The Lost Human Capital

Teacher Knowledge and Student Achievement in Africa

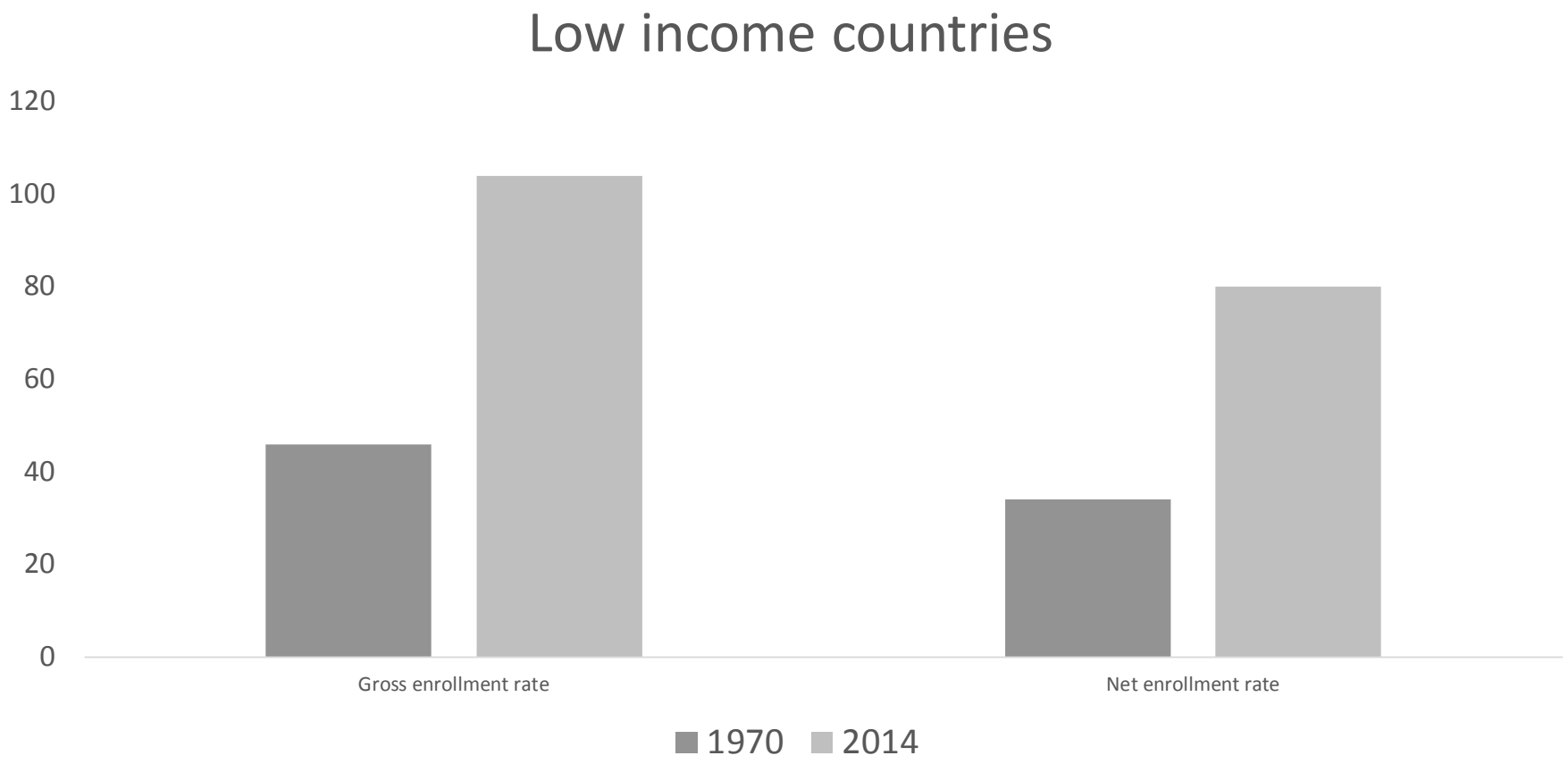
Tessa Bold
Institute for International Economic Studies (IIES), Stockholm University

Deon Filmer
World Bank
Ezequiel Molina
World Bank
Jakob Svensson
Institute for International Economic Studies (IIES), Stockholm University

