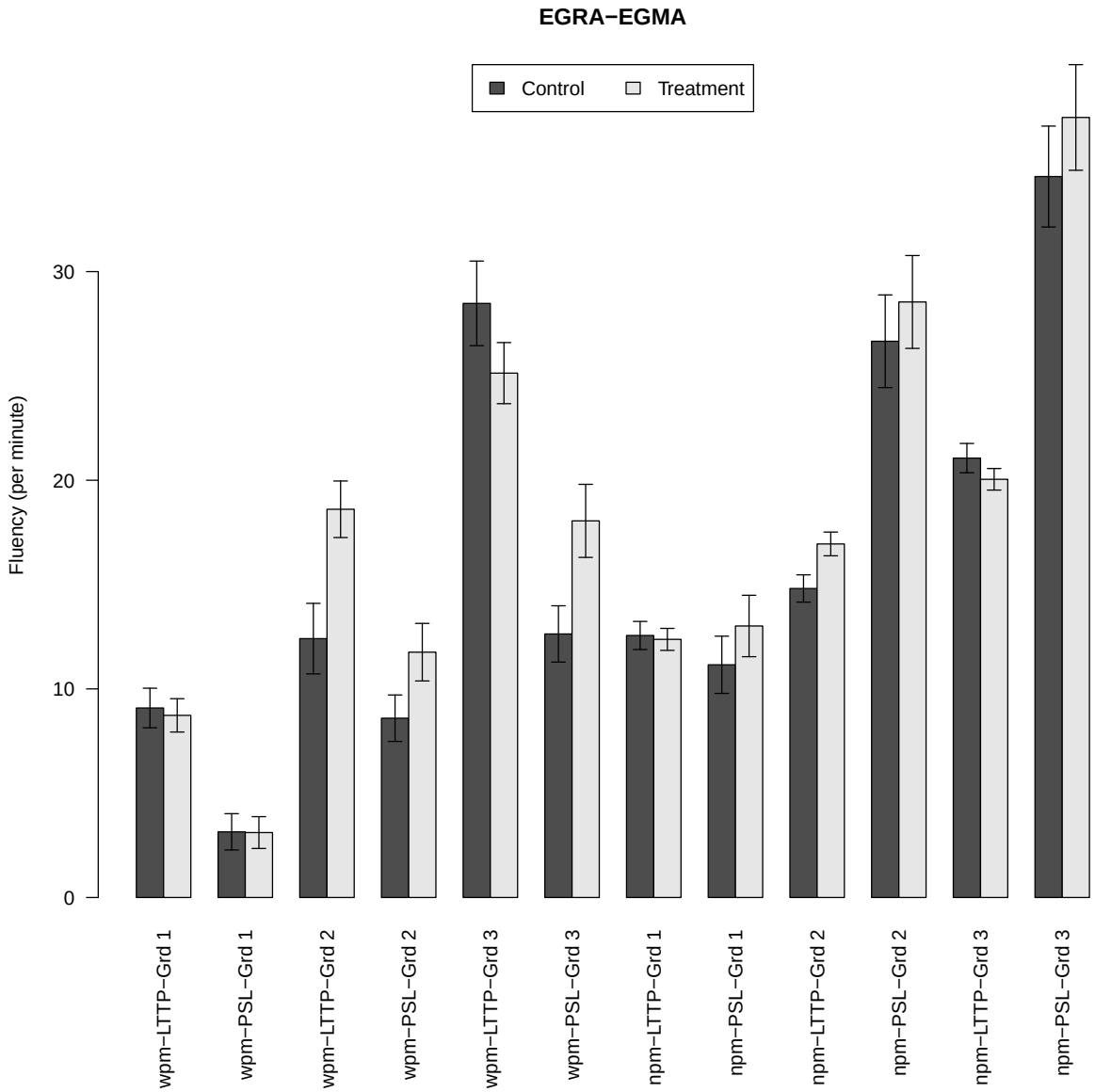
The test has some questions that are identical to those of other assessments, which allows us to compare absolute levels of learning: Two math questions taken from TIMSS released items (M031317 and M031316), two reading comprehension questions taken PrePIRLS released items (L11L01C and L11L02M), and the number and word identification matrices used during the The Liberia Teacher Training Program (LTTP)

program evaluation in Liberia (King et al., 2015).

Figure H.1 shows the average words per minute (wpm) and numbers per minute (npm) students in different grades achieved at the 2013 LTTP program midline, and at our own midline (for both treatment and control schools in both programs). Figures H.2 and H.3 show the results from 4th grade (students enrolled in 3rd grade in 2015/2016) students in treatment and control schools in the TIMSS items, as well as the average for every country in 2011. Finally, Figure H.4 show the results from 4th grade (students enrolled in 3rd grade in 2015/2016) students in treatment and control schools in the PrePIRLS items, as well as the average for every country in 2011.

