

Online Appendix to Making the Grade: The Trade-off between Efficiency and Effectiveness in Improving Student Learning

Jason T. Kerwin and Rebecca L. Thornton

January 30, 2018

[Click here for the latest version of this appendix](#)

List of Papers from McEwan (2015) used in Figures 3, 4, and 5

- Abeberese, A. B., Kumler, T. J., & Linden, L. (2012). *Improving reading skills by encouraging children to read: a randomized evaluation of the Sa Aklat Siskat reading program in the Philippines*. Unpublished manuscript, Columbia University, New York, NY.
- Banerjee, A., Banerji, R., Duflo, E., & Walton, M. (2012). *Effective pedagogies and a resistant education system: Experimental evidence on interventions to improve basic skills in rural India*. Unpublished manuscript, MIT, Cambridge, MA.
- Blimpo, M. P., & Evans, D. K. (2011). *School-based management and educational outcomes: Lessons from a randomized field experiment*. Unpublished manuscript, Stanford University, Stanford, CA.
- Cabezas, V., Cuesta, J. I., & Gallego, F. A. (2011). *Effects of short-term tutoring on cognitive and non-cognitive skills: Evidence from a randomized evaluation in Chile*. Unpublished manuscript, Pontificia Universidad Católica de Chile, Santiago.
- Cristia, J. P., Ibarrarán, P., Cueto, S., Santiago, A., & Severín, E. (2012). *Technology and child development: Evidence from the One Laptop per Child program* (Working Paper IDB-WP-304). Washington, DC: Inter-American Development Bank.
- 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, 1739-1774. doi:10.1257/aer.101.5.1739

- Duflo, E., Dupas, P., & Kremer, M. (2012a). *School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools* (Working Paper 17939). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w17939
- Duflo, E., Hanna, R., & Ryan, S. P. (2012b). Incentives work: Getting teachers to come to school. *American Economic Review*, 102,1241-1278. doi:10.1257/aer.102.4.1241
- Friedman, W., Gerard, F., & Ralaingita, W. (2010). *International independent evaluation of the effectiveness of Institut pour l'Education Populaire's "Read-Learn-Lead" (RLL) program in Mali: Mid-term report*. Research Triangle Park, NC: RTI International.
- Gardner, J. M., Grantham-McGregor, S., & Baddeley, A. (1996). *Trichuris trichiura* infection and cognitive function in Jamaican school children. *Annals of Tropical Medicine and Parasitology*, 90, 55-63.
- Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, 2, 205-227. doi:10.1257/app.2.3.205
- Grigorenko, E. L., Sternberg, R. J., Jukes, M., Alcock, K., Lambo, J., Ngorosho, D., Bundy, D. A. (2006). Effects of antiparasitic treatment on dynamically and statically tested cognitive skills over time. *Journal of Applied Developmental Psychology*, 27, 499-526. doi:10.1016/j.appdev.2006.08.005
- Lai, F., Zhang, L., Qu, Q., Hu, X., Shi, Y., Boswell, M., & Rozelle, S. (2012). *Does computer-assisted learning improve learning outcomes? Evidence from a randomized experiment in public schools in rural minority areas in Qinghai* (Working Paper No. 237). Stanford, CA: Rural Education Action Project.
- Lucas, A. M., McEwan, P. J., Ngware, M., & Oketch, M. (2012a). Improving early-grade literacy in East Africa: Experimental evidence from Kenya and Uganda. *Journal of Policy Analysis and Management*.
- Miguel, E., & Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72, 159-217. doi:10.1111/j.1468-0262.2004.00481.x
- Nokes, C., Grantham-McGregor, S. M., Sawyer, A. W., Cooper, E. S., Robinson, B. A., & Bundy, D. A. P. (1992). Moderate to heavy infections of *trichuris trichiura* affect cognitive function in Jamaican school children. *Parasitology*, 104, 539-547. doi:10.1017/S0031182000063800
- Oster, E., & Thornton, R. (2009). *Menstruation and education in Nepal* (Working Paper 14853). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w14853

