BGSE Development Summer School

Lecture 5:

What have we learned from RCTs: Education

Professors: Pamela Jakiela and Owen Ozier

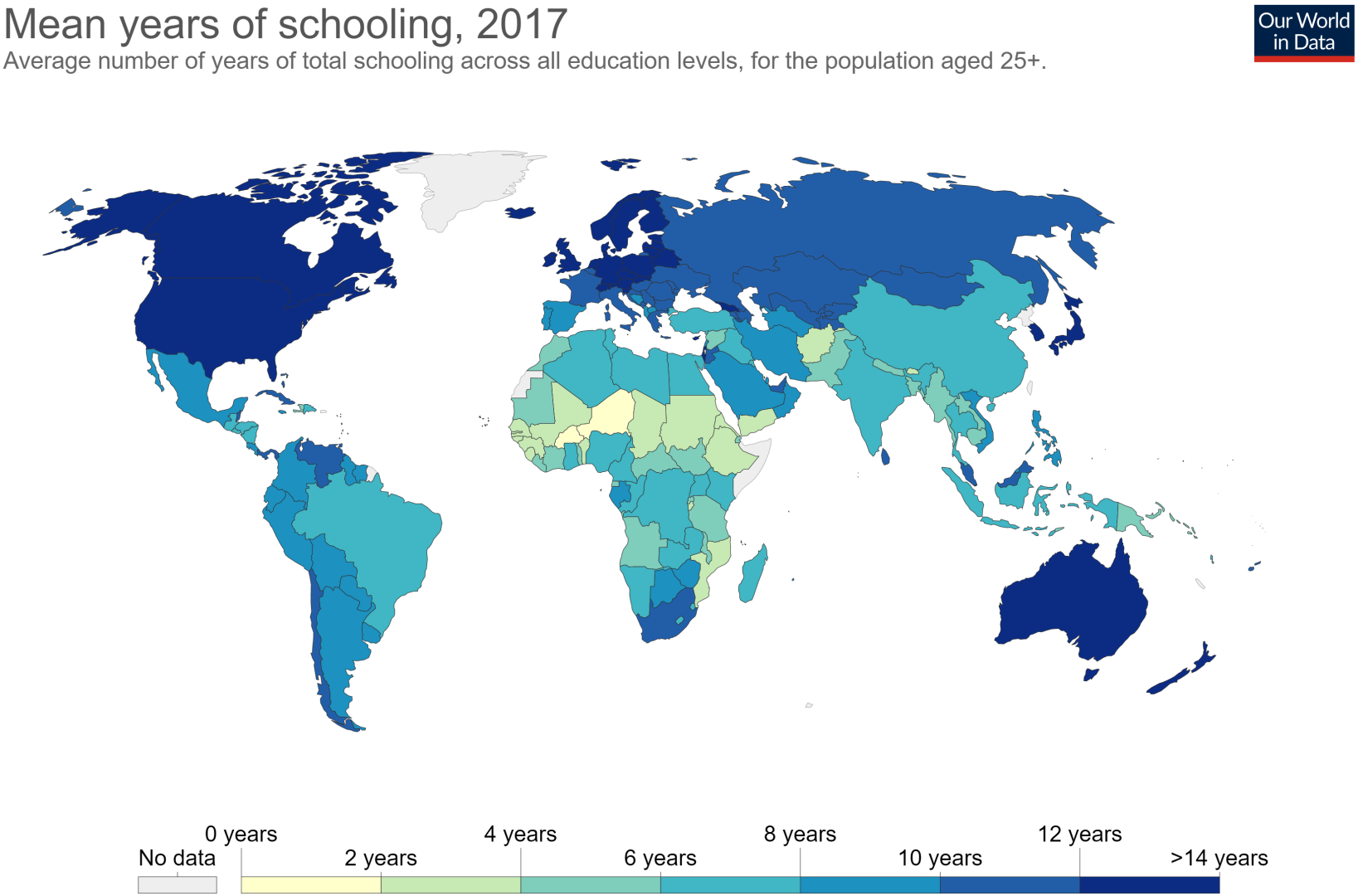
Some themes to bear in mind

- What to measure as an outcome
- When to measure it
- How to measure it
- What class of "interventions" to consider
- Why do things "work" and "not work" - a model?
- For each study, what **isn't** being studied?
- What won't be easy to study this way
- Mispronouncing "research"

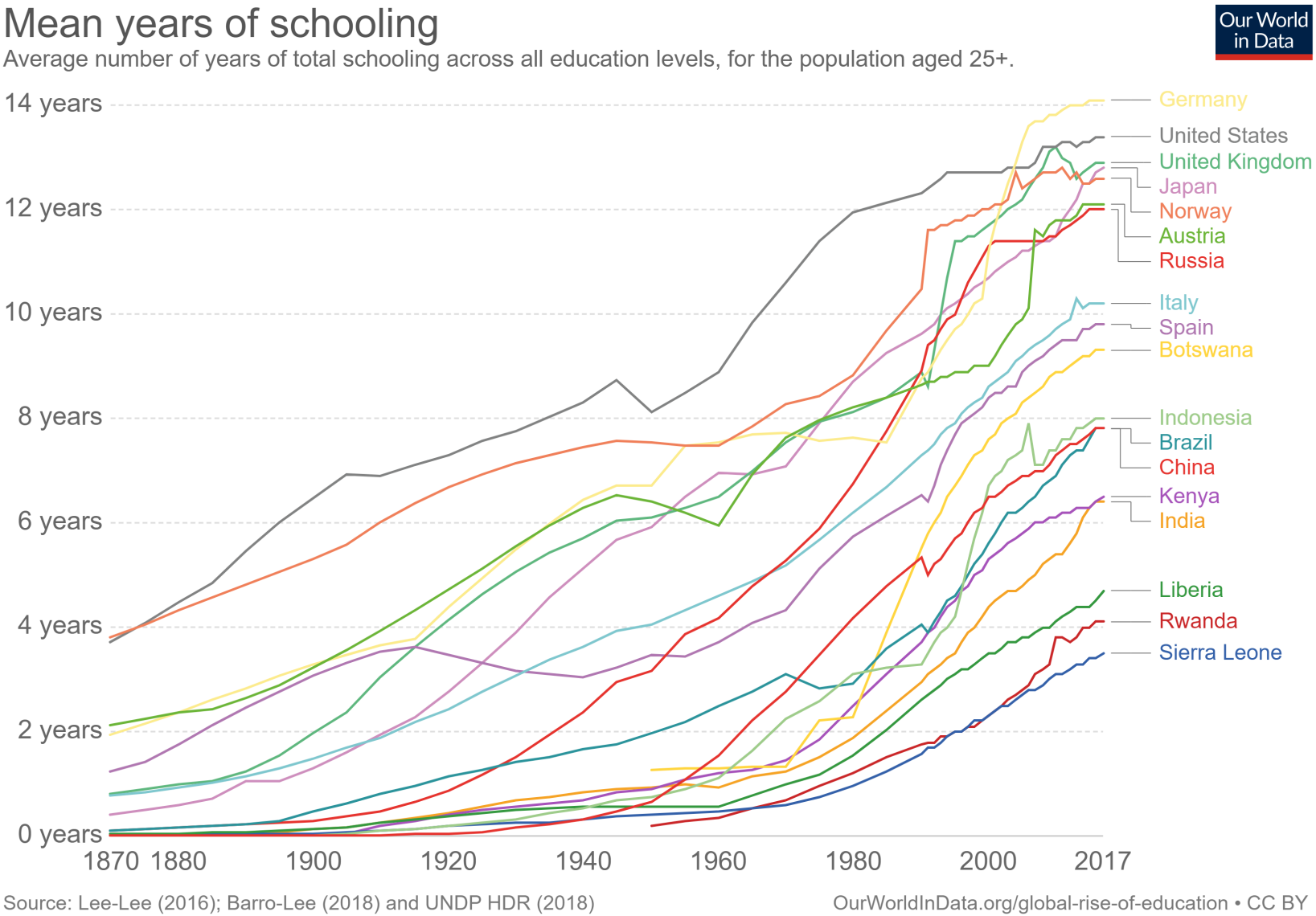
RCTs in Education, Part 1:

Background

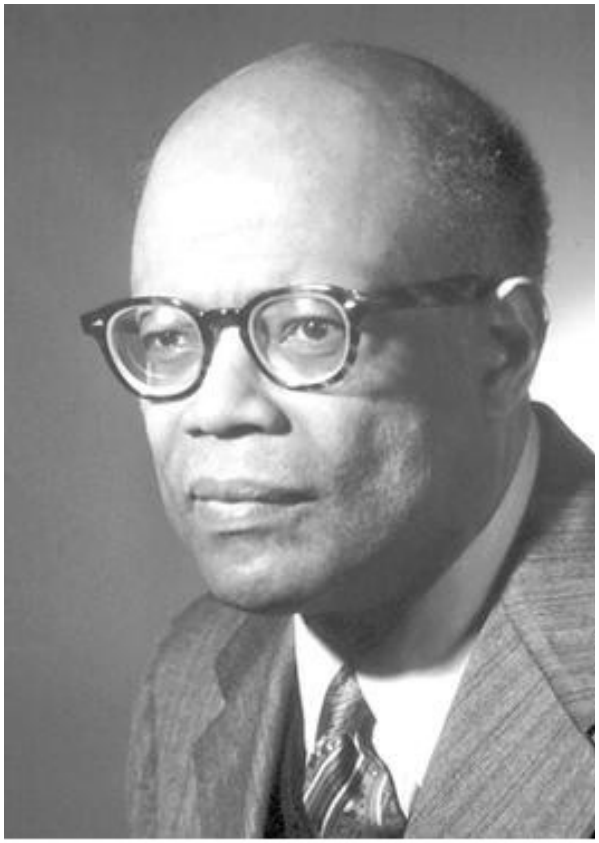
Attainment: Some History



Attainment: Some History



Arthur Lewis (1979 Nobel) on education



Education and Economic Development[1]

By

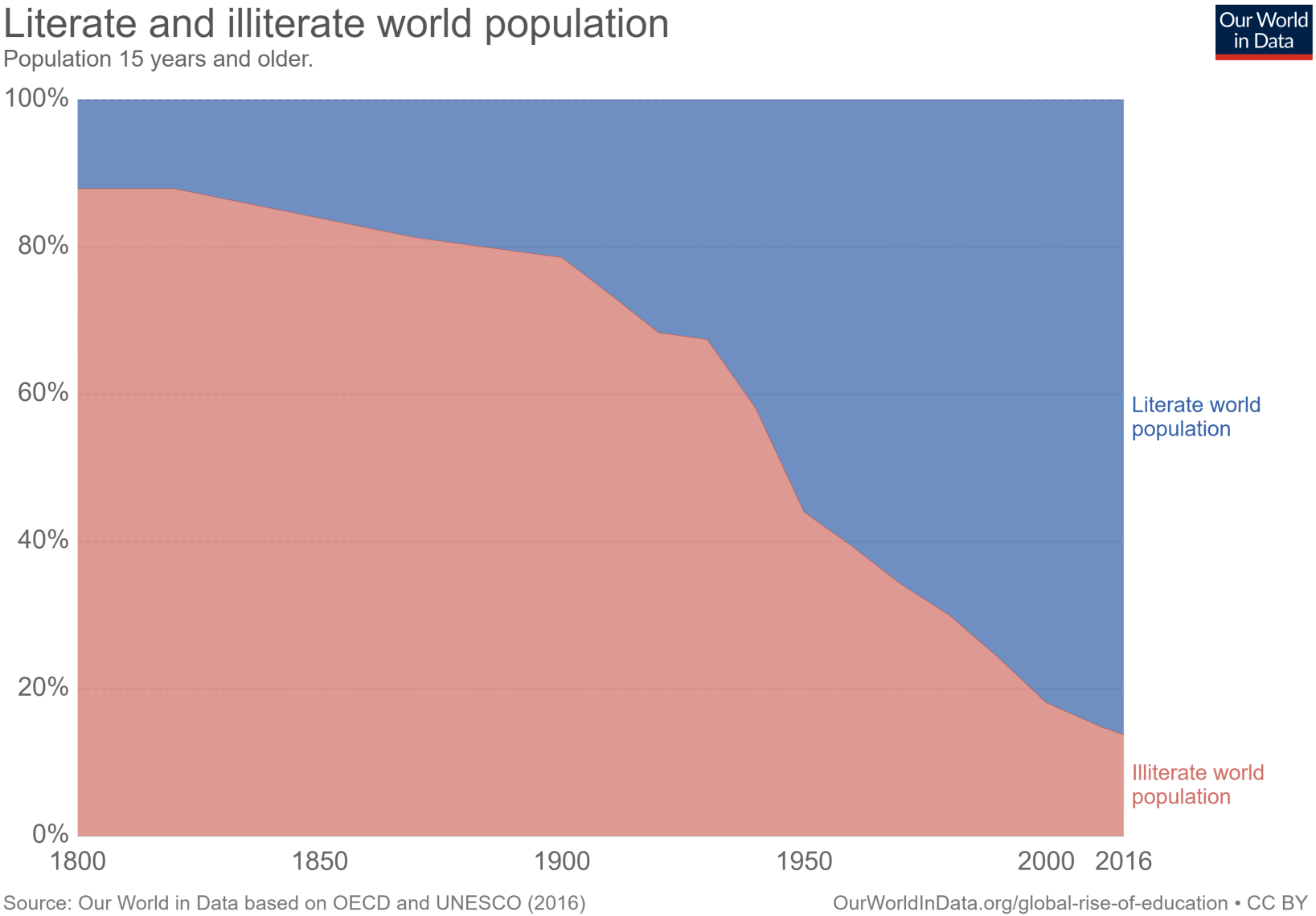
W. Arthur Lewis.

Poor countries cannot afford to pay for as much education as richer countries. They have therefore to establish priorities in terms both of quantity and of quality.

The requirements of economic development help in setting priorities, but they are not over-riding. Education was not invented in order to enable men to produce more goods and services. The purpose of education is to enable men to understand better the world in which they live, so that they may more fully express their potential capacities, whether spiritual, intellectual or material. Indeed, through the centuries the traditional attitude of "practical" men towards education has been that it unfits its recipients for useful work. Certainly, most people would agree that education is desirable even if it contributed nothing to material output.

From the standpoint of economic development, one may distinguish between types of education which increase productive capacity and types which do not. Teaching an African cook to read may increase his enjoyment of life, but will not necessarily make him a better cook. Education of the former kind I have called investment education, while the latter kind is called consumption education. From the standpoint of economic development, investment education has a high priority, but consumption education is on a par with other forms of consumption. The money spent on teaching cooks to read might equally be spent on giving them pure water supplies, or radios, or better housing, and must therefore compete in the context of all other possible uses of resources. In this perspective the needs of economic development help to determine the minimum amount which must be spent on education. How much to spend above this minimum depends on how rich the society is, and on competing claims. This article is confined to seeking to discover the nature and limits of investment education.