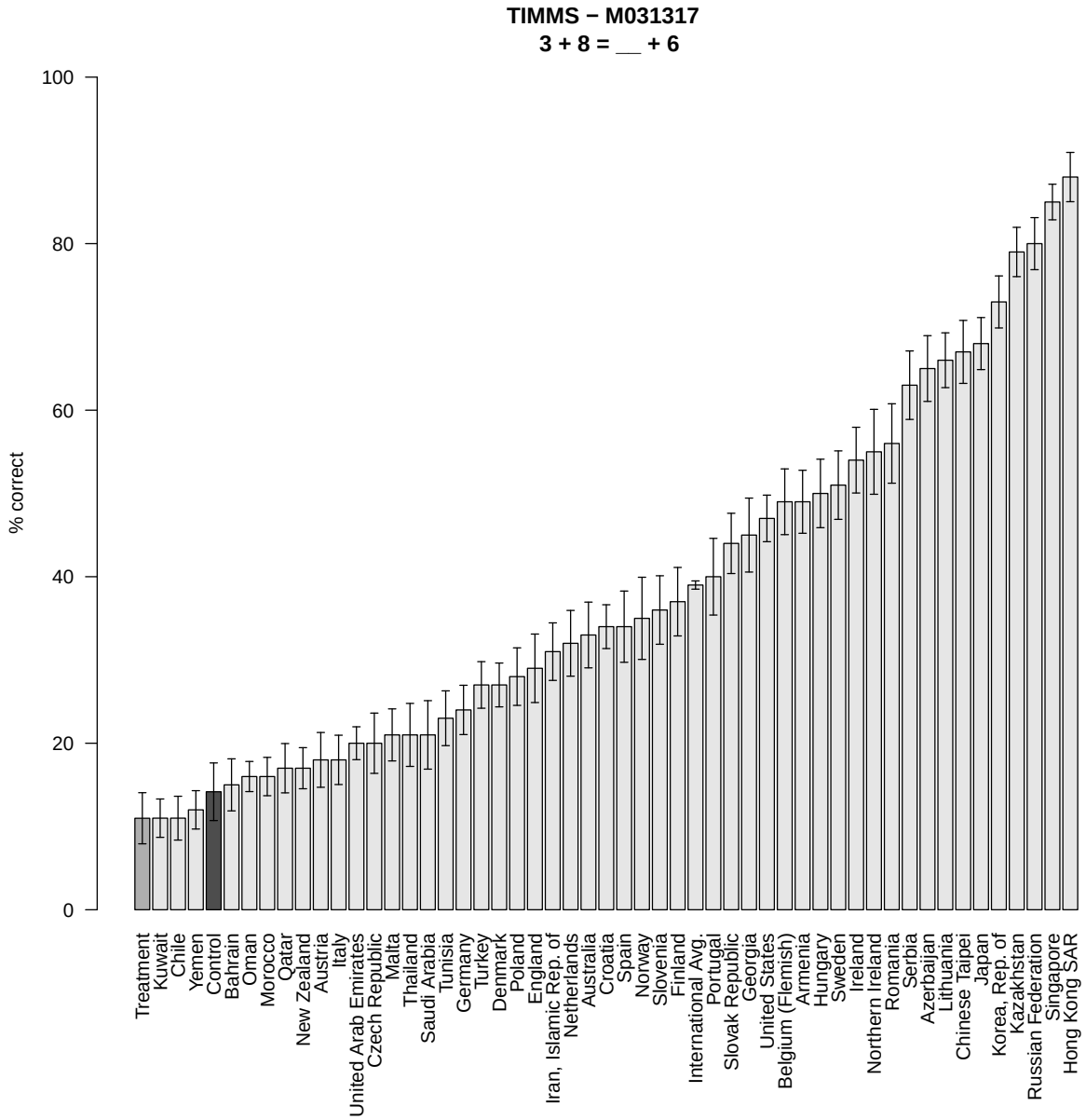
Note that absolute learning levels are low. Despite the positive treatment effect of PSL, students in treatment schools are still far behind their international peers. Either using the TIMSS or the PrePIRLS items, Liberia (both treatment and control schools) ranks at either the very bottom performer or near it. The issue is specially worrisome in English. Liberian students are well below other countries, specially taking into account PrePIRLS is specifically designed for countries where most children in the fourth grade are still developing fundamental reading skills (and thus, in most countries the PIRLS assessment is used).

Figure H.1: Comparison of PSL treatment effects on EGRA and EGMA with earlier USAID program (LTTP)



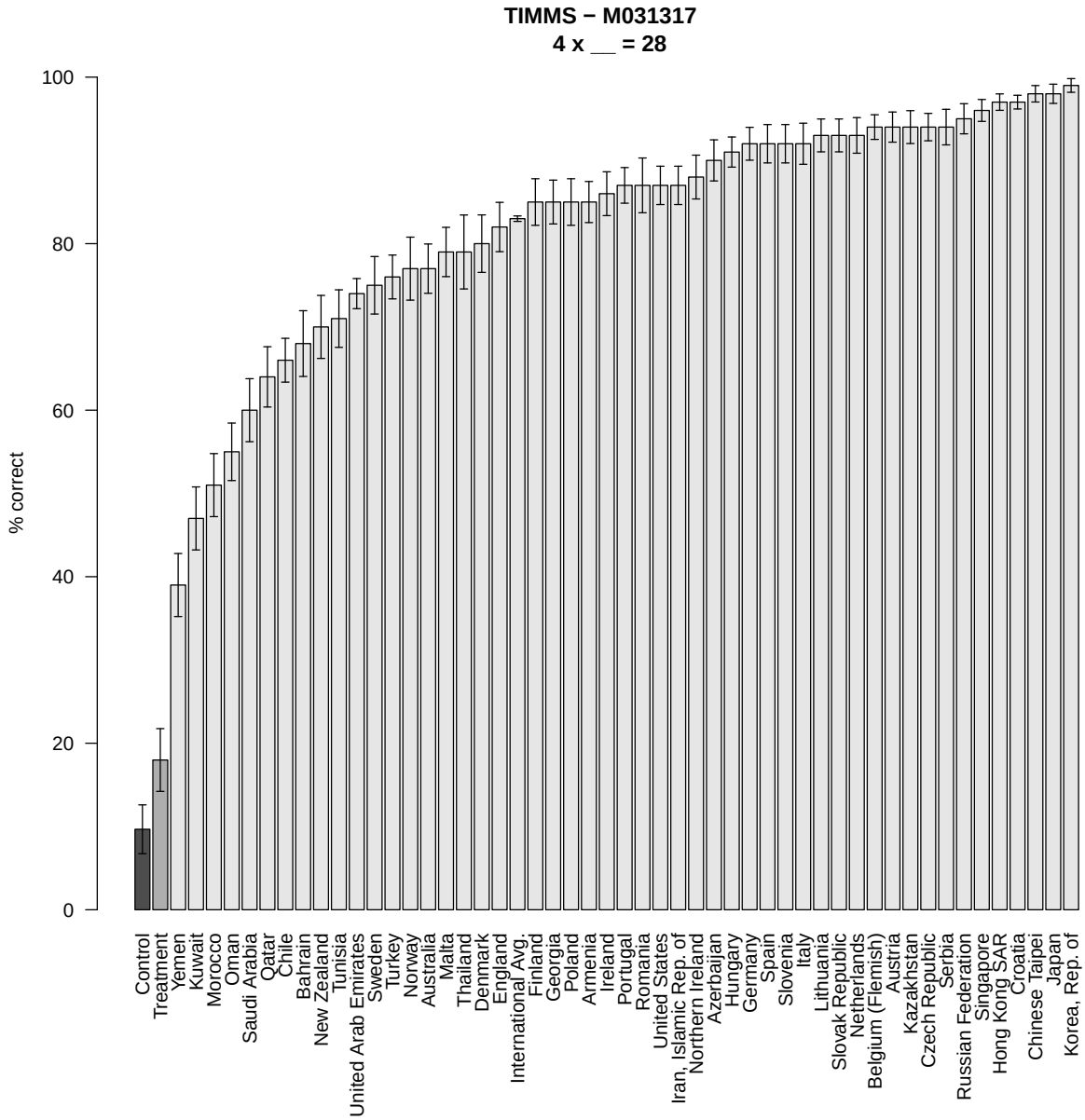
Note: Figures show the average number of words per minute (wpm) and numbers per minute (npm) in the LTTP evaluation and the PSL evaluation for students in Grades 1-3.