- Powell, C. A., Walker, S. P., Chang, S. M., Grantham-McGregor, S. M. (1998). Nutrition and education: A randomized trial of the effects of breakfast in rural primary school children. *American Journal of Clinical Nutrition*, 68, 873-879.
- Pradhan, M., Suryadarma, D., Beatty, A., Wong, M., Alishjabana, A., Gaduh, A., & Artha, R. P. (2011). *Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia* (Policy Research Working Paper 5795). Washington, DC: World Bank.
- Rico, J. A., Kordas, K., López, P., Rosado, J. L., García Vargas, G., Ronquillo, D., & Stolfus, R. J. (2006) Efficacy of iron and/or zinc supplementation on cognitive performance of lead-exposed Mexican schoolchildren: A randomized, placebo-controlled trial. *Pediatrics*, 117, e518-e527. doi:10.1542/peds.2005-1172
- Simeon, D. T., Grantham-McGregor, S. M., Callender, J. E., & Wong, M. S. (1995a). Treatment of *trichuris trichiura* infections improves growth, spelling scores and school attendance in some children. *Journal of Nutrition*, 125, 1875-1883.
- Simeon, D. T., Grantham-McGregor, S. M., & Wong, M. S. (1995b). *Trichuris trichiura* infection and cognition in children: Results of a randomized clinical trial. *Parasitology*, 110, 457-464. doi:10.1017/S0031182000064799
- Van Stuijvenberg, M. E., Kvalsig, J. D., Faber, M., Kruger, M., Kenoyer, D. G., & Spinnler Benadé, A. J. (1999). Effect of iron-, iodine-, and β -carotene-fortified biscuits on the micronutrient status of primary school children: a randomized controlled trial. *American Journal of Clinical Nutrition*, 69, 497-503.
- Watkins, W. E., Cruz, J. R., & Pollitt, E. (1996). The effects of deworming on indicators of school performance in Guatemala. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 90, 156-161. doi:10.1016/S0035-9203(96)90121-2

Appendix Figure A1
Classroom Observation Instrument

Specific Lesson Actions (Repeated for Second and Third Ten-Minute Window)

Time	Teacher actions	Pupil actions			
FIRST 10 minutes: <hr/> (start time) <hr/> (end time)	<u>Positive actions:</u> <input type="checkbox"/> Refers to TG or lesson plan while teaching <input type="checkbox"/> Moves freely around the classroom <input type="checkbox"/> Calls on individual pupils by name <input type="checkbox"/> Encourages pupil participation and keeps their attention <input type="checkbox"/> Brings pupils back on task when needed <input type="checkbox"/> Observes and records pupils' performance <u>Negative actions:</u> <input type="checkbox"/> Lesson does not appear planned <input type="checkbox"/> Remains at the front of the class <input type="checkbox"/> Does not call on individual pupils by name <input type="checkbox"/> Very little pupil participation and attention <input type="checkbox"/> Ignores or does not address pupils who are off task <input type="checkbox"/> Does not record pupil performance <u>Other:</u> % time speaking English _____% % time speaking LL _____% Minutes out of class _____ min. Minutes in class but not teaching _____ min. Minutes teaching _____ min.	Reading			
		<input type="checkbox"/> Sounds <input type="checkbox"/> Letters <input type="checkbox"/> Words <input type="checkbox"/> Sentences	<input type="checkbox"/> Whole class <input type="checkbox"/> Smaller group <input type="checkbox"/> Individual at seat <input type="checkbox"/> Individual at board	<input type="checkbox"/> On board <input type="checkbox"/> In primer <input type="checkbox"/> In reader <input type="checkbox"/> Other: _____	<input type="checkbox"/> English <input type="checkbox"/> LL
		Minutes on pupil reading tasks _____ min. % of pupils participating in reading task _____%			
		Writing			
		<input type="checkbox"/> Pictures <input type="checkbox"/> Letters <input type="checkbox"/> Words <input type="checkbox"/> Sentences <input type="checkbox"/> Name	<input type="checkbox"/> Air writing <input type="checkbox"/> Handwriting practice <input type="checkbox"/> Copying teacher text from the board <input type="checkbox"/> Writing own text	<input type="checkbox"/> On slate <input type="checkbox"/> On paper <input type="checkbox"/> On board	<input type="checkbox"/> English <input type="checkbox"/> LL
		Minutes on pupil writing tasks _____ min. % of pupils participating in writing task _____%			
		Speaking/Listening			
			<input type="checkbox"/> To a partner <input type="checkbox"/> To a small group <input type="checkbox"/> To the whole class <input type="checkbox"/> To the teacher		<input type="checkbox"/> English <input type="checkbox"/> LL
		Minutes on pupil speaking/listening tasks _____ min. % of pupils participating in speaking/listening task _____%			