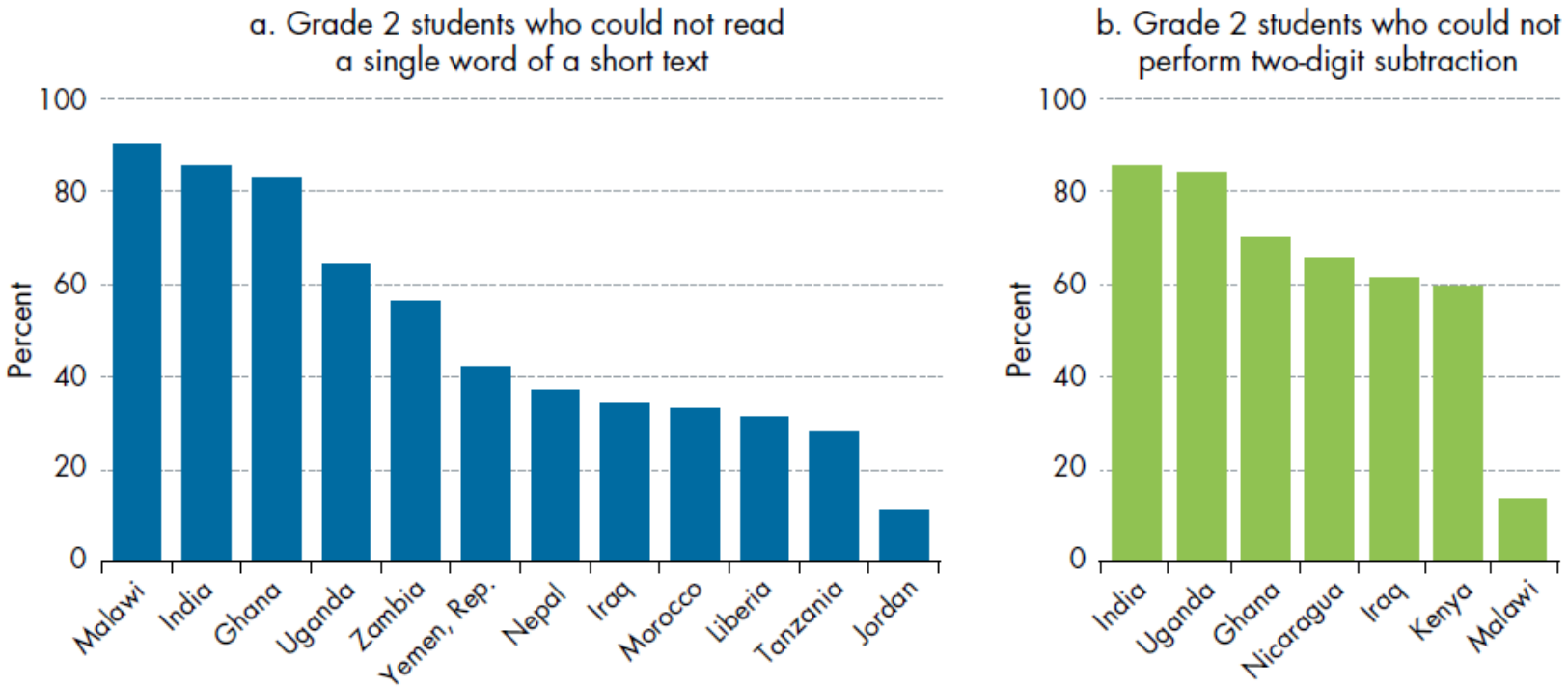
Literacy


Literate and illiterate world population
Population 15 years and older.
Our World in Data
100%
80%
60%
40%
20%
0%
1800
1850
1900
1950
2000
2016
Literate world population
Illiterate world population
Source: Our World in Data based on OECD and UNESCO (2016)
OurWorldInData.org/global-rise-of-education • CC BY

World Development Report 2018

Figure O.1 Shortfalls in learning start early

Percentage of grade 2 students who could not perform simple reading or math tasks, selected countries



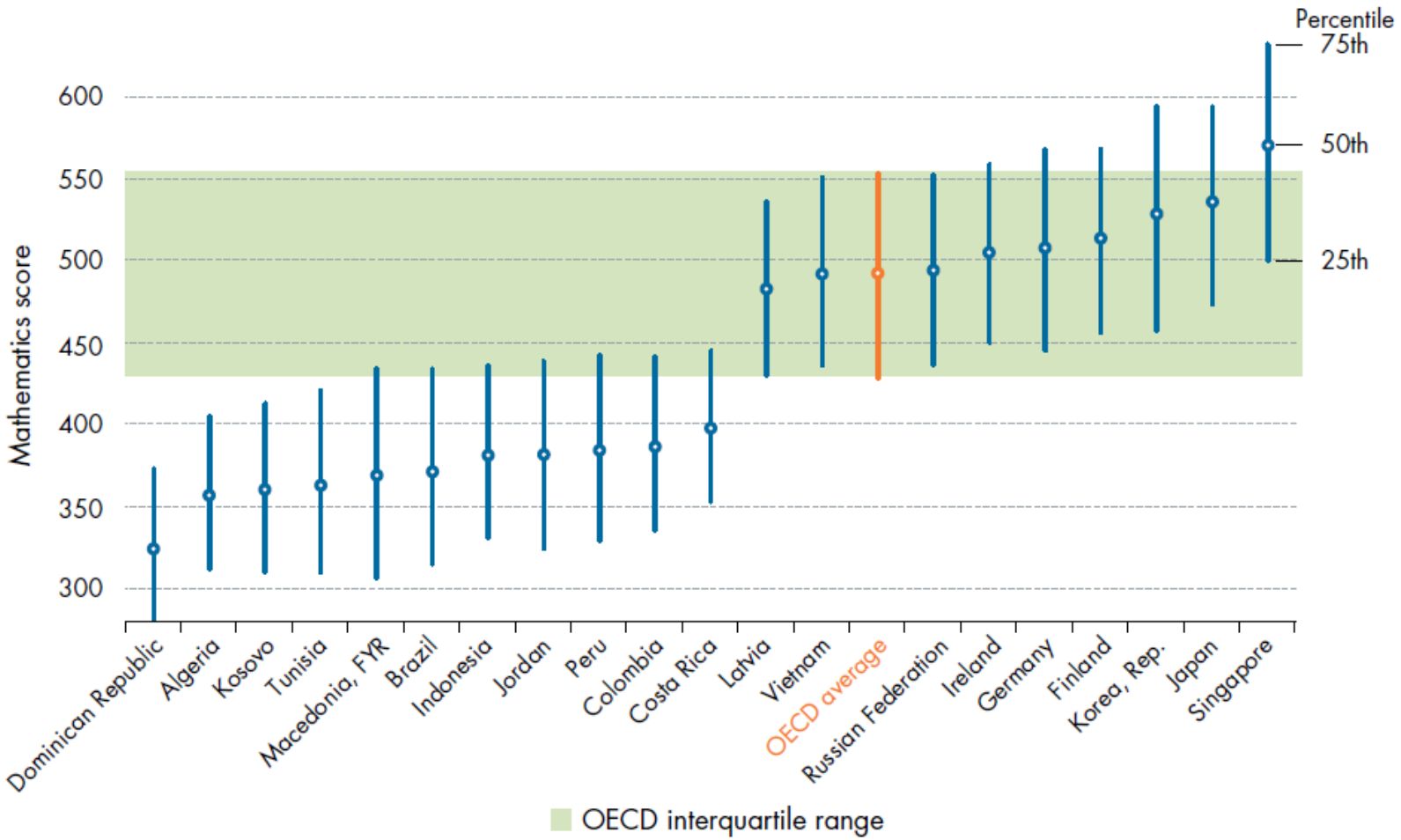
Sources: WDR 2018 team, using reading and mathematics data for Kenya and Uganda from Uwezo, Annual Assessment Reports, 2015 (http://www.uwezo.net/); reading and mathematics data for rural India from ASER Centre (2017); reading data for all other countries from U.S. Agency for International Development (USAID), Early Grade Reading Barometer, 2017, accessed May 30, 2017 (http://www.earlygradereadingbarometer.org/); and mathematics data for all other countries from USAID/RTI Early Grade Mathematics Assessment intervention reports, 2012–15 (https://shared.rti.org/sub-topic/early-grade-math-assessment-egma). Data at http://bit.do/WDR2018-Fig_O-1.

Note: These data typically pertain to selected regions in the countries and are not necessarily nationally representative. Data for India pertain to rural areas.

World Development Report 2018

Figure O.2 In several countries, the 75th percentile of PISA test takers performs below the 25th percentile of the OECD average

Performance of 25th, 50th, and 75th percentiles in 2015 PISA mathematics assessment, selected countries

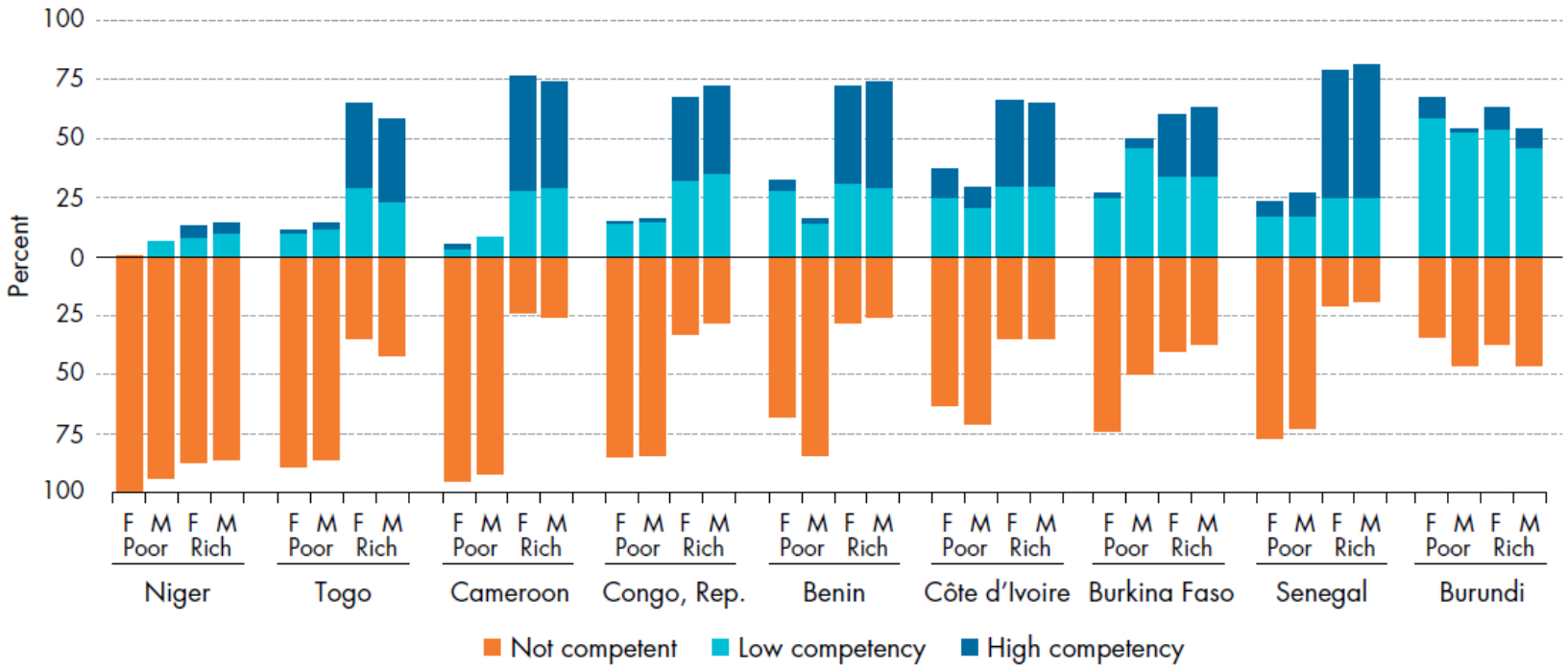


Source: WDR 2018 team, using data from Programme for International Student Assessment (PISA) 2015 (OECD 2016). Data at http://bit.do/WDR2018-Fig_O-2.

World Development Report 2018

Figure O.3 Children from poor households in Africa typically learn much less

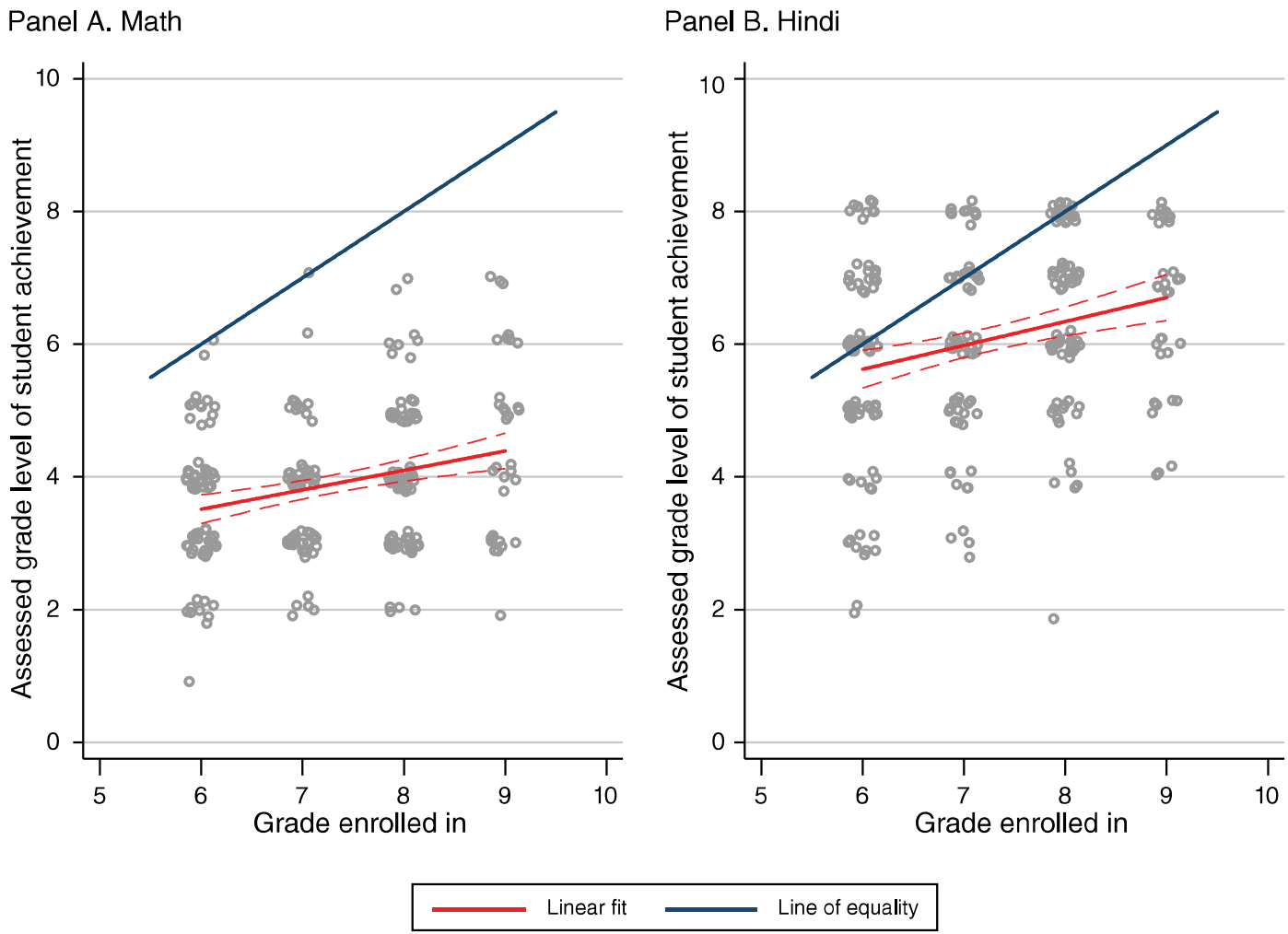
Percentage of grade 6 PASEC test takers in 2014 who scored above (blue) and below (orange) the sufficiency level on reading achievement: poorest and richest quintiles by gender, selected countries



Source: WDR 2018 team, using data from World Bank (2016b). Data at http://bit.do/WDR2018-Fig_O-3.