Figure H.2: International benchmark for mathematics proficiency (1 of 2)



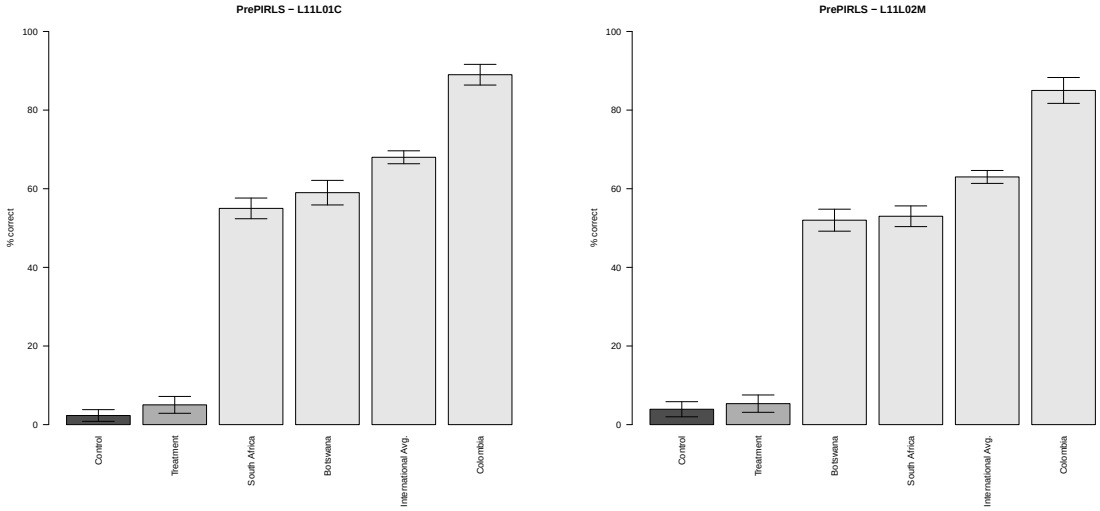
Note: Figures show the proportion of students with correct responses to this question in the PSL evaluation (only students in grade 3 in 2015/2016), and in TIMSS assesments. However, note that this question is multiple-choice in TIMSS and open-ended in our assesments.

Figure H.3: International benchmark for mathematics proficiency (2 of 2)



Note: Figures show the proportion of students with correct responses to this question in the PSL evaluation (only students in grade 3 in 2015/2016), and in TIMSS assesments. Note that this question is open-ended in TIMSS and in our assesments.

Figure H.4: International benchmark for reading proficiency



Note: Figures show the proportion of students with correct responses to this question in the PSL evaluation (only students in grade 3 in 2015/2016), and in PrePirls assesments. Note that question L11L01C is open-ended in TIMSS and in our assesments. Also note that question L11L02M is multiple-choice in TIMSS and open-ended in our assesments.

I Comparisons across contractors

It is important to note that the assignment of contractors to schools was not random. Contractors stated different preferences for locations and some volunteered to manage schools in more remote and marginalized areas. Thus, any heterogeneous effects by contractor or by contractor characteristics are not experimental. Figure I.1 shows the treatment and control schools allocated to each contractor. Table I.1 shows the difference in school (both treatment and control) characteristics across contractors.