Appendix Table A1
NULP Components by Study Arm

	Study Arm		
	Full-cost program	Reduced-cost program	Control
Number of Schools	12	14	12
Pedagogy			
Local Language-First Instruction	Yes	Yes	
NULP Instructional Model	Yes	Yes	
Books			
Leblango Primers	3 per student (1 for each term)	3 per student (1 for each term)	
Leblango Readers	3 per student (1 for each term)	3 per student (1 for each term)	
Leblango Teacher's Guides	1 per classroom	1 per classroom	
English Primers	3 per student (1 for each term)	3 per student (1 for each term)	
English Teacher's Guides	1 per classroom	1 per classroom	
Materials			
Slates	1 per student		
Wall Clocks	1 per classroom		
Training and Support for Teachers			
Leblango Orthography Training (5 Days, before school year)	Before school year, non- residential, taught by MT staff	Before school year, residential, taught by CCTs	
Literacy Methods Training (3-5 days, before each term)	1X/term, residential, taught by MT staff	1X/term, non-residential, taught by CCTs	
Saturday in-service training wkshps (1 Day, during each term)	2X/term, non-residential, taught by MT staff	2X/term, non-residential, taught by CCTs	
Classroom support supervision	3X/term from MT staff, 2X/term from CCTs	2X/term from CCTs	
Other			
Take a Book Home Activity	Early during first term		
Literacy & Local Language Radio Program		1X/month, available to whole community	

Appendix Table A2

Comparison of Arancibia, Popova, and Evans (2016) Indicators for Full-Cost and Reduced-Cost NULP

	Full-Cost	Reduced-Cost
Which organization designed the program?	2	2
Which organization is implementing the program?	2	2
Was program design based on a diagnostic or evaluation of some kind? If so, which one?	1	1
Program objectives	To create a culture of literacy and engage with people responsible for growing this culture. For students to learn the names of the letters of the alphabet.	
Targeting by geography	1	1
Targeting by subject	0	0
Targeting by grade	1	1
Targeting by years of experience	0	0
Targeting by skill gaps	0	0
Targeting by contract teachers	0	0
Do teachers have to pay some cost for the training (including their own transport cost)? If so, how much over one school year?	0	0
Does participation have any of these implications?	0	0
Is there a positive consequence if teachers are well evaluated?	0	0
Is there a negative consequence if teachers are poorly evaluated?	0	0
Did the program provide textbooks?	0	0
Did the program provide storybooks?	1	1
Did the program provide computers?	0	0
Did the program provide teacher manuals?	1	1
Did the program provide lesson plans/videos?	1	1
Did the program provide scripted lessons?	1	1
Did the program provide craft materials?	0	0
Did the program provide other reading materials - flashcards, word banks, reading pamphlets or similar?	1	1
Did the program provide software?	0	0
How many teachers received training under this program each year?	24	28
How many schools is the program being implemented in (at the time of the evaluation)?	12	14
How many years has the program been running (at the time of the evaluation)?	2	2
In the last year what percentage of the teachers who began the training dropped out before the end?	0	0
What is the primary focus of the training program?	2	2
What is the secondary focus of the training program?	1	1
What is the subject focus of the training program (if any)?	1	1
Did the training involve lectures?	1	1
Did the training involve discussion?	1	1
Did the training involve lesson enactment?	1	1
Did the training involve materials development?	0	0
Did the training involve training on how to conduct diagnostics?	1	1
Did the training involve lesson planning?	1	1
Did the training involve the use of scripted lessons?	1	1
Is it a cascade training model (i.e. one where program trainers train trainers who then train teachers)?	0	1
What is the most common profile of the direct trainers?	1	4
Is there a part of the training where teachers meet with trainers for several days in a row?	1	1
During this period, what is the total hours of teacher training they receive?	120	120
During this period, how many hours of lectures do they receive?	60	60
During this period, how many hours do they spend practicing with students?	0	0
During this period, how many hours do they spend practicing with other teachers?	60	60
Over how many weeks?	40	40
Where does this part of the training take place?	2	2
How many teachers are in each training session?	24	26
How many in-school follow-up support visits do teachers receive after the initial training (if any)?	9	6
What is the nature of these follow-up visits?	1	1
How many weeks of distance learning does the program include (if any)?	0	0
Over how many months?	9	9
Tested subject	Average	Average
Africa dummy	1	1
Interviewed	1	1