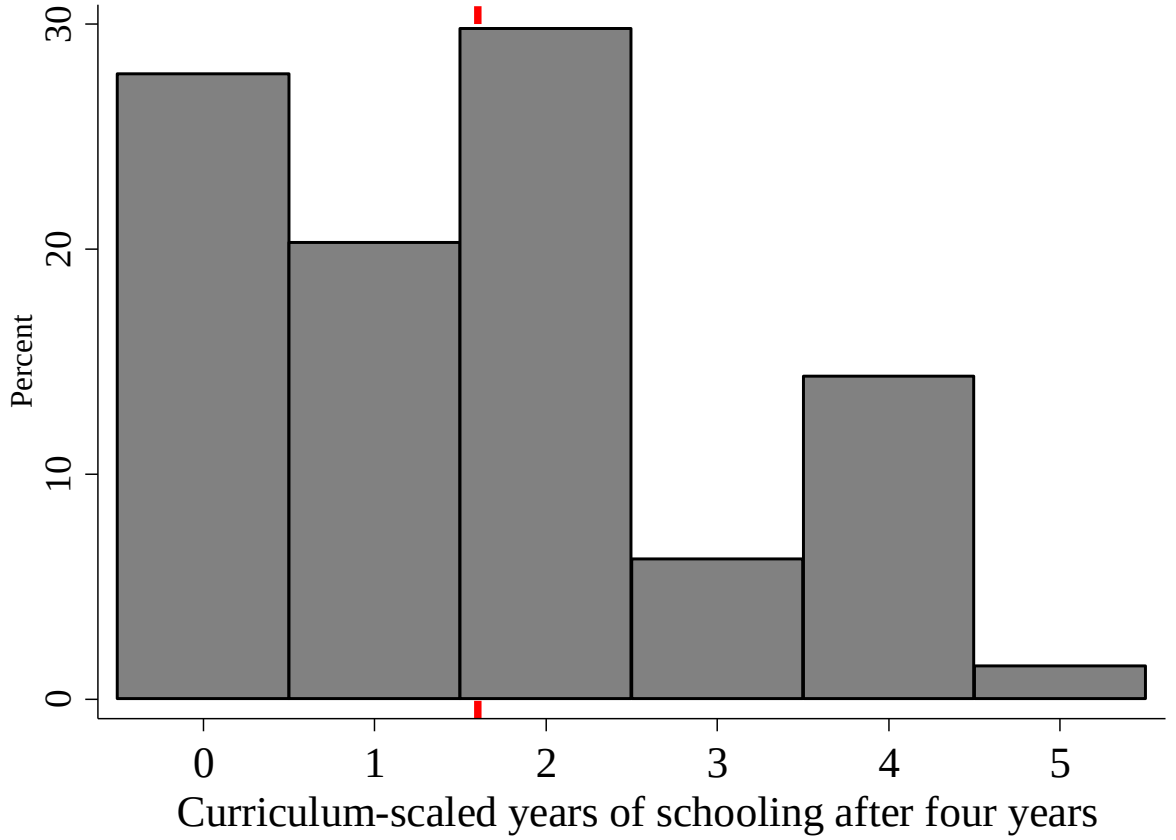
What do we know about education in LIC?

- Two facts

Large increase in enrollment (primary school)



but without much learning



but without much learning: Examples

- After more than 3 years of compulsory language teaching, 4 out of 5 students in Mozambique and Nigeria cannot read simple words of Portuguese and English, respectively.
- Only 1/4 of Indian students in grade 4 can manage tasks—such as basic subtraction—that are part of the curriculum for the second grade.
- Roughly half of the students in Uganda, after 3 years of mathematics teaching, cannot place numbers between 0-999 in order.
- “The global learning crisis” (UNESCO, 2013).
- **Here: put the spotlight on teachers (schools), realizing the problem of enrollment without learning is systemic**

What do we know and not know?

- A growing body of evidence—both from the teacher VA and the experimental literature—shows that teacher quality is a key determinant of student learning.
- Less is known about what specific dimensions of teacher quality matter
- Even less about how teachers perform along these dimensions
- And the extent to which this matters for student learning

Presentation today

- Discusses an ongoing research program intended to help fill this evidence/knowledge gap.
- Two parts:

  (1) Stylized facts (Bold, Svensson, et al, 2017a)

  (2) Structural estimates of the importance of teacher content knowledge (Bold, Filmer, Molina, Svensson, 2017b)

Presentation today: Part 1

Data collected through direct observations, unannounced visits, and tests, from primary schools in seven SSA countries (representing close to 40% of the region's total population):

(1) How much do teachers teach?

(2) What do teachers know?

(3) How well do teachers teach?

Presentation today: Part 2

- Present a statistical model of cumulative knowledge acquisition, accounting for imperfect persistence in learning btw grades
- Exploit the fact that we have test score data for both the current and previous year's teacher to identify two structural parameters: **contemporaneous effect** and **extent of fade out** of the teachers' impact in earlier grades
- Exploit within-student within-teacher variation to identify these effects empirically
- Combine structural parameters to estimate the **cumulative effect** of teacher knowledge on student achievement and calculate the learning achievements in a series of **counterfactual experiments**

Presentation today: Part 2

- We provide a complementary approach (less data heavy) to estimate the importance of teachers compared to that used in the value-added literature
- On basis of cross-sectional student data and one round of historic data on inputs, can estimate a reasonably tight bound on the cumulative effect of teacher content knowledge on student learning half way through primary school.