Note: Socioeconomic quintiles are defined nationally. "Not competent" refers to levels 0–2 in the original coding and is considered below the sufficiency level for school continuation; "low competency" refers to level 3; and "high competency" refers to level 4. F = female; M = male; PASEC = Programme d'Analyse des Systèmes Éducatifs de la Confemen.

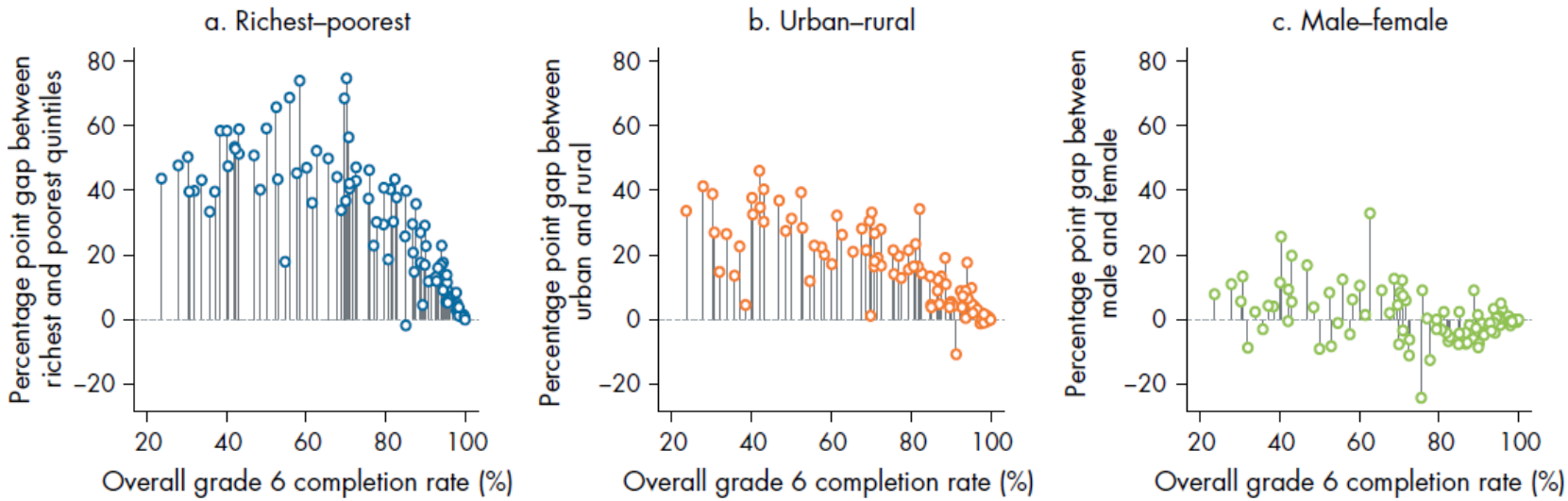
Muralidharan, Singh, Ganimian (2019 AER)



World Development Report 2018

Figure O.6 School completion is higher for richer and urban families, but gender gaps are more context-dependent

Gaps in grade 6 completion rates (percent) for 15- to 19-year-olds, by wealth, location, and gender



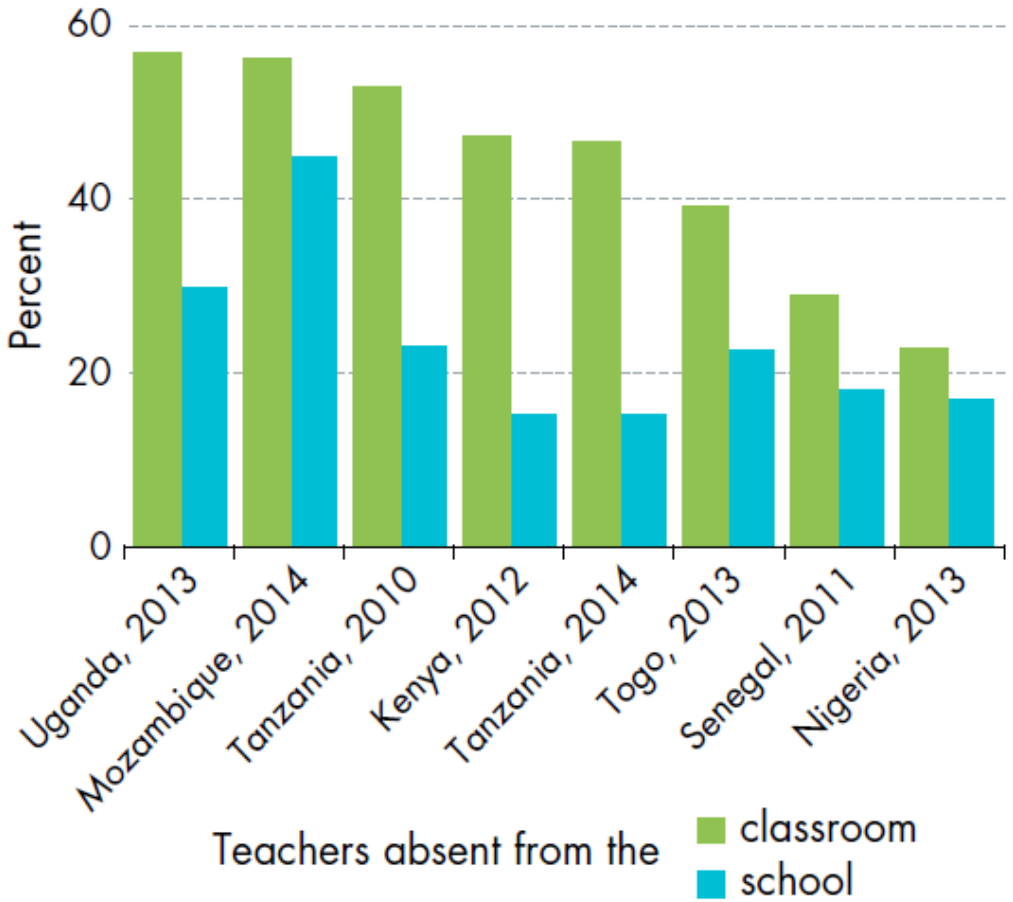
Source: WDR 2018 team, using data from Filmer (2016). Data at http://bit.do/WDR2018-Fig_O-6.

Note: The data presented are the latest available by country, 2005–14. Each vertical line indicates the size and direction of the gap for a country.

World Development Report 2018

Figure O.9 In Africa, teachers are often absent from school or from classrooms while at school

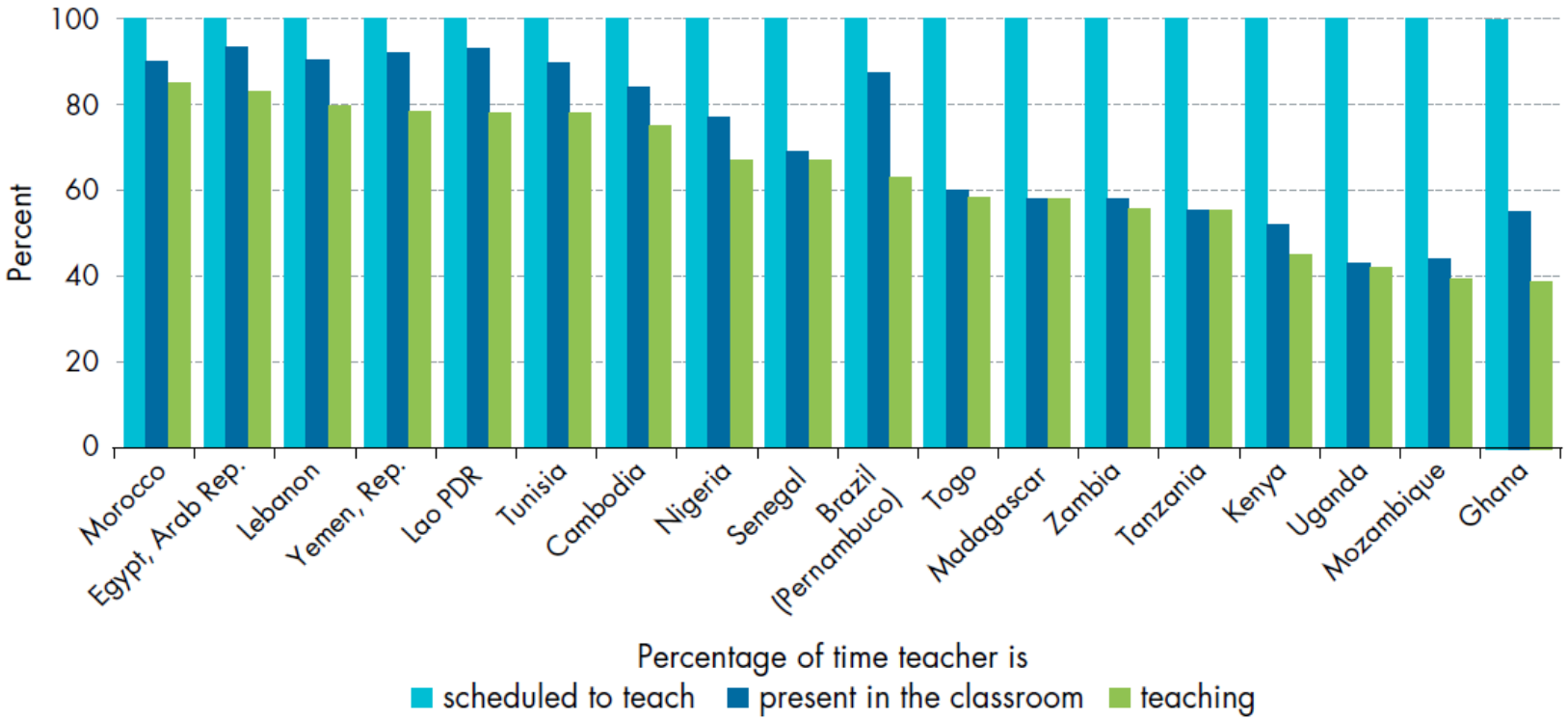
Percentage of teachers absent from school and from class on the day of an unannounced visit, participating countries



World Development Report 2018

Figure 3.10 A lot of official teaching time is lost

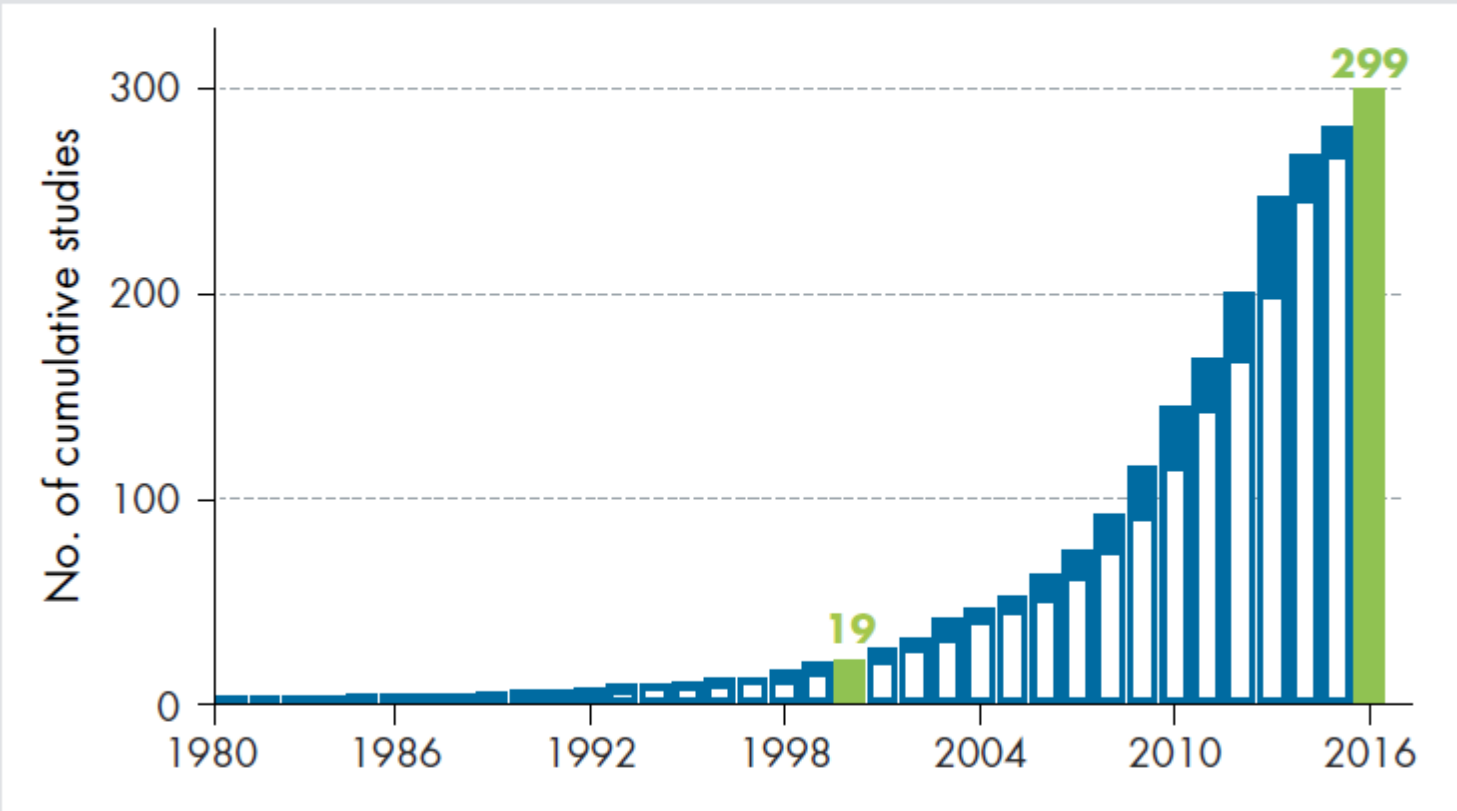
Percentage of time officially allocated to schooling that a teacher is scheduled to teach, is present in the classroom, and is actually teaching



Sources: WDR 2018 team, using data from Abadzi (2009): Brazil (Pernambuco state), Ghana, Morocco, and Tunisia; Benveniste, Marshall, and Araujo (2008): Cambodia; Benveniste, Marshall, and Santibañez (2007): Lao People's Democratic Republic; Millot and Lane (2002): Arab Republic of Egypt, Lebanon, and Republic of Yemen; World Bank (2016a): Madagascar; World Bank (2016b): Zambia; World Bank's Service Delivery Indicators, 2012–13 (http://www.worldbank.org/sdi): Kenya, Mozambique, Nigeria, Senegal, Tanzania, Togo, and Uganda. Data at http://bit.do/WDR2018-Fig_3-10.