Appendix Table A3
Baseline Covariate Means by Study Arm

	(1) Baseline Sample			(4) Longitudinal Sample			(7) Lost to Followup		
	Control	Full-Cost	Reduced-Cost	Control	Full-Cost	Reduced-Cost	Control	Full-Cost	Reduced-Cost
Present at Endline	0.795	0.808	0.741	1.000	1.000	1.000	0.000	0.000	0.000
Male	0.486	0.509	0.474	0.488	0.524	0.479	0.475	0.447	0.460
Age	7.018	7.078	7.017	7.013	7.052	7.000	7.041	7.191	7.066
<u>EGRA</u>									
PCA EGRA score index	-0.000	0.006	-0.075	0.000	0.039	-0.085	-0.000	-0.130	-0.045
1(any correct)	0.396	0.386	0.368	0.394	0.406	0.378	0.402	0.301	0.341
Letter name knowledge (letters per minute)	1.150	1.190	1.274	1.180	1.377	1.206	1.033	0.400*	1.469
Initial sound identification (sounds identified)	0.153	0.123	0.070	0.161	0.148	0.046	0.122	0.017	0.138
Familiar word reading (words per minute)	0.169	0.182	0.044	0.168	0.225	0.025	0.171	0.000	0.099
Invented word reading (words per minute)	0.094	0.132	0.029	0.084	0.163	0.008	0.130	0.000	0.088
Oral reading fluency (words per minute)	0.503	0.552	0.126	0.508	0.684	0.037	0.480	0.000**	0.382
Reading comprehension (questions correct)	0.327	0.318	0.266	0.327	0.342	0.272	0.325	0.217	0.249
<u>Writing Test</u>									
PCA writing score index	0.000	-0.011	-0.027	0.010	-0.008	-0.024	-0.039	-0.022	-0.036
1(any correct)	0.212	0.330	0.186	0.237	0.355	0.195	0.114	0.226	0.160
African name (surname) writing	0.180	0.323	0.181	0.201	0.348*	0.193	0.098	0.217	0.149
English name (given name) writing	0.127	0.043	0.054*	0.145	0.043	0.058	0.057	0.043	0.044
Ideas	0.005	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000
Organization	0.002	0.002	0.000	0.002	0.002	0.000	0.000	0.000	0.000
Voice	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Word choice	0.057	0.023	0.016	0.069	0.023	0.019	0.008	0.026	0.006
Sentence fluency	0.005	0.000	0.001	0.006	0.000	0.002	0.000	0.000	0.000
Conventions	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Baseline Sample includes 1,900 students who were tested at baseline. Longitudinal Sample includes 1,481 students who were tested at baseline as well as endline. Lost to Followup includes 419 students who were tested at baseline but not at endline. Stars indicate randomization inference p-values for a test of the null hypothesis of no difference between each NULP variant and the control group, conditioning on stratification cell indicators and the date of the baseline exam: * p<0.05, ** p<0.01, *** p<0.001.