Figure I.1: Geographical distribution of contractors across the country

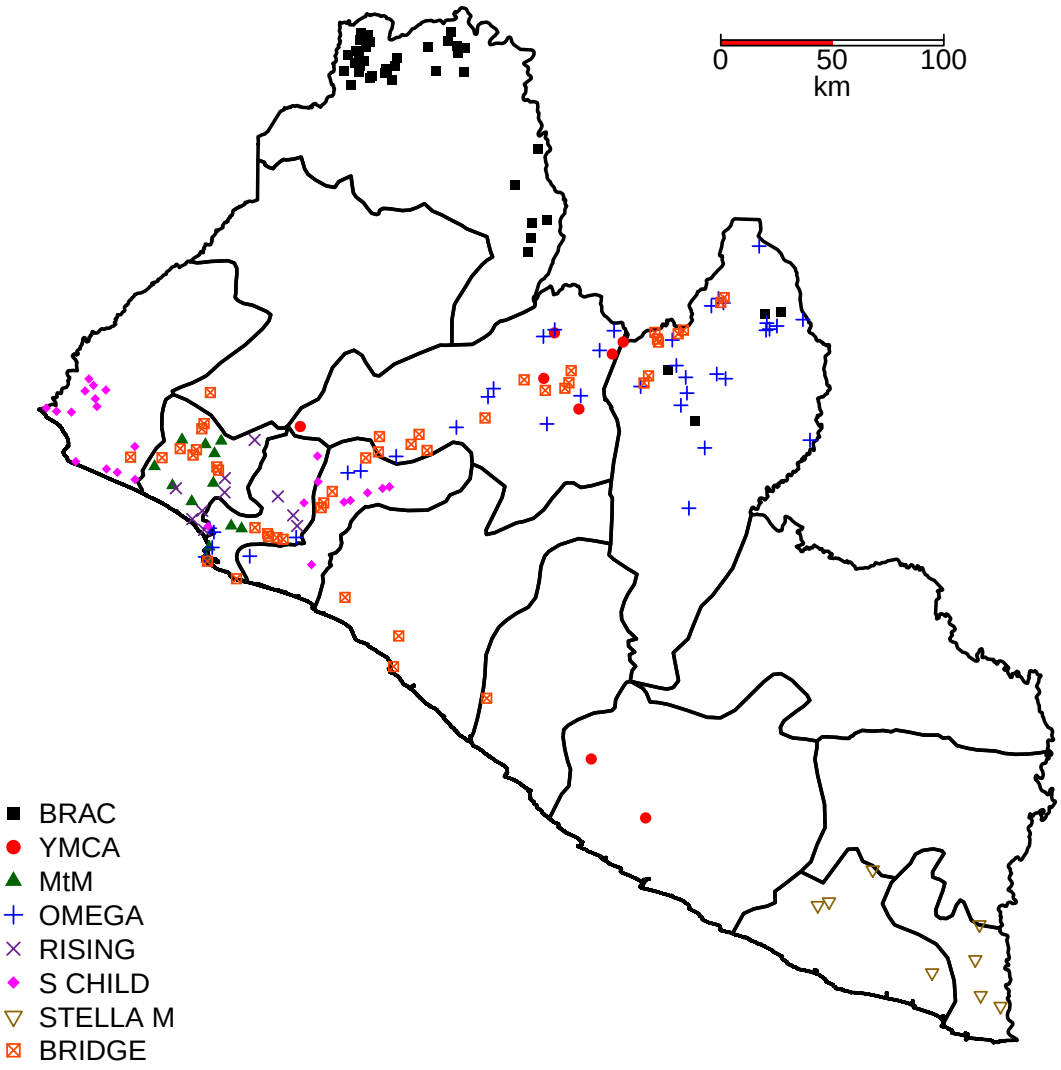


Table I.1: Pre-treatment EMIS characteristics of treatment schools by contractor

	BRAC	BRIDGE	MtM	OMEGA	RISING	SCHILD	STELLAM	YMCA	Total
Students: ECE	126.14	178.50	106.78	158.37	123.67	154.86	115.17	115.43	146.94
	(12.18)	(18.27)	(11.04)	(9.55)	(18.21)	(11.62)	(13.80)	(21.66)	(6.04)
Students: Primary	152.20	225.08	140.33	115.14	120.00	109.36	99.00	110.43	148.28
	(11.72)	(35.58)	(43.47)	(7.96)	(14.47)	(7.57)	(16.13)	(20.35)	(9.68)
Students	278.34	403.58	247.11	273.51	243.67	264.23	214.17	225.86	295.22
	(19.59)	(39.60)	(46.23)	(13.21)	(26.78)	(14.53)	(29.01)	(32.47)	(11.97)
Classrooms per 100 students	0.97	1.28	2.16	0.56	1.90	1.11	0.00	1.45	1.07
	(0.26)	(0.20)	(0.95)	(0.20)	(0.66)	(0.33)	(0.00)	(0.66)	(0.12)
Teachers per 100 students	2.97	2.49	3.95	3.17	3.55	2.76	3.21	3.17	2.98
	(0.19)	(0.17)	(1.11)	(0.18)	(0.62)	(0.26)	(0.29)	(0.45)	(0.11)
Textbooks per 100 students	139.13	75.74	58.67	96.39	120.84	83.64	68.20	75.67	96.63
	(16.65)	(11.50)	(23.96)	(22.27)	(42.49)	(19.15)	(15.53)	(24.30)	(7.90)
Chairs per 100 students	6.19	25.42	38.68	15.56	34.82	23.20	15.49	41.69	20.33
	(2.23)	(3.30)	(11.89)	(2.94)	(9.86)	(7.27)	(11.59)	(16.75)	(2.04)
Food from Gov or NGO	0.03	0.39	0.67	0.31	0.78	0.64	0.67	0.00	0.36
	(0.03)	(0.08)	(0.17)	(0.08)	(0.15)	(0.10)	(0.21)	(0.00)	(0.04)
Solid building	0.26	0.61	0.33	0.14	0.67	0.41	0.00	0.71	0.37
	(0.07)	(0.08)	(0.17)	(0.06)	(0.17)	(0.11)	(0.00)	(0.18)	(0.04)
Water pump	0.31	0.64	0.56	0.71	0.89	0.73	0.83	0.71	0.62
	(0.08)	(0.08)	(0.18)	(0.08)	(0.11)	(0.10)	(0.17)	(0.18)	(0.04)
Latrine/toilet	0.78	0.87	0.81	0.88	0.89	0.91	0.93	0.86	0.86
	(0.07)	(0.06)	(0.13)	(0.05)	(0.08)	(0.06)	(0.07)	(0.14)	(0.03)
Distance to MoE (in KM)	239.70	111.15	35.07	180.22	35.00	75.80	379.11	180.20	154.29
	(2.75)	(13.11)	(6.86)	(15.88)	(4.51)	(4.44)	(11.26)	(19.03)	(7.99)
Observations	40	45	8	12	38	10	24	8	185

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

J How is this report different from contractors' internal monitor and evaluation reports?

Four contractors have produced internal monitor and evaluation reports. Here we address why our results are different from those in these reports. There are four major differences: the counterfactulal, student sorting/selection, treatment of statistical inference, and how student learning is measured.

First, three out four reports lack a counterfactual. Student learning is compared between two points in time; however, there is no measure of what learning would have taken place in the absence of treatment (i.e., had contractor not taken over management in these school). Bridge International Academies' report uses non-experimental methods to create a counterfactual. The observable characteristics of these "control" schools are different from those of control schools (e.g., Bridge schools are on average 4 minutes walking from main road on average vs 26 minutes for "control schools). This is unsurprising, since the schools in the "control" list were filtered out from the original 299 list in Section 2.2.1 by Brdige's restrictions on what schools they were willing to work in.