World Development Report 2018

Figure S4.1 The number of experimental and quasi-experimental studies of interventions to improve learning has mushroomed in recent decades

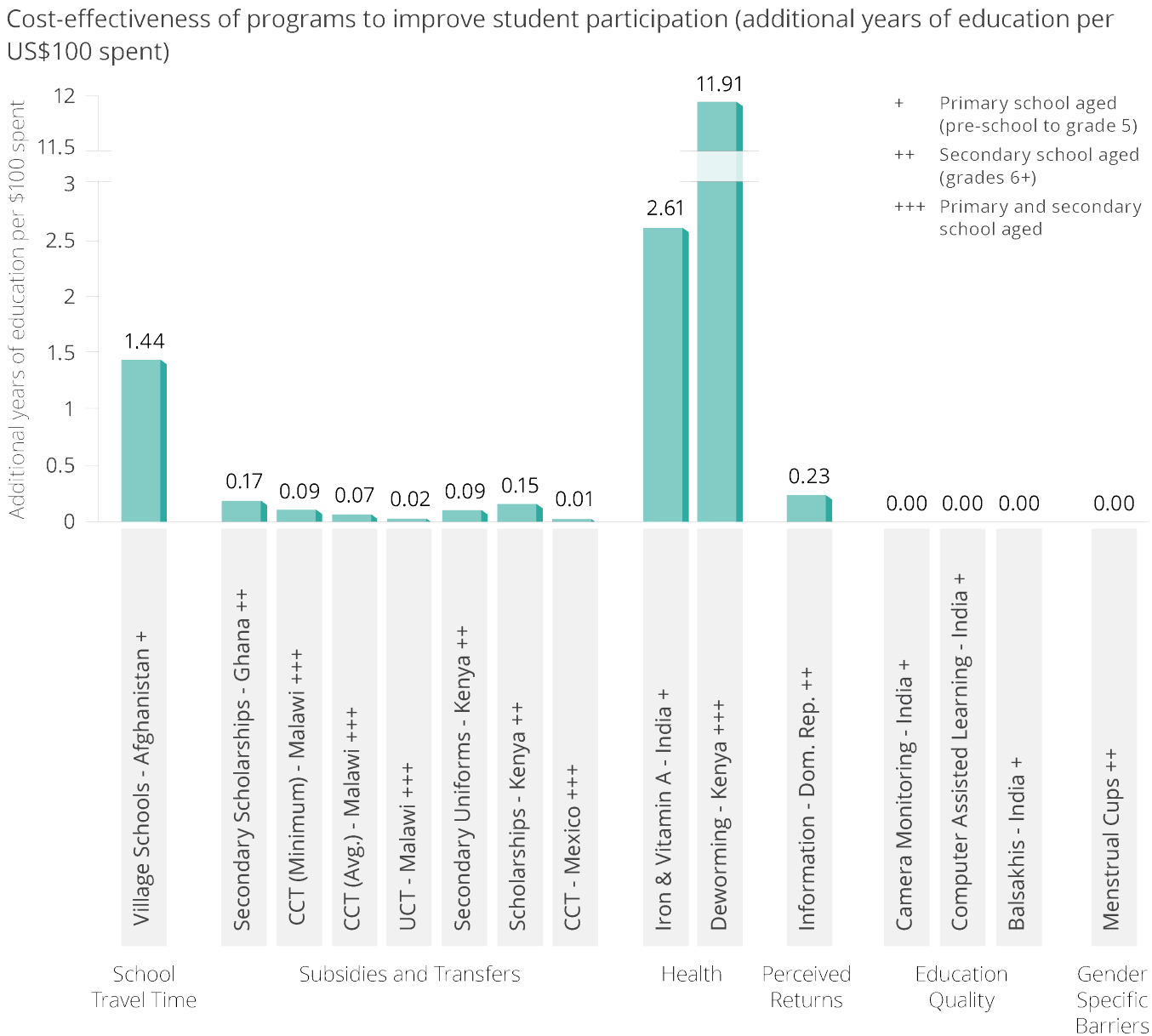


Sources: WDR 2018 team, using data from 3ie (2016) and Evans and Popova (2016b). Data at http://bit.do/WDR2018-Fig_S4-1.

RCTs in Education, Part 2:

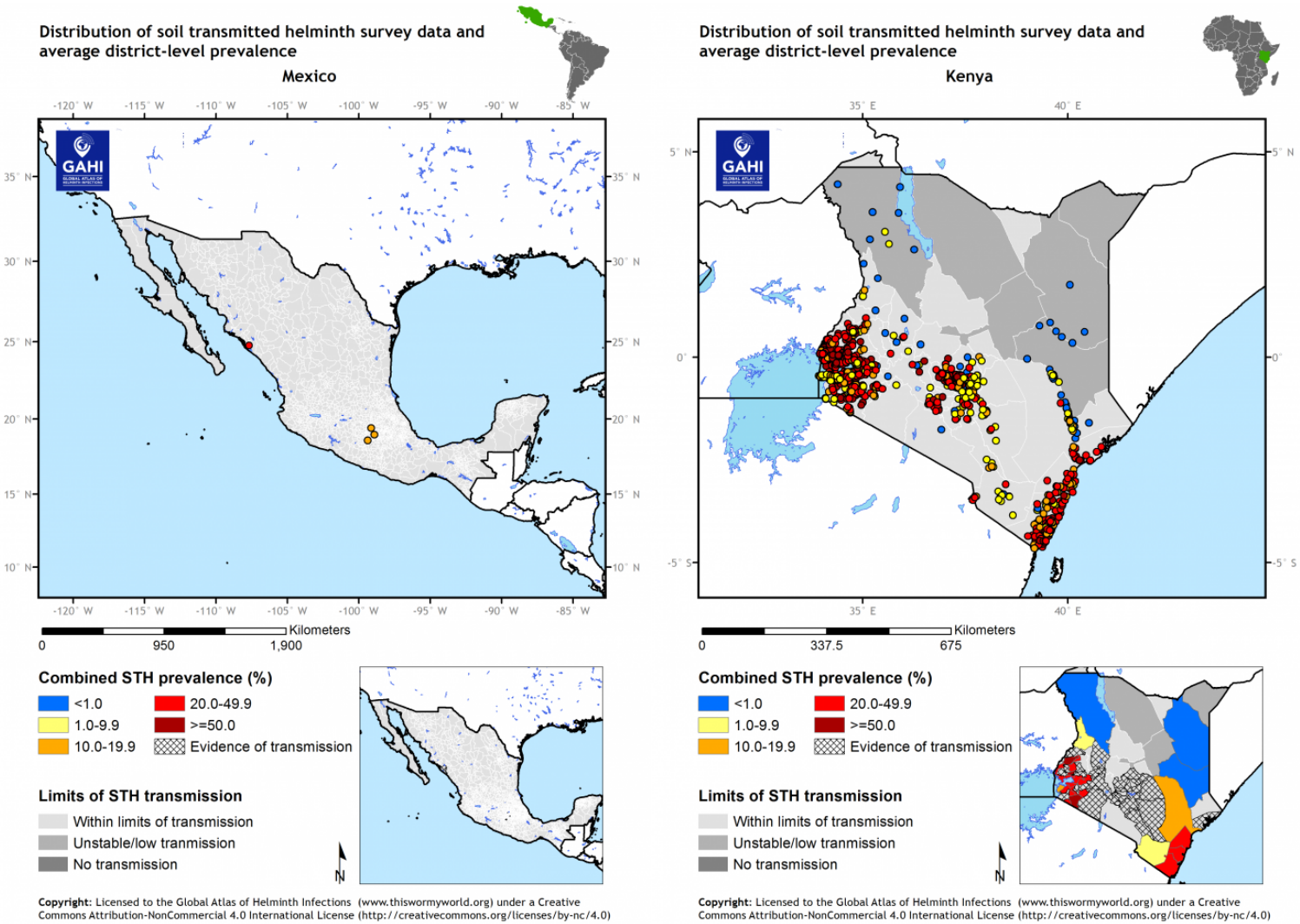
Evidence:
things that work...
well, poorly, or not at all

JPAL Cost-Effectiveness ("participation")



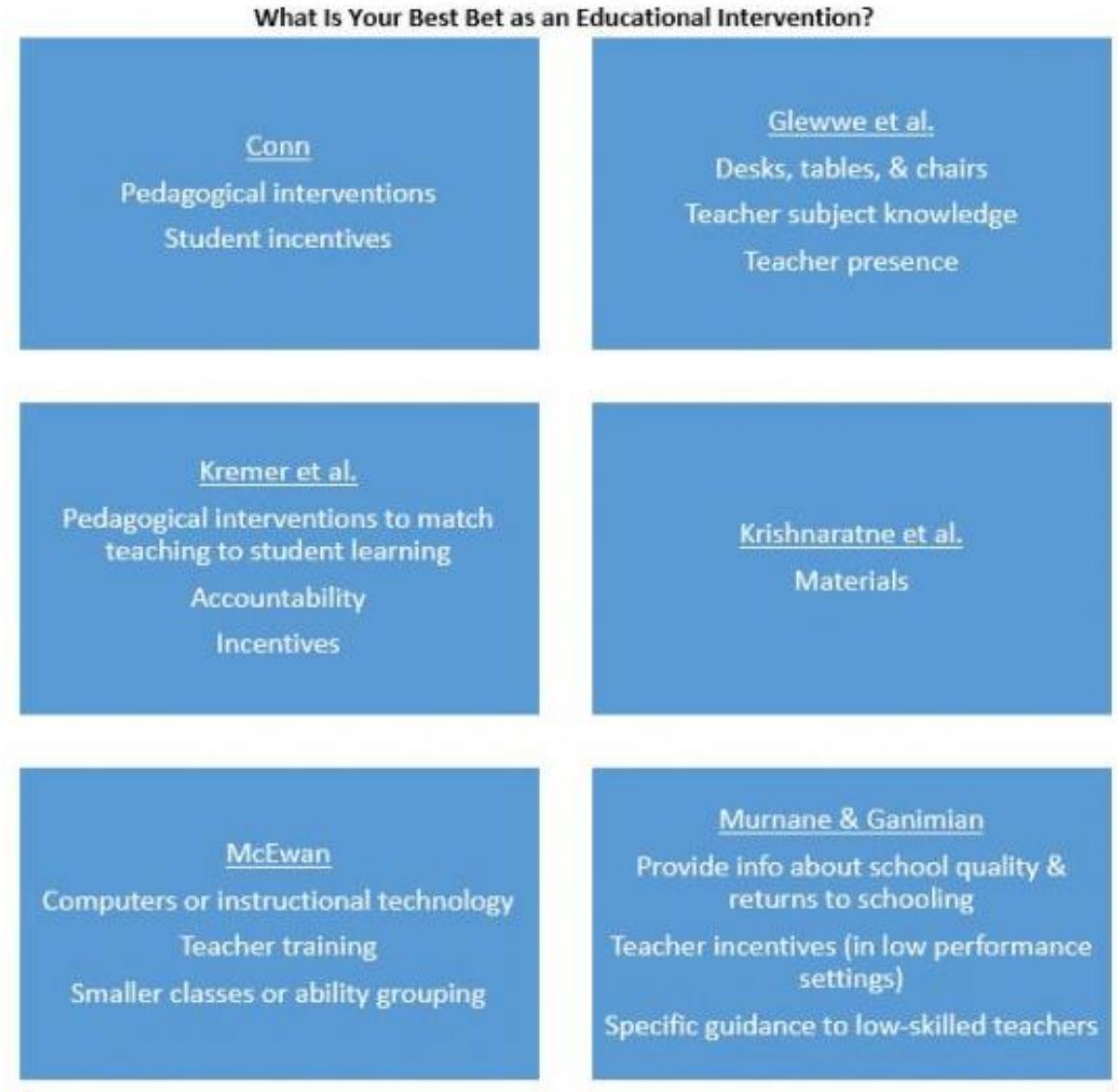
From JPAL (2019): "Reducing costs to increase school participation"

Caveats: validity, objective function



From GAHI (2019)

Evans and Popova 2016



From Evans and Popova blog:
"227 studies later, what actually works to improve learning in developing countries?"