Appendix Table A4
Factor Loadings for Classroom Management Indices

	(1) Keeps Students Focused	(2) Solid Lesson Plan	(3) Active Throughout Classroom
<u>Teacher Actions:</u>			
Refers to Teacher's Guide	0.01	0.34	0.05
Moves Freely Around Classroom	0.00	-0.03	0.32
Calls on Individuals	0.02	0.09	0.13
Brings Students Back on Task	0.48	-0.01	0.13
Observes/ Records Performance	0.02	0.07	0.27
Lesson Not Planned	0.01	-0.31	0.05
Very Little Participation	-0.06	-0.13	-0.01
Ignores Off-Task Students	-0.42	0.06	0.19
Share of Time Speaking Leblango	-0.02	-0.03	-0.06
Share of Variance Explained	0.81	0.31	0.25

Notes: This table presents the rotated factor loadings for the three indices of classroom management techniques used in the paper. We retain all factors that explain at least 10% of the variance of the data, and apply a varimax rotation to the resulting set of selected factors. We then give each factor a descriptive name based on which of the underlying behaviors it loads on.

Appendix Table A5
Factor Loadings for Reading Pedagogy Indices

	(1)	(2)	(3)	(4)	(5)
			Basic	Leblango	
	Sounds and Letters Only	Whole Language On Board	Elements in Breakout Sessions	Sentences in Reader	Paragraphs in Primer
<u>Students are Reading:</u>					
Sounds	0.27	0.01	-0.02	0.07	0.10
Letters	0.41	0.04	0.01	0.09	0.01
Words	0.01	0.05	0.17	-0.10	-0.02
Sentences	-0.29	0.08	-0.02	0.25	0.14
Whole Paragraphs In Smaller Groups	0.00	0.03	0.16	-0.06	0.14
Individually at Seats	-0.05	0.06	0.26	-0.01	-0.02
Individually on Board	0.03	0.02	0.27	0.07	0.03
Whole Group on Board	-0.03	0.08	-0.06	0.07	-0.17
In Primer	0.01	0.52	0.00	-0.02	0.06
In Reader	0.00	-0.20	-0.05	-0.05	0.27
From Other Text	0.03	-0.13	0.14	0.24	-0.13
Percent of Students Participating	-0.04	-0.06	0.17	-0.10	-0.18
Share in Leblango	0.02	-0.03	0.08	-0.02	0.16
	0.03	0.02	0.04	0.29	-0.01
Share of Variance Explained	0.49	0.35	0.27	0.19	0.15

Notes: This table presents the rotated factor loadings for the five indices of reading pedagogy used in the paper. We retain all factors that explain at least 10% of the variance of the data, and apply a varimax rotation to the resulting set of selected factors. We then give each factor a descriptive name based on which of the underlying behaviors it loads on.

Appendix Table A6
Factor Loadings for Writing Pedagogy Indices

	(1) Pictures, Words, and Stories	(2) Copying Teacher's Text	(3) Leblango Practice on Slates	(4) Pictures and Letters on Paper, High-Energy	(5) Leblango Sentences and Handwriting
Students are Writing:					
Pictures	0.15	-0.04	0.11	0.12	-0.14
Letters	-0.50	0.04	0.15	0.11	-0.08
Words	0.10	0.11	0.04	-0.07	-0.04
Sentences	0.04	0.05	-0.02	0.03	0.34
Their Names	0.06	0.00	0.24	0.00	0.07
Air Writing	-0.22	-0.13	0.00	-0.05	0.04
Handwriting Practice	-0.01	0.02	0.15	0.03	0.26
Copying Teacher's Text from Board	0.05	0.44	0.09	0.03	-0.04
Writing Own Text	0.12	-0.34	0.08	0.07	-0.07
On Slate	0.01	0.00	0.31	-0.11	-0.03
On Paper	0.06	0.06	-0.11	0.39	0.04
On Board	0.00	-0.02	-0.11	-0.22	-0.01
Percent of Students Participating	-0.01	0.01	0.08	0.14	-0.12
Share in Leblango	-0.05	-0.06	0.18	0.01	0.11
Share of Variance Explained	0.46	0.31	0.21	0.16	0.12

Notes: This table presents the rotated factor loadings for the five indices of writing pedagogy used in the paper. We retain all factors that explain at least 10% of the variance of the data, and apply a varimax rotation to the resulting set of selected factors. We then give each factor a descriptive name based on which of the underlying behaviors it loads on.