Second, none of the reports take into account student selection and sorting. Is possible contractors avoided (although we find no evidence of this) certain students (e.g., poorer), or that some parents (e.g., richer) were attracted by PSL schools. Similarly, is possible that the population of students enrolled changes during the school year (e.g., students lagging behind may drop-out in the middle of the school year). If the student population changes across schools and time, changes in test scores could reflect these changes in the population of students.

Third, three out of four reports lack statistical inference. They report means without standard errors,

and thus is impossible to determine whether differences in student test scores are due to "chance" and/or measurement error, or whether they are likely to reflect true changes in learning outcomes. While Bridge's report does present standard errors, these are likely under estimated. Since they do not have enough schools to cluster standard errors at the school level (the level of treatment), they cluster at the school-grade level. This level of clustering yields "enough" clusters, but are likely underestimated of true standard errors. An alternative would be to see how they standard errors behave if they collapse their data at the school level.

Finally, all contractors used EGRA/EGMA to measure student learning. As show in section F our test is more comprehensive, including conceptual questions, modules for abstract reasoning and executive function, and dictation exercises.

Table J.1: Summary of contractors' internal monitor and evaluation reports

	Bridge International Academies[a]	Rising Academies[b]	More Than Me[c]	Street Child[d]
Research design	Difference-in-difference: 6 Bridge PSL and 6 matched-controls schools	Comparison across time	Comparison across time	Comparison across time
Key claims	Reading: average +0.77 SD with respect to "control"; Math: average +0.18 SD with respect to "control"	Over school year, students have increased + 0.75 SD on literacy and + 0.95 SD on numeracy	Comparing baseline to midline. Words/min: G3 +41%. G2 +53%. G1 +300%. Numeracy: some growth for G1/G2/decline for G3	Average 11% increase in scores comparing baseline to midline, driven by maths and G1 students
Sample size	12 schools (6 Bridge/6 "control"). 658 students in both baseline and midline (out of 848 at baseline)	All 5 Rising schools. All G1-G6 students tested, 350 students (varies by collection time)	All 6 MtM schools. All 613 students at baseline, random sample of 193 G1-G3 at midline	All 12 Street Child schools 15% of students in G1-G6 (300 students) sampled
Counterfactual	Six public schools chosen based on similarity and proximity, but not part of the RCT control group	None	None	None
Quantitative data	Selected EGRA/EGMA subtasks. Principal survey and classroom observation.	EGRA/EGMA subtasks. Sample of ECE were tested at baseline only using IDELA	4 EGRA/3 EGMA subtasks	Curriculum based questions. EGRA/EGMA subtasks.
Timeline	Baseline Sept/Oct 2016, midline January.[1]	Four rounds of assessment so far. September 2016 to May 2017[2]	Baseline 24-28 Oct 2016. Midline 9-16 March 2017	Baseline and midline 5 months apart. No dates given
Qualitative data	Not mentioned	Staff/student perception surveys. Qualitative surveys with parents and communities	Not mentioned	Not mentioned
Data collection by	Local enumerators trained by Bridge and Pencils of Promise	Rising central team staff and trained enumerators	Unclear	Unclear
Limitations	Visits to comparison schools in morning/Bridge in afternoon. Bridge students younger, had fewer years of ECE and less likely to have eaten dinner. Bridge schools 4 minutes walk from main road on average, vs 26 minutes for comparison. 22% attrition average. Lower attrition of higher-literacy students at Bridge. Clustered standard errors at grade-school level, but treatment at school level. High number of zero scores (bottom-coding) reduce sample variance and inflate SD changes	Comparison of Rising students at midline with students from grade above at baseline may be misleading because of "Summer Learning Loss". Sample composition changed over the year. Time limits were not imposed for EGRA/EGMA. No confidence intervals (or standard errors) reported. High % of zero scores (bottom coding, which inflates treatment effects)	Underpowered midline sample Sample composition changed over the year. No confidence intervals (or standard errors) reported. High % of zero scores (bottom coding, which inflates treatment effects)	Very little statistical detail provided. Sample composition changed over the year. No confidence intervals (or standard errors) reported

This table was prepared in collaboration with Avi Ahuja and Benjamin Tan.

[a] The report can be found at http://moe.gov.lr/wp-content/uploads/2017/07/BIA-Learning-in-Liberia_Mid-Year-Results_Full-Report_2017.06.20_FINAL.pdf

[b] The report can be found at http://moe.gov.lr/wp-content/uploads/2017/07/RAN-Liberia-Student-Assessment-Interim-Progress-Report-June-2017-FINAL.pdf

[c] The report can be found at http://moe.gov.lr/wp-content/uploads/2017/07/More-Than-Me-Midline-Assessment-Narrative-compressed.pdf

[d] The report can be found at http://moe.gov.lr/wp-content/uploads/2017/07/Street-Child-Progress-Report-compressed.pdf

[1] Planned endline in June/July 2017

[2] Planned endline in June 2017

K Key performance indicators

The contracts stipulated key performance indicators (KPIs) for all contractors (see Table K.1). While these KPIs are used to measure contractor's performance, the evaluation was never intended to serve as the main form of measurement of these KPIs. Instead, the government captures data only on treatment schools to keep contractors accountable for results. For the sake of completeness, the tables below show the key performance indicators (KPIs) for each contractor. Unlike most tables in this document, these tables only include compliant schools (and their control group counterparts). Note that Bridge's MOU specified slightly different indicators,[57] but we present tables using the KPIs in the contracts for the other seven operators to make them comparable.

Note that the KPIs measure retention of students during the academic year, and not across years (see Table K.1). We added retention measurements across academic years to the tables. Note that literacy and numeracy are measured in standard deviations. We do not include a measure of overall school quality. We measure community engagement by whether parents attend schools meetings or not. We measure adherence to teacher code (imperfectly) as the proportion of students that claim teachers never hit them, and as the proportionf of class time used for instruction. Finally, we do not include a measure of adherence to the national curriculum.