Evans and Popova 2016

Their summary of "what works," seen in multiple reviews:

- Pedagogical interventions that match teaching to individual student learning levels
  - Example: Banerjee, Cole, Duflo, and Linden (2007)
  - Example: Duflo, Dupas, and Kremer (2011)
  - Example: Piper and Korda (2010)
- Individualized, repeated teacher training, associated with a specific method or task
- Accountability-boosting interventions

Banerjee, Cole, Duflo, Linden 2007 (India; Grades 3,4)

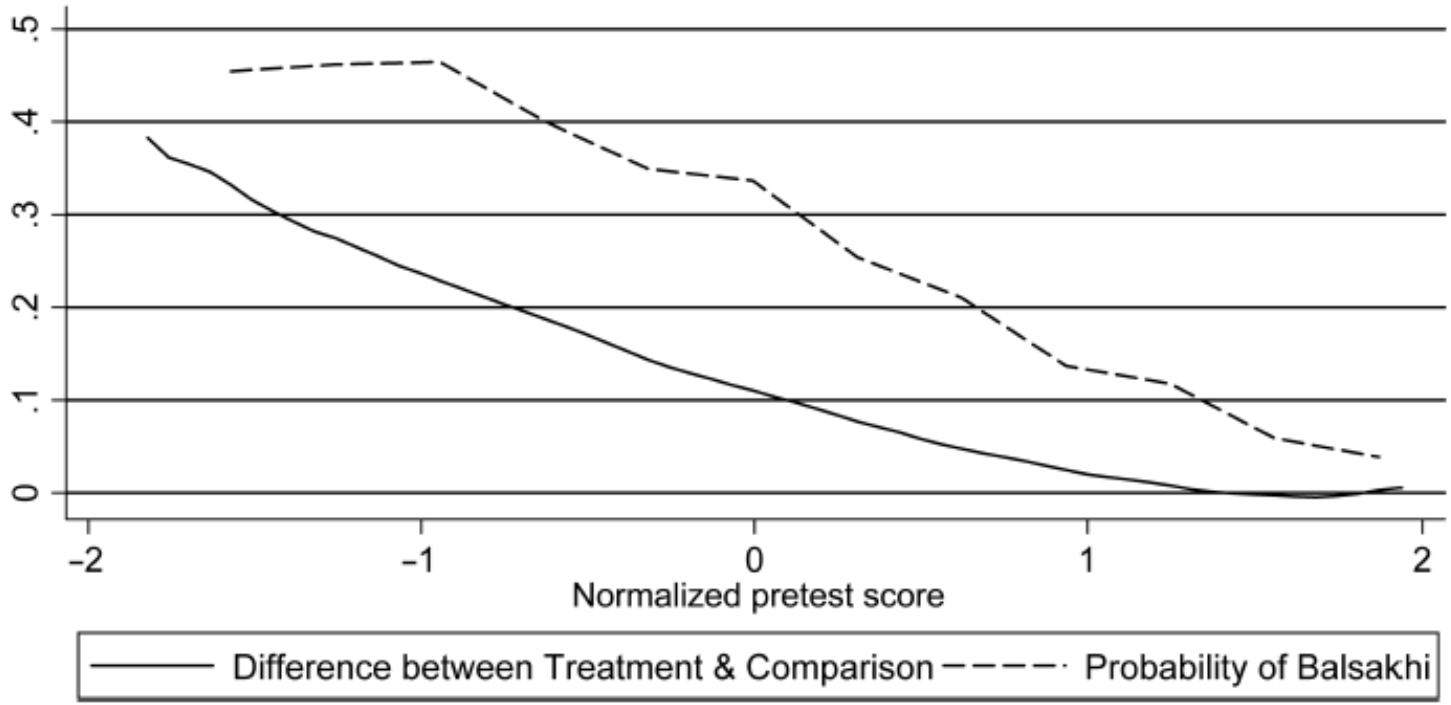


FIGURE I

Program Effect and Assignment Probability as a Function of Pretest Score

Note: The dashed line presents the probability a child is assigned to a balsakhi as a function of her place in the pretest score distribution. The solid line presents the difference in test score gains between children in treatment and comparison groups as a function of their place in the pretest score distribution. The values are computed using locally weighted regressions with a bandwidth of 1.5.

Banerjee, Cole, Duflo, Linden 2007 (India; Grades 3,4)

TABLE IV
IMPACT OF THE CAL PROGRAM, BY YEAR

	Number of observations	Dependent variable: Test score improvement (posttest − pretest)		
		Math	Language	Total
	(1)	(2)	(3)	(4)
A: Effect of the CAL program				
Vadodara both years	11,255	0.394	−0.025	0.191
		(0.074)	(0.082)	(0.083)
Vadodara Year 2	5,732	0.347	0.013	0.208
		(0.076)	(0.069)	(0.074)
Vadodara Year 3	5,523	0.475	−0.005	0.225
		(0.068)	(0.042)	(0.051)
B: Balsakhi and CAL program: Main effects and interactions (Vadodara, Year 2)				
CAL	5,732	0.408	0.017	0.242
		(0.087)	(0.084)	(0.087)
Balsakhi	—	0.371	0.229	0.315
		(0.112)	(0.104)	(0.112)
CAL * balsakhi	—	−0.144	−0.020	−0.086
		(0.141)	(0.134)	(0.141)

This table reports the impact of the CAL program. In Panel A, each cell represents a separate regression, of test score gain on a dummy for treatment school, controlling for initial pretest score. In Panel B, each column represents a regression, of test score improvement on a dummy for the CAL program, a dummy for the Balsakhi program, and an interaction term, as well as a control for initial pretest score. Standard errors, clustered at the school-grade level, are given in parentheses. Normalized test score improvement is the difference between posttest and pretest. The total score is the sum of the normalized math and language scores.

Duflo, Dupas, Kremer 2011 (Kenya; Grade 1)

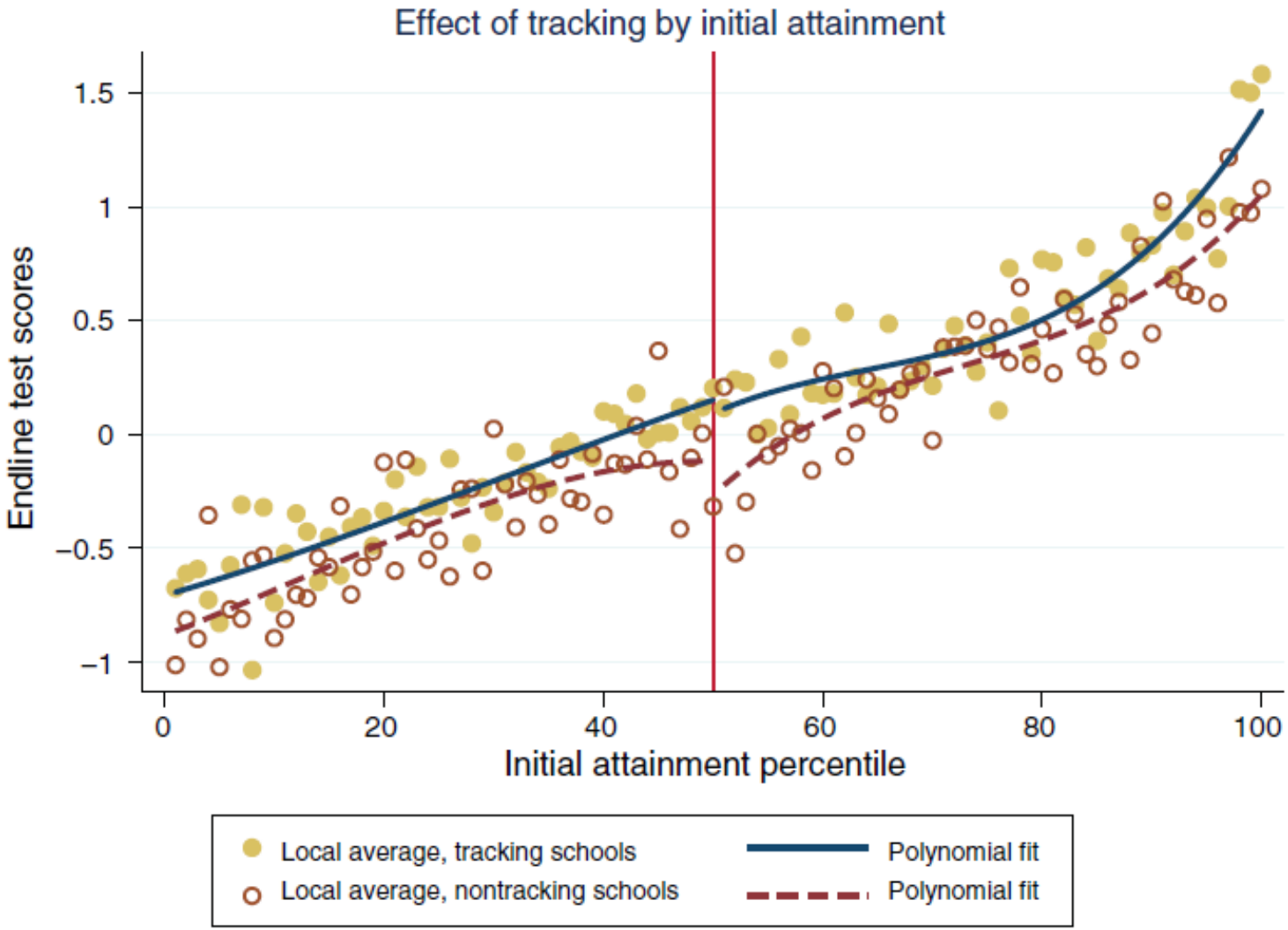
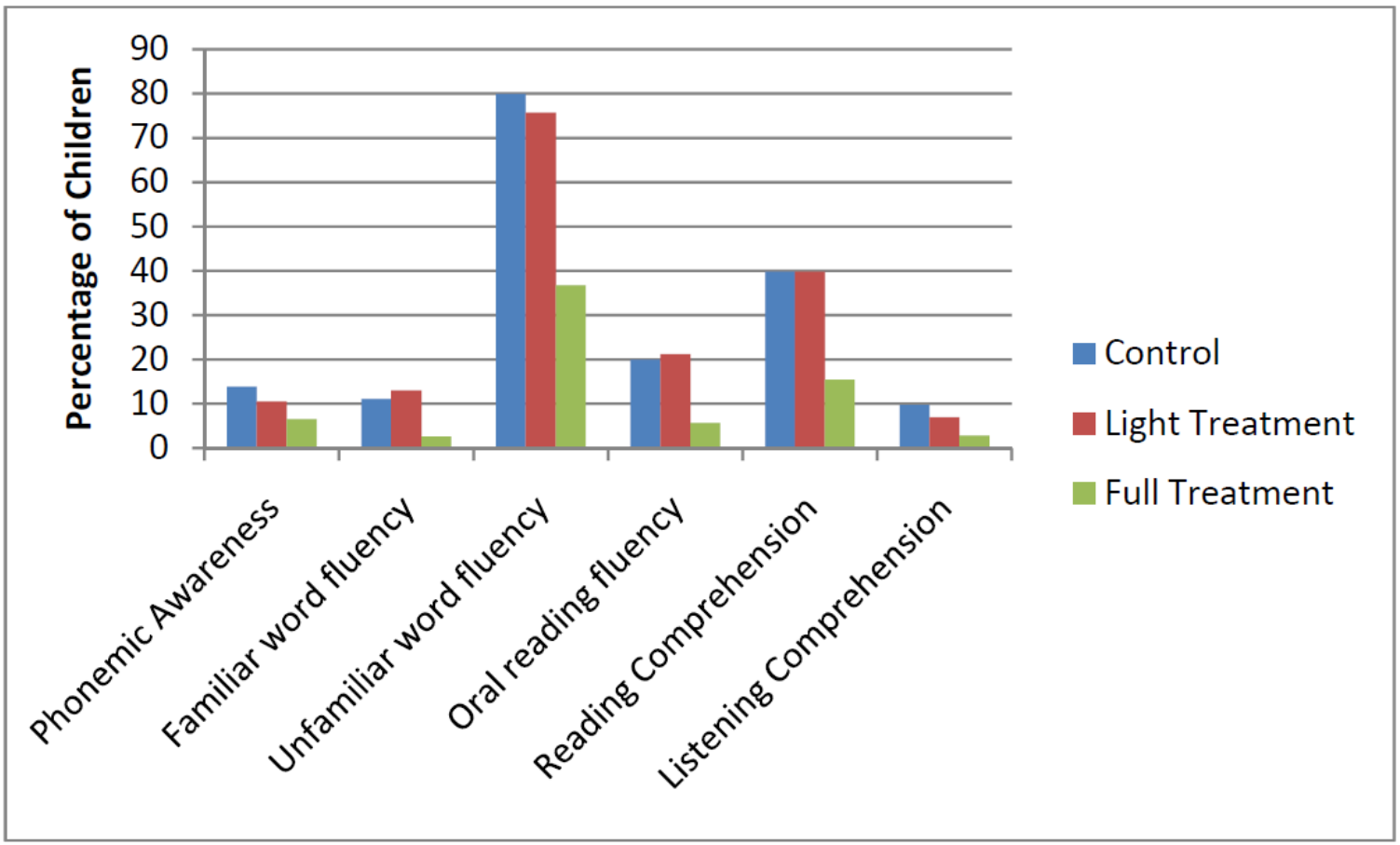


FIGURE 3. LOCAL POLYNOMIAL FITS OF ENDLINE SCORE BY INITIAL ATTAINMENT

Piper and Korda 2010 (Liberia; Grades 2,3)

Figure 2: Zero Scores, by Treatment Group and Section



Basic things that might be a good idea

We use these things all the time in schools we have attended.
Shouldn't policymakers just make sure people have these things?

- Textbooks
- Light
- Computers

Great evidence on things that didn't work

Some informative highlights

- Textbooks at the wrong level
- Textbooks locked in cabinets
- Substitute light sources
- Computers without content

Books: Glewwe, Kremer, and Moulin (2009)

American Economic Journal: Applied Economics 2009, 1:1, 112–135
http://www.aeaweb.org/articles.php?doi=10.1257/app.1.1.112

Many Children Left Behind? Textbooks and Test Scores in Kenya†

By Paul Glewwe, Michael Kremer, and Sylvie Moulin*

A randomized evaluation in rural Kenya finds, contrary to the previous literature, that providing textbooks did not raise average test scores. Textbooks did increase the scores of the best students (those with high pretest scores) but had little effect on other students. Textbooks are written in English, most students' third language, and many students could not use them effectively. More generally, the curriculum in Kenya, and in many other developing countries, tends to be oriented toward academically strong students, leaving many students behind in societies that combine a centralized educational system; the heterogeneity in student preparation associated with rapid educational expansion; and disproportionate elite power.