Appendix Table A7

Program Impacts on Early Grade Reading Assessment Scores, Without Controlling for Baseline Scores
(in SDs of the Control Group Endline Score Distribution)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PCA Leblango EGRA Score Index [†]	Letter Name Knowledge	Initial Sound Recognition	Familiar Word Recognition	Invented Word Recognition	Oral Reading Fluency	Reading Comprehension
Full-cost program	0.654***	1.043***	0.649***	0.382***	0.233	0.484**	0.449**
S.E.	(0.127)	(0.163)	(0.129)	(0.091)	(0.097)	(0.121)	(0.110)
R.I. p-value	[0.004]	[0.004]	[0.007]	[0.004]	[0.135]	[0.015]	[0.028]
Reduced-cost program	0.110	0.418	0.064	-0.012	0.021	0.058	0.034
S.E.	(0.102)	(0.181)	(0.096)	(0.074)	(0.069)	(0.081)	(0.084)
R.I. p-value	[0.367]	[0.104]	[0.513]	[0.862]	[0.790]	[0.516]	[0.730]
Number of students	1460	1476	1481	1474	1471	1467	1481
Adjusted R-squared	0.118	0.175	0.096	0.056	0.037	0.063	0.051
Difference between treatment effects	0.544***	0.624**	0.585***	0.393***	0.213	0.426**	0.415**
S.E.	(0.124)	(0.159)	(0.127)	(0.092)	(0.093)	(0.115)	(0.120)
R.I. p-value	[0.006]	[0.017]	[0.007]	[0.001]	[0.127]	[0.012]	[0.031]
Raw (unadjusted) values [§]							
Control group mean	0.144	5.973	0.616	0.334	0.358	0.611	0.216
Control group SD	1.000	9.364	1.920	2.207	2.762	4.163	0.437

Notes: Longitudinal sample includes 1,478 students from 38 schools who were tested at baseline as well as endline. All regressions control for stratification cell indicators. Heteroskedasticity-robust standard errors, clustered by school, in parentheses. Randomization inference p-values, clustered by school and stratified by stratification cell, in brackets; * p<0.05, ** p<0.01, *** p<0.001.

[†] PCA Leblango EGRA Score Index is constructed by weighting each of the 6 test modules (columns 2 through 7) using the first principal component of the 2013 endline control-group data as in Black and Smith (2006). The index is normalized by subtracting the baseline control-group mean and dividing by the endline control-group standard deviation, so that the control group mean for the index shows the control group's progress over the course of the year. Estimated effects are comparable for an alternative index that uses the unweighted mean across (normalized) test modules instead.

[§] Control Group Mean and SD are the raw (unstandardized) means and SDs computed using the endline data for control-group observations in the estimation sample.

Appendix Table A8

Program Impacts on Writing Test Scores, Without Controlling for Baseline Scores
(in SDs of the Control Group Endline Score Distribution)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PCA Writing Score Index †	African Name (Surname)	English Name (Given Name)	Ideas	Organization	Voice	Word Choice	Sentence Fluency	Conventions	Presentation
Full-cost program	0.449*	0.922***	1.312***	0.163	0.441	0.152	0.175	0.383	0.221	0.139
S.E.	(0.144)	(0.107)	(0.143)	(0.171)	(0.207)	(0.156)	(0.153)	(0.207)	(0.173)	(0.150)
R.I. p-value	[0.064]	[0.001]	[0.001]	[0.536]	[0.173]	[0.539]	[0.466]	[0.231]	[0.385]	[0.558]
Reduced-cost program	-0.159	0.435**	0.450**	-0.274	-0.316	-0.313***	-0.262	-0.330	-0.253	-0.330***
S.E.	(0.122)	(0.119)	(0.147)	(0.144)	(0.177)	(0.134)	(0.124)	(0.177)	(0.156)	(0.129)
R.I. p-value	[0.421]	[0.011]	[0.021]	[0.150]	[0.155]	[0.006]	[0.102]	[0.104]	[0.297]	[0.007]
Number of students	1373	1447	1374	1475	1475	1474	1474	1475	1475	1475
Adjusted R-squared	0.352	0.240	0.236	0.174	0.304	0.177	0.200	0.302	0.164	0.171
Difference between treatment effects	0.631***	0.577**	0.837***	0.435***	0.758***	0.465***	0.436***	0.711***	0.474***	0.469***
S.E.	(0.149)	(0.136)	(0.156)	(0.151)	(0.173)	(0.118)	(0.150)	(0.175)	(0.151)	(0.115)
R.I. p-value	[0.000]	[0.014]	[0.001]	[0.005]	[0.000]	[0.003]	[0.006]	[0.001]	[0.005]	[0.003]
Raw (unadjusted) values [§]										
Control group mean	0.482	0.593	0.350	0.141	0.286	0.164	0.166	0.267	0.116	0.175
Control group SD	1.000	0.685	0.533	0.372	0.594	0.393	0.416	0.590	0.339	0.396