Table K.1: Key performance indicators

KPI	Note
Enrolment as % of school capacity	Capacity will be determined as 65 children per grade[1]
Gender parity	% of female students relative to male
% of students retained during the academic year	Measured by drop-out of students from 05/09/16
% of teachers retained during the academic year	Measured by teacher transfers from 05/09/16
Pupil attendance rate	Average attendance of all pupils
Teacher attendance rate	Average attendance of all teaching staff
Literacy rates as measured by external evaluation	
Numeracy rates as measured by external evaluation	
Overall school quality rating	Measured during school inspection
Community engagement rating	Measured during school inspection
Adherence to teacher code of conduct	Based on review of teaching policies and management
Lesson plans adhere to national curriculum	Based on review of lesson planning

[1] For Bridge the capacity is set at 45.

[57] Specifically, the measurements for Bridge are: Teacher attendance, pupil attendance, gender parity, number of books per classroom, and effective number of hours of actual teaching time implemented in a day (Ministry of Education - Republic of Liberia, 2016b)

Table K.2: key performance indicators for BRAC

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	44.04	50.68	6.64	6.64
	(4.27)	(4.19)	(5.98)	(5.18)
Gender parity (Female/Male)	0.87	0.91	0.04	0.04
	(0.03)	(0.04)	(0.05)	(0.05)
Panel B: Retention				
% of students retained (same academic year)	86.36	85.64	-0.72	0.87
	(4.08)	(2.60)	(4.76)	(4.66)
% of students retained (across academic years)	91.37	93.48	2.11	2.62
	(3.41)	(1.48)	(3.31)	(3.56)
% of teachers retained (across academic years)	71.88	77.46	5.57	5.57
	(4.11)	(3.42)	(5.35)	(5.05)
% of teachers retained (same academic year)	87.91	93.38	5.47	5.47
	(3.53)	(1.82)	(3.97)	(4.29)
Panel C: Attendance				
Pupil attendance %	29.48	49.79	20.32***	20.32***
	(5.80)	(4.74)	(7.49)	(5.50)
Teacher attendance %	44.58	52.52	7.94	7.94
	(5.72)	(6.18)	(8.42)	(7.45)
Panel D: Student attainment				
Math	-0.03	0.05	0.08	0.09
	(0.09)	(0.10)	(0.14)	(0.11)
English	-0.21	-0.03	0.18	0.19**
	(0.07)	(0.09)	(0.11)	(0.09)
Panel E: Community engagement				
(mean) e1a_meeting	0.72	0.80	0.08	0.08
	(0.05)	(0.05)	(0.08)	(0.08)
Panel F: Teacher standards				
Teacher never hits students (%)	48.70	47.63	-1.07	-1.07
	(3.52)	(3.30)	(4.82)	(4.10)
Instruction (% of class time)	30.00	62.00	32.00***	32.00**
	(8.34)	(6.63)	(10.65)	(12.00)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.3: key performance indicators for Bridge International Academies

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	77.40	94.01	16.61	16.36
	(9.26)	(6.77)	(11.47)	(10.15)
Gender parity (Female/Male)	0.84	0.90	0.06	0.05
	(0.05)	(0.05)	(0.07)	(0.07)
Panel B: Retention				
% of students retained (same academic year)	82.57	71.60	-10.97	-3.99
	(3.17)	(5.84)	(9.28)	(9.20)
% of students retained (across academic years)	88.55	85.14	-3.42	0.05
	(2.68)	(3.47)	(5.86)	(6.15)
% of teachers retained (across academic years)	79.46	28.67	-50.79***	-50.62***
	(5.22)	(4.48)	(6.88)	(7.52)
% of teachers retained (same academic year)	82.25	81.12	-1.12	-0.90
	(4.21)	(2.95)	(5.14)	(5.64)
Panel C: Attendance				
Pupil attendance %	41.66	49.31	7.65	6.76
	(5.67)	(4.07)	(6.97)	(5.94)
Teacher attendance %	45.81	63.05	17.25***	16.79***
	(4.79)	(4.02)	(6.25)	(5.92)
Panel D: Student attainment				
Math	0.12	0.44	0.32***	0.37***
	(0.06)	(0.07)	(0.10)	(0.11)
English	0.19	0.44	0.25**	0.26*
	(0.09)	(0.07)	(0.12)	(0.15)
Panel E: Community engagement				
(mean) e1a_meeting	0.82	0.86	0.04	0.04
	(0.05)	(0.04)	(0.07)	(0.05)
Panel F: Teacher standards				
Teacher never hits students (%)	58.22	61.37	3.15	3.76
	(3.24)	(4.13)	(5.25)	(4.86)
Instruction (% of class time)	48.64	49.57	0.93	0.45
	(8.20)	(6.49)	(10.45)	(9.66)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.4: key performance indicators for the Youth Movement for Collective Action

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	31.07	45.64	14.57	14.57
	(6.92)	(8.71)	(11.12)	(7.97)
Gender parity (Female/Male)	0.72	0.98	0.26**	0.26**
	(0.05)	(0.09)	(0.10)	(0.08)
Panel B: Retention				
% of students retained (same academic year)	90.07	92.85	2.78	2.89
	(4.12)	(2.67)	(4.60)	(6.30)
% of students retained (across academic years)	92.05	94.71	2.66	2.49
	(3.84)	(2.92)	(4.53)	(5.61)
% of teachers retained (across academic years)	95.83	75.60	-20.24	-20.24
	(4.17)	(10.93)	(11.69)	(14.30)
% of teachers retained (same academic year)	87.50	87.86	0.36	0.36
	(7.98)	(5.96)	(9.96)	(6.45)
Panel C: Attendance				
Pupil attendance %	15.91	29.48	13.57	13.57*
	(14.72)	(19.22)	(24.21)	(7.11)
Teacher attendance %	32.14	61.39	29.25	29.25
	(4.86)	(16.61)	(17.31)	(17.81)
Panel D: Student attainment				
Math	-0.03	0.16	0.19	0.18
	(0.04)	(0.11)	(0.11)	(0.11)
English	-0.23	0.28	0.51**	0.49**
	(0.14)	(0.13)	(0.18)	(0.15)
Panel E: Community engagement				
(mean) e1a_meeting	0.78	0.74	-0.04	-0.04
	(0.11)	(0.06)	(0.12)	(0.08)
Panel F: Teacher standards				
Teacher never hits students (%)	58.93	45.18	-13.75	-13.75
	(9.79)	(4.64)	(10.84)	(7.32)
Instruction (% of class time)	5.00	25.00	20.00	20.00
	(5.00)	(15.00)	(15.81)	(11.55)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.5: key performance indicators for More than Me