Books: Sabarwal, Evans, Marshak (2014)

The Permanent Input Hypothesis

The Case of Textbooks and (No) Student Learning in Sierra Leone

Shwetlena Sabarwal
David K. Evans
Anastasia Marshak

Abstract

A textbook provision program in Sierra Leone demonstrates how volatility in the flow of government-provided learning inputs to schools can induce storage of these inputs by school administrators to smooth future consumption. This process in turn leads to low current utilization of inputs for student learning. A randomized trial of a public program providing textbooks to primary schools had modest positive impacts on teacher behavior but no impacts on student performance. In many treatment schools, student access to textbooks did not actually increase because a large majority of the books were stored rather than distributed to students. At the same time, the propensity to save books was positively correlated with uncertainty on the part of head teachers regarding government transfers of books. The evidence suggests that schools that have high uncertainty with respect to future transfers are more likely to store a high proportion of current transfers. These results show that reducing uncertainty in school input flows could result in higher current input use for student learning. For effective program design, public policy programs must take forward-looking behavior among intermediate actors into account.

Lanterns: Kudo, Shonchoy, and Takahashi (2017)

The World Bank Economic Review, 2017, 1–25
doi: 10.1093/wber/lhw073
Article

OXFORD

Can Solar Lanterns Improve Youth Academic Performance? Experimental Evidence from Bangladesh

Yuya Kudo,* Abu S. Shonchoy, and Kazushi Takahashi

Abstract

We conducted an experimental intervention in unelectrified areas of northern Bangladesh to investigate the effectiveness of solar products in improving children's educational achievement. We found that treated households substituted solar lanterns for kerosene-based lighting products, helping to decrease total household expenditure. Solar lanterns increased the children's home-study hours, particularly at night and before exams. The solar lanterns initially led to an increase in school attendance, but this effect diminished over time. However, the increased study hours and initial improvement in school attendance did not translate into improved academic performance. Varying the number of solar products within the treated households did not alter these results. Analyses that exploited the school grade treatment intensity also provided no evidence suggesting that spillover effects explained the "no academic performance effects." These findings suggest that improving the home-study environment solely through the provision of solar products may have a limited impact on children's educational achievement.

Computers: Cristia, Ibarrarán, Cueto, Santiago, Severín

American Economic Journal: Applied Economics 2017, 9(3): 295–320
https://doi.org/10.1257/app.20150385

Technology and Child Development: Evidence from the One Laptop per Child Program†

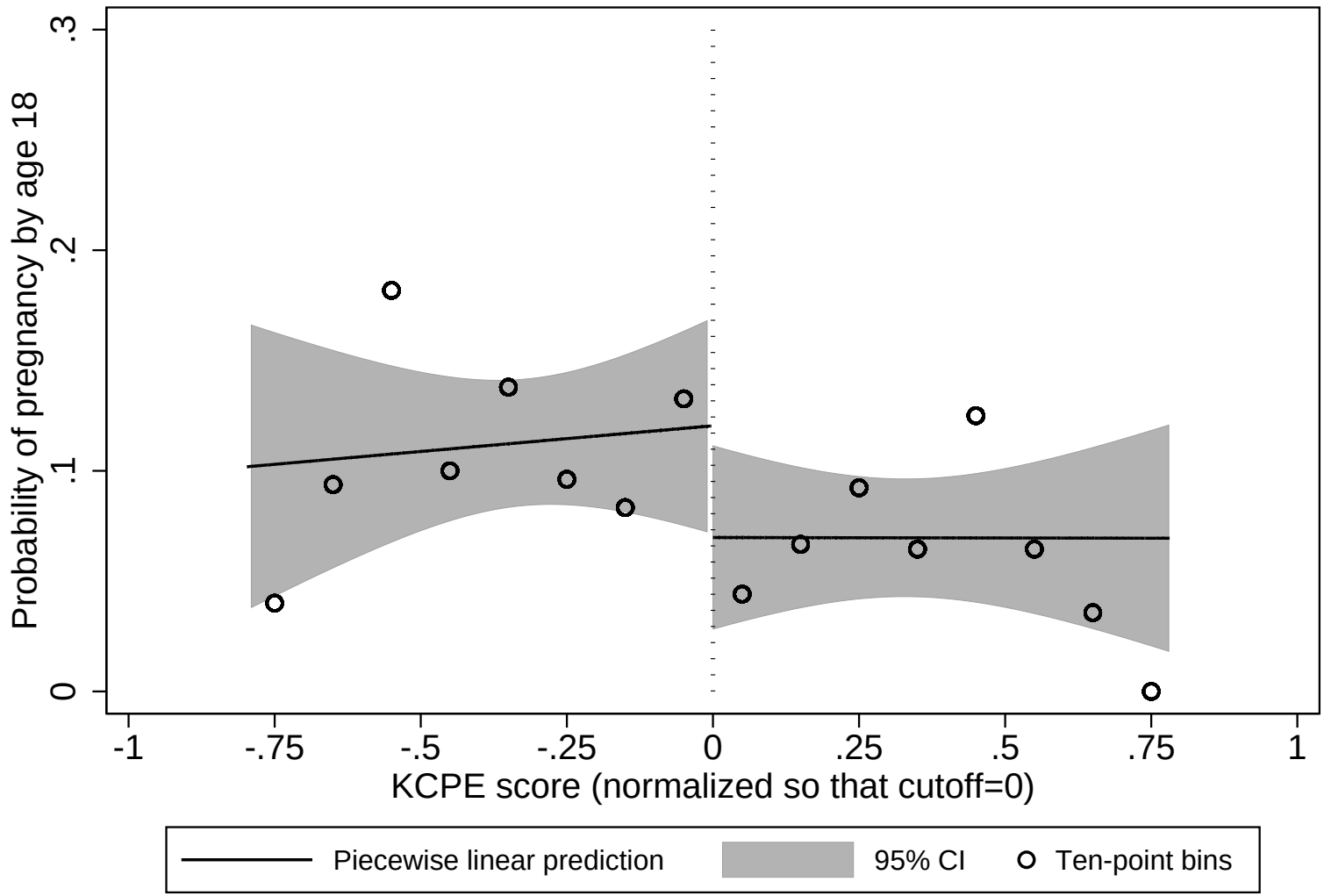
By Julian Cristia, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín*

This paper presents results from a large-scale randomized evaluation of the One Laptop per Child program, using data collected after 15 months of implementation in 318 primary schools in rural Peru. The program increased the ratio of computers per student from 0.12 to 1.18 in treatment schools. This expansion in access translated into substantial increases in use of computers both at school and at home. No evidence is found of effects on test scores in math and language. There is some evidence, though inconclusive, about positive effects on general cognitive skills. (*JEL* H52, I21, I24, I28, O15)

RCTs in Education, Part 3:

Evidence:
quasi-experiments vs experiments

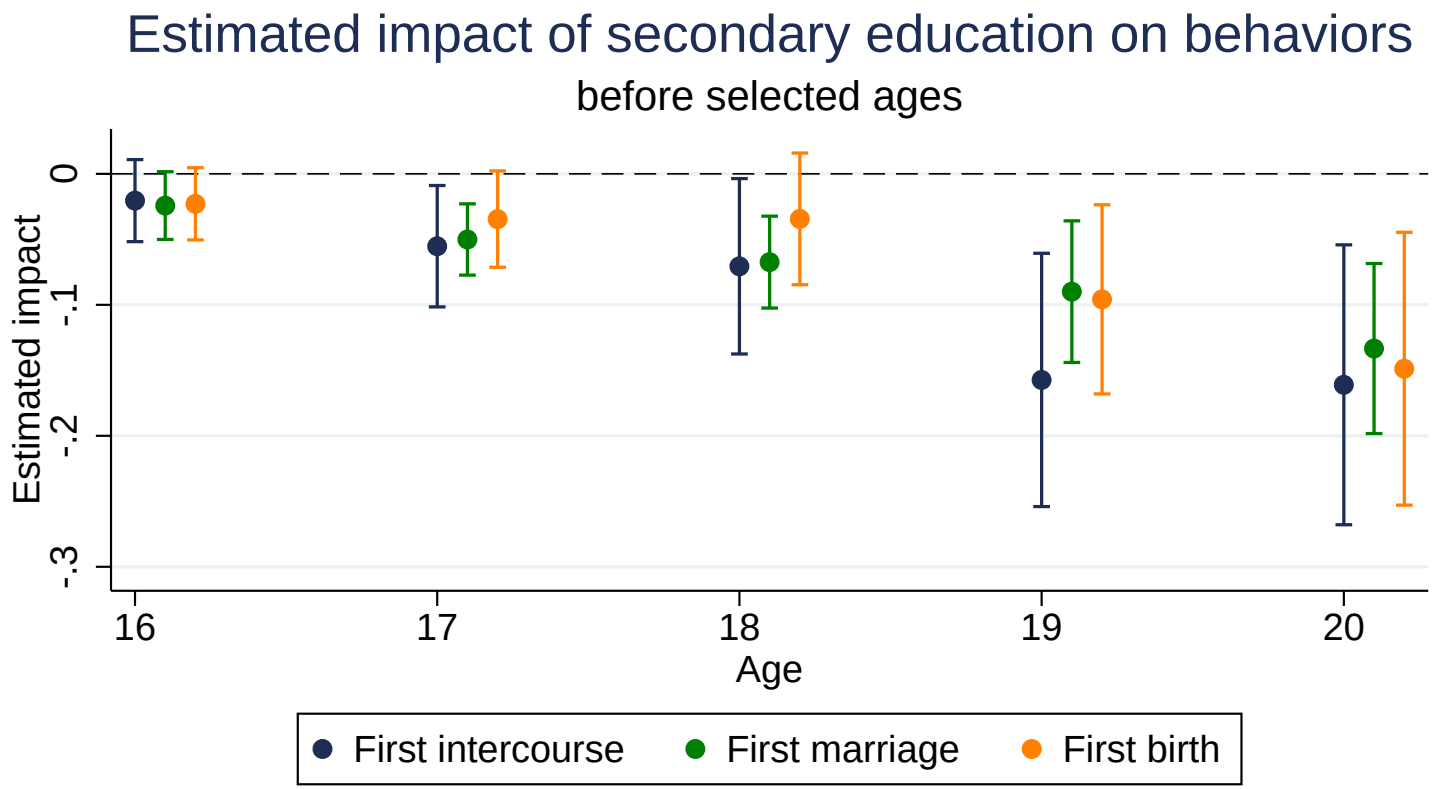
Ozier (2018) secondary school, Kenya, RD



KLPS2 data; dots are midpoints of 10-point bins

Brudevold-Newman: secondary school, Kenya, DD

Figure 8: Fertility behavior coefficients



Each point represents the coefficient on years of education from separate regressions where the dependent variables are binary indicators for whether individuals participated in each behavior before age X. Years of education is instrumented with cohort * county level exposure. The bars denote the corresponding 95% confidence intervals, with standard errors clustered by county. The F-statistics for first intercourse and first marriage are 10.46, 10.46, 10.46, 12.43, and 14.38 for age 16, 17, 18, 19, and 20, respectively. First birth F-statistics are 18.08, 18.08, 18.08, 22.76, and 13.34.

Duflo, Dupas, Kremer: secondary school, Ghana, RCT

Table 5: Marriage, Reproductive Health and Health Behaviors

	Combined		
	All	Female	Male
	(1)	(2)	(3)
Ever lived with partner(2016)			
Treatment effect	-0.063	-0.091	-0.027
Standard error	(0.020)***	(0.028)***	(0.028)
Comparison mean	0.241	0.344	0.134
p-value on equality of effect *(5)=(6)=(8)=(9): .426*		*(2)=(3): .106*	
Ever pregnant/had a pregnant partner (2016)			
Treatment effect	-0.071	-0.107	-0.023
Standard error	(0.024)***	(0.032)***	(0.031)
Comparison mean	0.403	0.582	0.213
p-value on equality of effect *(5)=(6)=(8)=(9): .317*		*(2)=(3): .061**	
Number of children ever had (2016)			
Treatment effect	-0.130	-0.217	-0.030
Standard error	(0.040)***	(0.054)***	(0.054)
Comparison mean	0.519	0.814	0.212
p-value on equality of effect *(5)=(6)=(8)=(9): .092**		*(2)=(3): .014***	

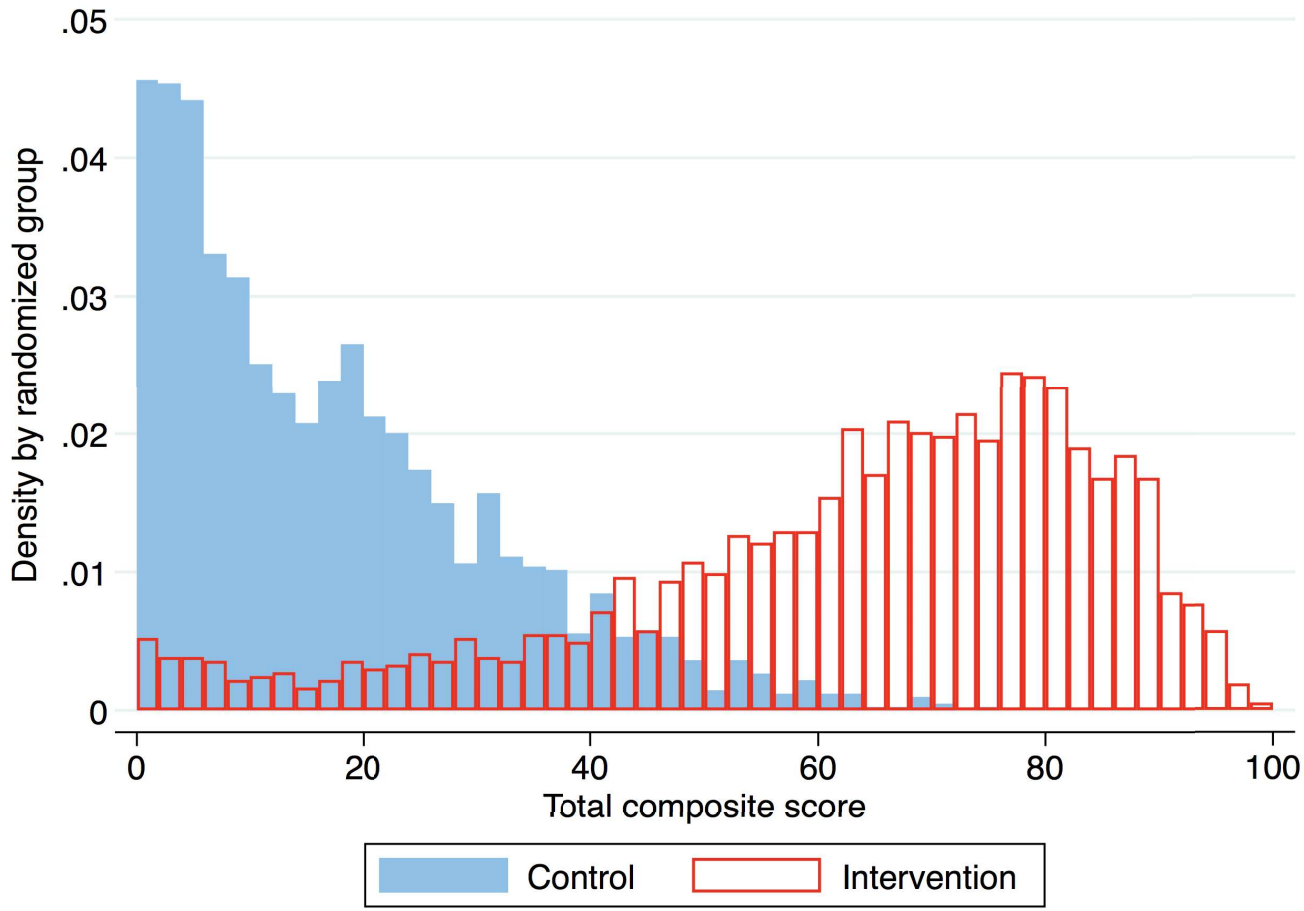
RCTs in Education, Part 4:

Evidence:
the big effects

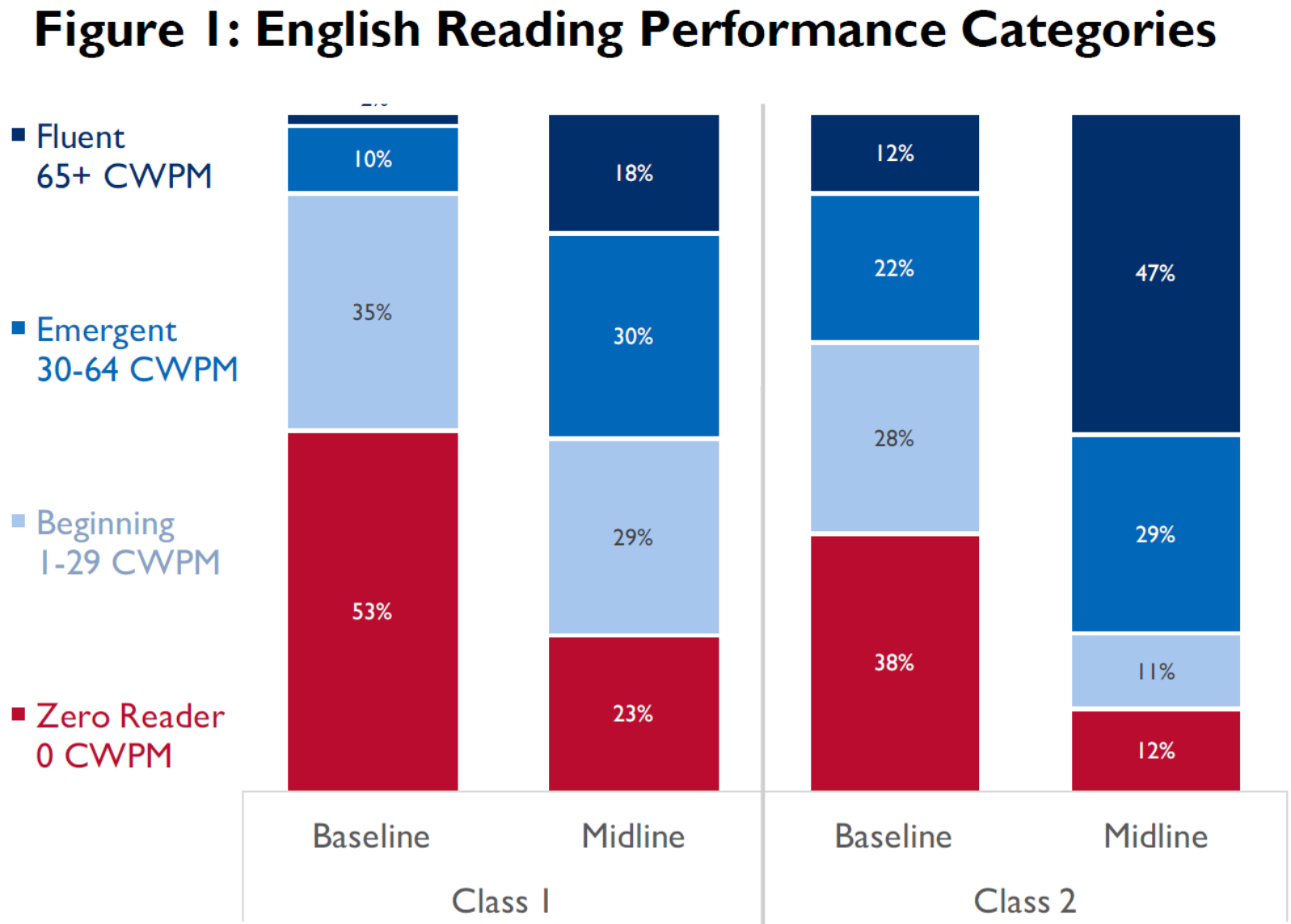
Eble et al 2019: $800, 3SD effect, The Gambia

Remember: MDE falls with root N. But if high dose yields high response:

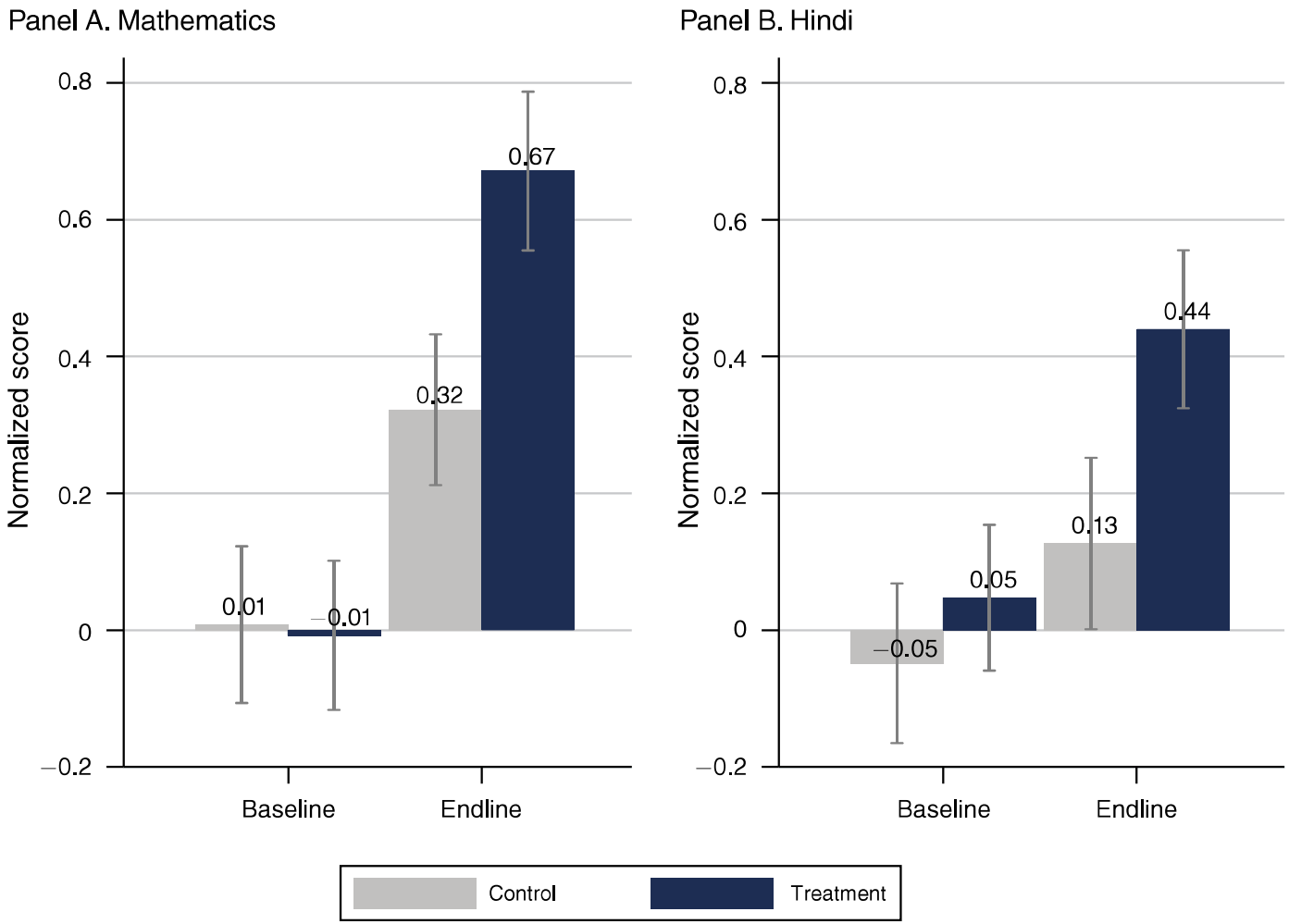
Primary outcome - composite test score



Piper et al (various years) - Kenya, nationwide



Muralidharan, Singh, Ganimian (2019 AER)



Muralidharan, Singh, Ganimian (2019 AER)

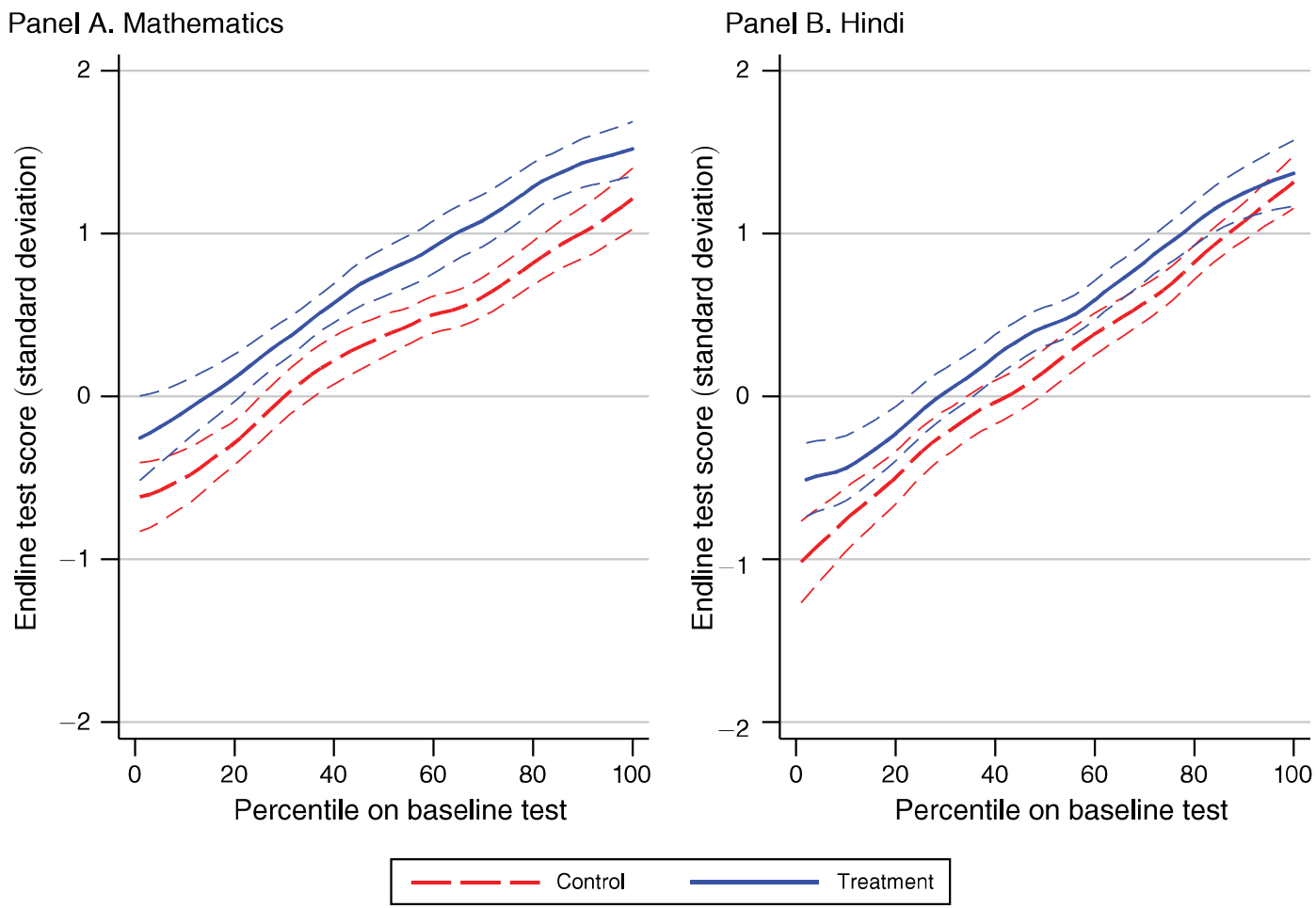


FIGURE 3. NONPARAMETRIC INVESTIGATION OF TREATMENT EFFECTS BY BASELINE PERCENTILES

RCTs in Education, Part 5:

Evidence: accountability*

Muralidharan, Sundararaman (2011): incentives

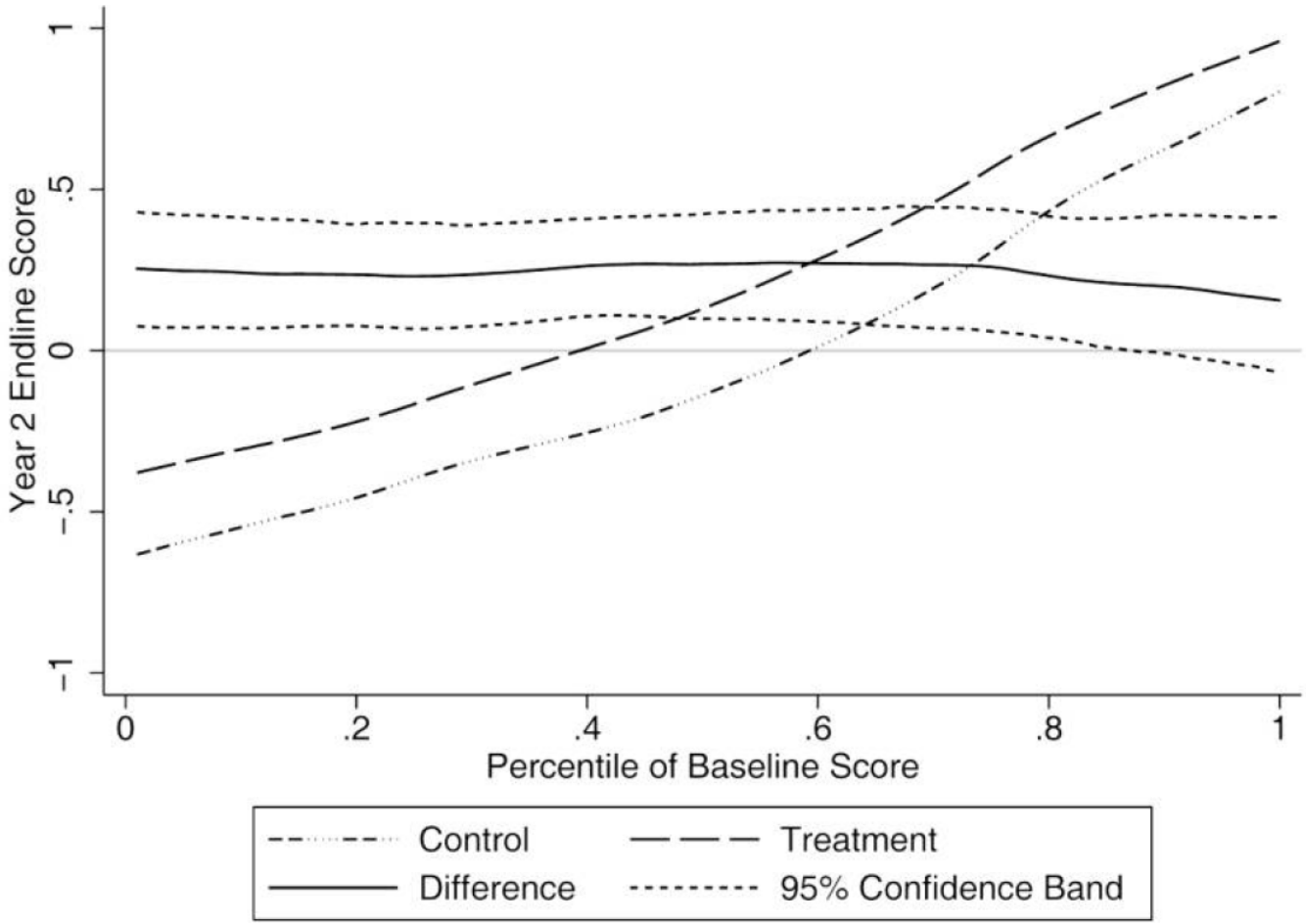


FIG. 3.—Nonparametric treatment effects by percentile of baseline score

Mbiti, Muralidharan, Romero, Schipper, Manda, Rajani

INPUTS, INCENTIVES, AND COMPLEMENTARITIES IN EDUCATION: EXPERIMENTAL EVIDENCE FROM TANZANIA*

ISAAC MBITI
KARTHIK MURALIDHARAN
MAURICIO ROMERO
YOUDI SCHIPPER
CONSTANTINE MANDA
RAKESH RAJANI

We present results from a large-scale randomized experiment across 350 schools in Tanzania that studied the impact of providing schools with (i) unconditional grants, (ii) teacher incentives based on student performance, and (iii) both of the above. After two years, we find (i) no impact on student test scores from providing school grants, (ii) some evidence of positive effects from teacher incentives, and (iii) significant positive effects from providing both programs.

Gilligan, Karachiwalla, Kasirye, Lucas, Neal

EDUCATOR INCENTIVES AND EDUCATIONAL TRIAGE IN RURAL PRIMARY SCHOOLS

Daniel O. Gilligan
Naureen Karachiwalla
Ibrahim Kasirye
Adrienne M. Lucas
Derek Neal

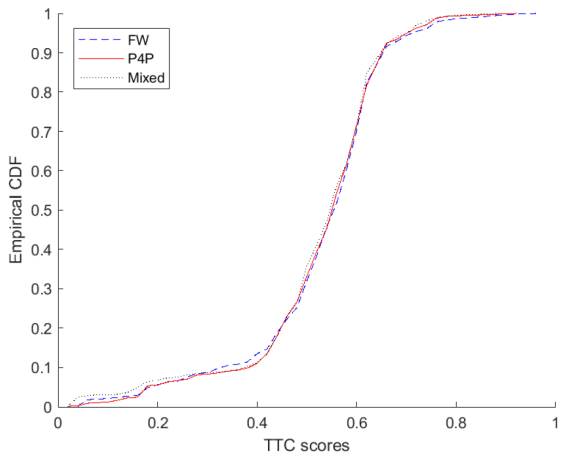
- Only 56% of the baseline students in our control schools were still attending school regularly.
- Our PFP treatment increased this attendance rate by four percentage points to 60%.
- Nonetheless, the test score results from our follow-up at the end of the P6 school year contain no statistically significant evidence that PFP treatment improved overall math scores.

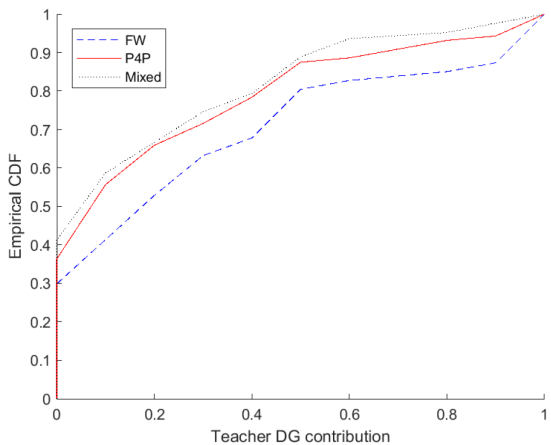
Although these results appear at odds with our strong attendance results, we find that when we separately examine the roughly 50% of our schools that provide math textbooks for P6 students, PFP raises maths achievement on items at or near the P6 grade level by more than 0.1 standard deviations. Further, the overall four percentage point increase in attendance in our second follow-up data is driven almost entirely by a seven percentage point increase in attendance among students in schools that provide math books for P6 students.

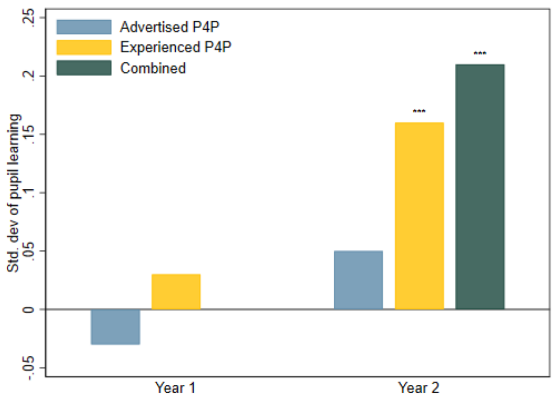
Leaver, Ozier, Serneels, Zeitlin

Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary schools

Clare Leaver*, Owen Ozier†, Pieter Serneels‡ and Andrew Zeitlin§



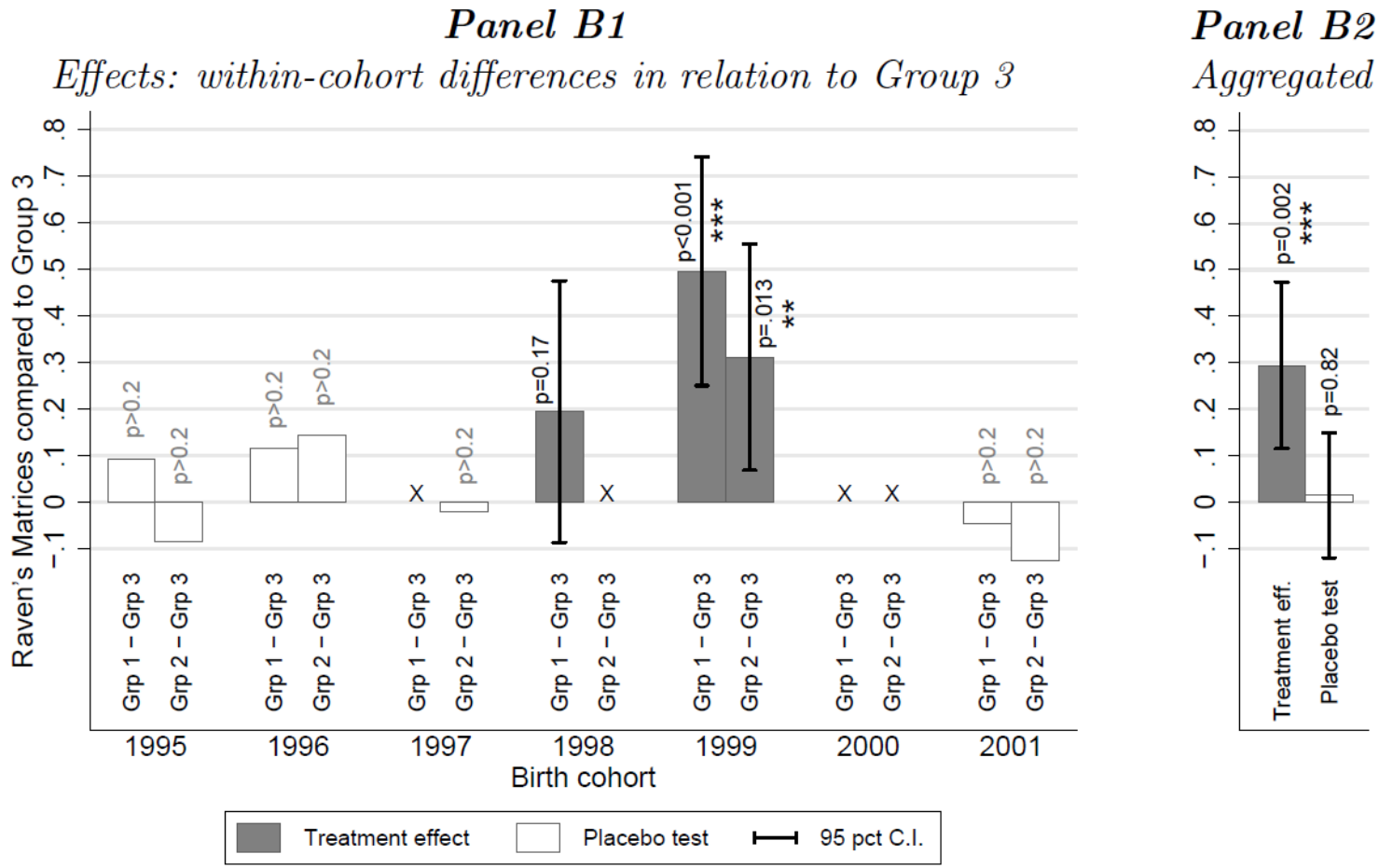




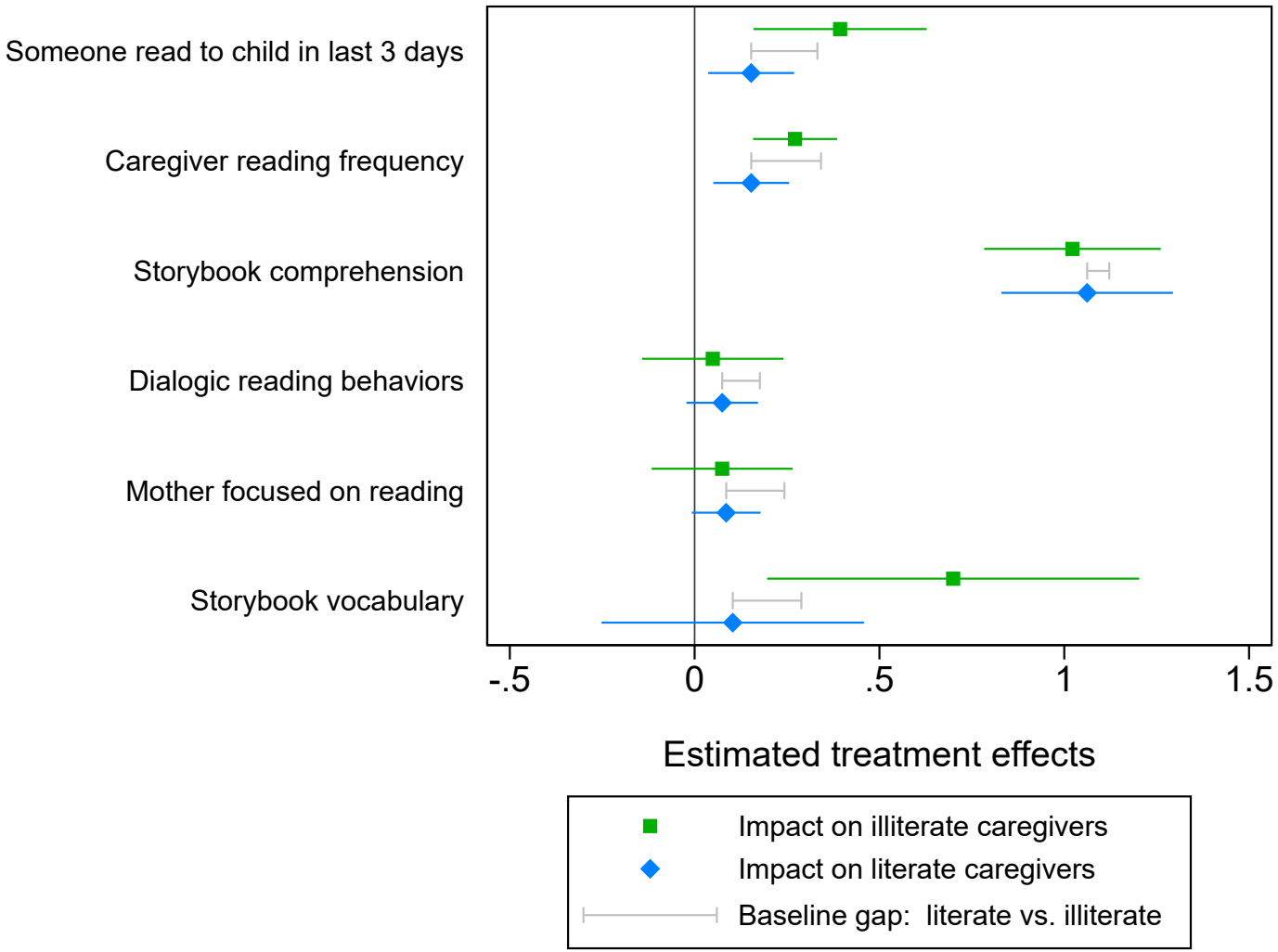
RCTs in Education, Part 6:

Evidence:
intervening early

Ozier (2018): early health



Knauer, Jakiela, Ozier, Aboud, Fernald: early books



Much more

Less covered by these slides and/or published research:

- How to improve teacher training
- The role for the private sector and markets
- Long-term (labor market?) impacts (Lewis, Sen)
- Other early childhood interventions, good and bad

Thanks!