Notes: Longitudinal sample includes 1,478 students from 38 schools who were tested at baseline as well as endline. All regressions control for stratification cell indicators. Heteroskedasticity-robust standard errors, clustered by school, in parentheses. Randomization inference p-values, clustered by school and stratified by stratification cell, in brackets; * p<0.05, ** p<0.01, *** p<0.001.

† PCA Writing Score Index is constructed by weighting each of the 9 test modules (columns 2 through 10) using the first principal component of the 2013 endline control-group data as in Black and Smith (2006). The index is normalized by subtracting the baseline control-group mean and dividing by the endline control-group standard deviation, so that the control group mean for the index shows the control group's progress over the course of the year. Estimated effects are comparable for an alternative index that uses the unweighted mean across (normalized) test modules instead.

§ Control Group Mean and SD are the raw (unstandardized) means and SDs computed using the endline data for control-group observations in the estimation sample.

Appendix Table A9

Program Impacts on Writing Test Scores, Excluding Stratification Cell for School That Completed Test in English
(in SDs of the Control Group Endline Score Distribution)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PCA Writing Score Index †	African Name (Surname)	English Name (Given Name)	Ideas	Organization	Voice	Word Choice	Sentence Fluency	Conventions	Presentation
Full-cost program	0.613***	0.933***	1.364***	0.372*	0.701***	0.350**	0.351*	0.638***	0.435**	0.328**
S.E.	(0.108)	(0.117)	(0.150)	(0.109)	(0.129)	(0.091)	(0.114)	(0.130)	(0.110)	(0.088)
R.I. p-value	[0.006]	[0.001]	[0.001]	[0.056]	[0.003]	[0.028]	[0.082]	[0.003]	[0.020]	[0.030]
Reduced-cost program	-0.004	0.473**	0.527***	-0.093	-0.079	-0.130**	-0.107	-0.093	-0.050	-0.155**
S.E.	(0.076)	(0.125)	(0.149)	(0.078)	(0.088)	(0.060)	(0.078)	(0.085)	(0.082)	(0.060)
R.I. p-value	[0.960]	[0.011]	[0.004]	[0.309]	[0.328]	[0.024]	[0.197]	[0.217]	[0.608]	[0.021]
Number of students	1262	1336	1263	1361	1361	1360	1360	1361	1361	1361
Adjusted R-squared	0.315	0.234	0.241	0.153	0.319	0.165	0.151	0.302	0.146	0.158
Difference between treatment effects	0.618***	0.460**	0.837***	0.464***	0.780***	0.480***	0.458***	0.731***	0.485***	0.484***
S.E.	(0.117)	(0.144)	(0.162)	(0.130)	(0.146)	(0.091)	(0.127)	(0.147)	(0.130)	(0.090)
R.I. p-value	[0.004]	[0.040]	[0.004]	[0.001]	[0.000]	[0.000]	[0.008]	[0.000]	[0.002]	[0.000]
Raw (unadjusted) values [§]										
Control group mean	0.222	0.527	0.274	0.061	0.131	0.084	0.075	0.108	0.037	0.098
Control group SD	0.585	0.671	0.486	0.239	0.338	0.278	0.264	0.310	0.190	0.298

Notes: Longitudinal sample includes 1,361 students from 35 schools who were tested at baseline as well as endline and are not from the stratification cell where one school conducted the writing test in English. All regressions control for stratification cell indicators as well as baseline values of the outcome variable, except for "Presentation" (column 10) which was not included in the baseline scores. Heteroskedasticity-robust standard errors, clustered by school, in parentheses. Randomization inference p-values, clustered by school and stratified by stratification cell, in brackets; * p<0.05, ** p<0.01, *** p<0.001.