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	28.09	35.73	7.64	7.64
	(4.71)	(4.49)	(6.51)	(8.72)
Gender parity (Female/Male)	0.90	0.77	-0.13	-0.13
	(0.10)	(0.08)	(0.13)	(0.08)
Panel B: Retention				
% of students retained (same academic year)	94.62	95.01	0.39	2.06
	(5.53)	(1.26)	(4.17)	(6.52)
% of students retained (across academic years)	94.79	95.56	0.77	2.61
	(5.27)	(1.21)	(3.98)	(6.05)
% of teachers retained (across academic years)	63.61	55.85	-7.76	-7.76
	(15.70)	(11.95)	(19.73)	(23.76)
% of teachers retained (same academic year)	82.34	76.30	-6.04	-6.04
	(7.61)	(10.32)	(12.82)	(13.66)
Panel C: Attendance				
Pupil attendance %	19.50	63.26	43.76**	43.76***
	(9.18)	(14.17)	(16.89)	(10.53)
Teacher attendance %	26.11	67.58	41.46**	41.46***
	(10.56)	(8.93)	(13.82)	(12.35)
Panel D: Student attainment				
Math	0.06	0.36	0.30	0.37**
	(0.13)	(0.11)	(0.18)	(0.13)
English	0.17	0.56	0.39	0.47*
	(0.18)	(0.14)	(0.23)	(0.22)
Panel E: Community engagement				
(mean) e1a_meeting	0.82	0.86	0.04	0.04
	(0.08)	(0.08)	(0.11)	(0.12)
Panel F: Teacher standards				
Teacher never hits students (%)	41.52	81.06	39.54***	39.54***
	(6.35)	(3.99)	(7.50)	(10.04)
Instruction (% of class time)	33.33	50.00	16.67	16.67**
	(16.26)	(16.93)	(23.48)	(5.58)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.6: key performance indicators for Omega Schools

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	47.18	64.64	17.46**	16.52**
	(4.33)	(5.39)	(6.92)	(7.30)
Gender parity (Female/Male)	0.84	0.87	0.03	0.04
	(0.04)	(0.04)	(0.05)	(0.05)
Panel B: Retention				
% of students retained (same academic year)	85.25	88.04	2.79	1.28
	(3.79)	(2.69)	(5.05)	(3.84)
% of students retained (across academic years)	89.54	95.03	5.49	3.58
	(3.25)	(1.30)	(3.63)	(3.46)
% of teachers retained (across academic years)	76.61	82.40	5.79	8.31
	(5.71)	(3.12)	(6.52)	(7.75)
% of teachers retained (same academic year)	87.87	88.56	0.69	-1.27
	(3.32)	(2.56)	(4.19)	(4.69)
Panel C: Attendance				
Pupil attendance %	32.50	39.75	7.25	4.44
	(5.73)	(6.49)	(8.65)	(7.55)
Teacher attendance %	50.67	62.46	11.79	13.28
	(5.92)	(6.06)	(8.47)	(9.17)
Panel D: Student attainment				
Math	0.07	0.04	-0.02	-0.05
	(0.07)	(0.09)	(0.12)	(0.10)
English	-0.01	-0.07	-0.06	-0.10
	(0.09)	(0.13)	(0.17)	(0.11)
Panel E: Community engagement				
(mean) e1a_meeting	0.77	0.76	-0.01	-0.02
	(0.05)	(0.06)	(0.08)	(0.05)
Panel F: Teacher standards				
Teacher never hits students (%)	44.40	55.92	11.52*	11.40*
	(3.45)	(4.71)	(5.83)	(5.64)
Instruction (% of class time)	30.53	48.24	17.71	14.12
	(7.15)	(8.54)	(11.14)	(9.32)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.7: key performance indicators for Rising Academies

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	44.72	42.09	-2.62	0.09
	(7.43)	(4.20)	(8.59)	(5.36)
Gender parity (Female/Male)	0.90	0.92	0.02	0.02
	(0.06)	(0.06)	(0.09)	(0.08)
Panel B: Retention				
% of students retained (same academic year)	90.31	94.57	4.26	9.47
	(6.67)	(2.66)	(6.25)	(9.69)
% of students retained (across academic years)	94.77	94.94	0.17	2.29
	(2.80)	(2.25)	(3.29)	(3.24)
% of teachers retained (across academic years)	66.11	65.56	-0.56	7.92
	(15.07)	(5.33)	(16.16)	(11.82)
% of teachers retained (same academic year)	90.56	80.81	-9.74	-7.38
	(5.80)	(6.92)	(8.99)	(11.82)
Panel C: Attendance				
Pupil attendance %	28.07	57.89	29.82**	34.03*
	(11.88)	(3.97)	(12.66)	(15.32)
Teacher attendance %	25.16	67.86	42.70**	43.55**
	(7.97)	(11.40)	(13.81)	(17.36)
Panel D: Student attainment				
Math	0.10	0.64	0.54**	0.51**
	(0.15)	(0.07)	(0.18)	(0.19)
English	0.27	0.81	0.55*	0.45*
	(0.22)	(0.10)	(0.24)	(0.20)
Panel E: Community engagement				
(mean) e1a_meeting	0.97	0.79	-0.18***	-0.21***
	(0.03)	(0.04)	(0.05)	(0.04)
Panel F: Teacher standards				
Teacher never hits students (%)	54.76	53.80	-0.96	-5.34
	(4.81)	(5.77)	(7.47)	(7.78)
Instruction (% of class time)	30.00	75.00	45.00*	37.50*
	(18.97)	(8.66)	(21.04)	(18.87)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.8: key performance indicators for Stella Maris