† PCA Writing Score Index is constructed by weighting each of the 9 test modules (columns 2 through 10) using the first principal component of the 2013 endline control-group data as in Black and Smith (2006). The index is normalized by subtracting the baseline control-group mean and dividing by the endline control-group standard deviation, so that the control group mean for the index shows the control group's progress over the course of the year. Estimated effects are comparable for an alternative index that uses the unweighted mean across (normalized) test modules instead.

§ Control Group Mean and SD are the raw (unstandardized) means and SDs computed using the endline data for control-group observations in the estimation sample.

Appendix Table A10

Program Impacts on Oral English Test Scores & English Word Recognition
(in SDs of the Control Group Endline Score Distribution)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PCA Oral English Score Index†	Test 1 (Vocab.)	Test 1 (Count)	Test 2a (Vocab.)	Test 2a (Phrase Structure)	Test 2b (Vocab.)	Test 2b (Phrase Structure)	Test 3 (Vocab., Expressive: Objects)	Test 3 (Vocab., Expressive: People)	Reading English Words‡
Full-cost program	0.145	0.157	-0.118	-0.034	0.045	0.025	-0.114	0.306*	0.295	-0.290
S.E.	(0.099)	(0.099)	(0.097)	(0.095)	(0.114)	(0.100)	(0.113)	(0.105)	(0.117)	(0.135)
R.I. p-value	[0.356]	[0.279]	[0.338]	[0.813]	[0.804]	[0.856]	[0.518]	[0.093]	[0.134]	[0.248]
Reduced-cost program	-0.087	0.001	-0.115	-0.020	-0.113	-0.154	-0.213	-0.023	-0.099	-0.209
S.E.	(0.091)	(0.082)	(0.091)	(0.103)	(0.092)	(0.095)	(0.119)	(0.095)	(0.086)	(0.140)
R.I. p-value	[0.489]	[0.994]	[0.382]	[0.909]	[0.378]	[0.249]	[0.223]	[0.845]	[0.327]	[0.342]
Number of students	1481	1481	1481	1481	1481	1481	1481	1481	1481	1481
Adjusted R-squared	0.346	0.164	0.163	0.205	0.186	0.279	0.092	0.238	0.188	0.274
Difference between treatment effects	0.233*	0.156	-0.002	-0.014	0.158	0.179	0.098	0.330**	0.394***	-0.080
S.E.	(0.092)	(0.099)	(0.072)	(0.092)	(0.089)	(0.092)	(0.092)	(0.104)	(0.093)	(0.108)
R.I. p-value	[0.084]	[0.276]	[0.984]	[0.901]	[0.158]	[0.155]	[0.426]	[0.046]	[0.005]	[0.596]
Raw (unadjusted) values§										
Control group mean	0.101	2.048	0.294	0.501	0.807	1.826	2.092	2.327	1.585	1.792
Control group SD	1.000	1.888	0.620	0.911	1.209	1.928	2.217	2.133	1.839	4.184

Notes: Longitudinal sample includes 1,478 students from 38 schools who were tested at baseline as well as endline. All regressions control for stratification cell indicators and baseline values of the outcome variable except for Recognition of Printed English Words (column 10), which was not administered at baseline; missing values of control variables are dummied out. Heteroskedasticity-robust standard errors, clustered by school, in parentheses. Randomization inference p-values, clustered by school and stratified by stratification cell, in brackets; * p<0.05, ** p<0.01, *** p<0.001.

† PCA Oral English Score Index is constructed by weighting each of the 8 test modules (columns 2 through 9) using the first principal component of the 2013 endline control-group data as in Black and Smith (2006). The index is normalized by subtracting the baseline control-group mean and dividing by the endline control-group standard deviation, so that the control group mean for the index shows the control group's progress over the course of the year. Estimated effects are comparable for an alternative index that uses the unweighted mean across (normalized) test modules instead.

‡ Reading English Words is not part of the Oral English examination (and is not included in the computation of the overall PCA index).

§ Control Group Mean and SD are the raw (unstandardized) means and SDs computed using the endline data for control-group observations in the estimation sample.