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	29.19	46.75	17.56	17.56
	(4.28)	(14.50)	(15.12)	(10.43)
Gender parity (Female/Male)	0.83	0.88	0.06	0.16
	(0.12)	(0.01)	(0.13)	(0.08)
Panel B: Retention				
% of students retained (same academic year)	90.70	95.78	5.08	4.97
	(2.83)	(1.73)	(3.44)	(4.73)
% of students retained (across academic years)	94.21	96.81	2.60	1.32
	(4.06)	(2.01)	(4.63)	(6.16)
% of teachers retained (across academic years)	64.29	73.20	8.92	8.92
	(12.02)	(5.85)	(13.37)	(9.07)
% of teachers retained (same academic year)	81.55	81.09	-0.46	-0.46
	(13.48)	(7.75)	(15.55)	(9.41)
Panel C: Attendance				
Pupil attendance %	26.48	31.17	4.69	4.69
	(15.73)	(12.80)	(20.28)	(10.08)
Teacher attendance %	22.50	78.72	56.22***	56.22**
	(13.15)	(8.71)	(15.77)	(21.39)
Panel D: Student attainment				
Math	-0.68	-0.81	-0.12	-0.14
	(0.29)	(0.19)	(0.34)	(0.44)
English	-0.66	-0.85	-0.19	-0.20
	(0.31)	(0.17)	(0.35)	(0.39)
Panel E: Community engagement				
(mean) e1a_meeting	0.41	0.36	-0.05	-0.05
	(0.12)	(0.05)	(0.13)	(0.16)
Panel F: Teacher standards				
Teacher never hits students (%)	36.01	45.23	9.22	9.22
	(12.48)	(3.75)	(13.03)	(10.21)
Instruction (% of class time)	0.00	32.50	32.50*	32.50*
	(0.00)	(16.52)	(16.52)	(16.52)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table K.9: key performance indicators for Street Child

	(1) Control	(2) Treatment	(3) Difference	(4) Difference (F.E)
Panel A: Enrollment				
Enrolment as % of school capacity	40.06	52.61	12.56*	12.56**
	(3.33)	(5.54)	(6.47)	(5.96)
Gender parity (Female/Male)	0.88	0.91	0.02	0.02
	(0.05)	(0.04)	(0.07)	(0.06)
Panel B: Retention				
% of students retained (same academic year)	93.26	93.39	0.13	0.23
	(1.90)	(1.95)	(2.66)	(3.24)
% of students retained (across academic years)	94.91	94.78	-0.13	0.60
	(1.29)	(1.81)	(2.37)	(2.84)
% of teachers retained (across academic years)	69.52	70.97	1.45	1.45
	(5.88)	(6.78)	(8.97)	(9.07)
% of teachers retained (same academic year)	87.51	87.84	0.34	0.34
	(4.25)	(3.85)	(5.73)	(5.26)
Panel C: Attendance				
Pupil attendance %	30.34	50.22	19.88	19.88*
	(9.26)	(8.19)	(12.36)	(10.38)
Teacher attendance %	32.08	58.79	26.71**	26.71***
	(8.14)	(5.00)	(9.55)	(8.43)
Panel D: Student attainment				
Math	-0.15	0.16	0.31*	0.30
	(0.14)	(0.09)	(0.17)	(0.18)
English	0.05	0.29	0.24	0.25*
	(0.13)	(0.07)	(0.16)	(0.13)
Panel E: Community engagement				
(mean) e1a_meeting	0.72	0.82	0.10	0.10
	(0.10)	(0.05)	(0.11)	(0.09)
Panel F: Teacher standards				
Teacher never hits students (%)	38.01	58.33	20.33***	20.33***
	(4.76)	(4.75)	(6.72)	(6.52)
Instruction (% of class time)	30.00	42.50	12.50	12.50
	(10.59)	(8.27)	(13.43)	(15.48)

This table presents the mean and standard error of the mean (in parenthesis) for the control (Column 1) and treatment (Column 2) groups, 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). Data is collapse at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

L Full list of schools

Table L.1 summarizes the difference between schools in our main sample, and the set of schools actually managed by PSL contractors.

Table L.1: Number of schools by contractor

	(1)	(2)	(3)	(4)	(5) (1)-(2)+(3)+(4)	(6) [(1)-(2)]/(1)
	Randomly assigned	Non compliant	Replacement	Outside sample	Managed	% compliant in sample
BRAC	20	0	0	0	20	100%
Bridge	23	0	0	2	25	100%
YMCA	4	0	0	0	4	100%
MtM	6	2	2	0	6	67%
Omega	19	2	0	0	17	89%
Rising	5	1	0	1	5	80%
Stella	4	4	0	0	0	0%
St. Child	12	2	2	0	12	83%

Note: Table shows the number of schools randomly assigned to treatment originally (Column 1), the schools that either did not meet criteria determined by the Ministry of Education or were refused by contractors (Column 2). For schools that did not meet the criteria determined by the Ministry of Education replacement schools were provided, presenting each contractor with a new list of counterparts and informing them, as before, that they would operate one of each pair of schools (but not which one). Replacement schools are show in Column 3. Column 4 has non-randomly assigned schools given to some contractors. Column 5 shows the final number of schools managed by a contractor. Finally, the last column shows the percent of schools actually managed by the contractor that are in our main sample.

The list below shows all schools involved in the PSL evaluation program. School ID is the EMIS code for the school, contractor indicates the contractor that each "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" is equal to one for schools in the original RCT list, and "Final" 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 L.2: School list

School ID	contractor	Treatment	GroupID	Original	Final	PSL
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

Continued on next page

School ID	contractor	Treatment	GroupID	Original	Final	PSL
20027	BRIDGE	0	7	1	1	0
20057	BRIDGE	1	7	1	1	1
20167	YMCA	1	8	1	1	1
20182	YMCA	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
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

Continued on next page

School ID	contractor	Treatment	GroupID	Original	Final	PSL
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
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

Continued on next page

School ID	contractor	Treatment	GroupID	Original	Final	PSL
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
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	YMCA	0	62	1	1	0
20200	YMCA	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

Continued on next page

School ID	contractor	Treatment	GroupID	Original	Final	PSL
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	YMCA	0	79	1	1	0
150082	YMCA	1	79	1	1	1
100111	STELLAM	0	80	1	1	0
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

Continued on next page

School ID	contractor	Treatment	GroupID	Original	Final	PSL
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	YMCA	1	95	1	1	1
20239	YMCA	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