Context

- Rapidly expanding but weakly governed, primary education sector in SSA.
- Most of the increase in enrollment has taken place in the public sector, which remains the dominant actor in the sector (80%).
- Huge increase in the number of teachers, which has risen from 500,000 primary school teachers in 1970 to almost 2.8 million in 2009.
- The salaries of these teachers make up more than 70 % of the expenditure in education and approximately 12 % of total government expenditure in SSA.

Data: The Service Delivery Indicators program

- SDI program has collected data from a total of 7 countries: Kenya (2012), Mozambique (2014), Nigeria (2013), Senegal (2010), Tanzania (2010, 2014), Togo (2013), and Uganda (2013).
- Primary schools with a 4^{th}-grade class formed the sampling frame.
- Samples designed to provide representative estimates for teacher effort, knowledge, and skills in public primary schools
- Representative data were also collected for private primary schools.
- The surveys collected a broad set of school, teacher, and student specific information, with an approach that relies as much as possible on direct observation, rather than on respondent reports.
- In total data on 2,600 schools, over 21,000 teachers and 24,000 students

How much do teachers teach?

How Much Time Do Teachers Teach?

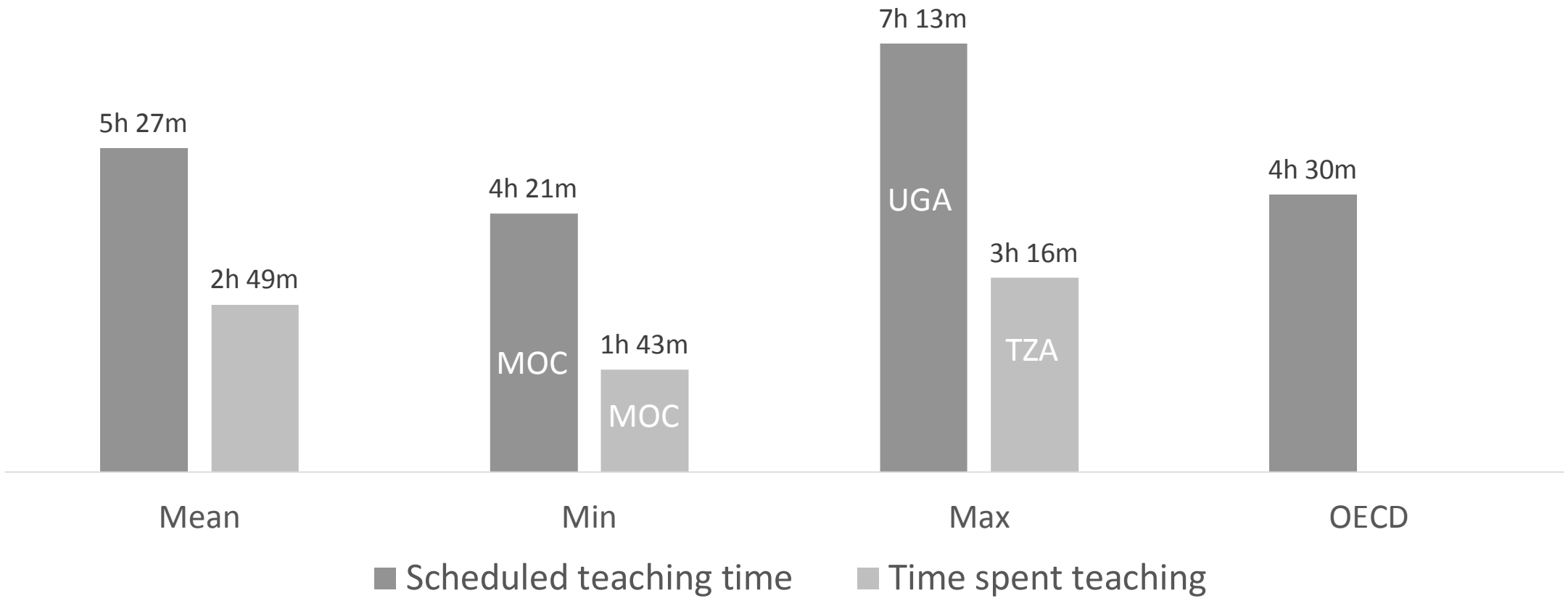
- Measurement: an extended approach of Chaudhury et al. (2006)
- In each school, during a first announced visit, up to 10 teachers were randomly selected from the teacher roster.
- An unannounced visit followed to identify whether the selected teachers were in the school (**absence from school**), and if so, if they were in class teaching (**absence from class**).
- Combine data on absenteeism from class, data on scheduled time of a teaching day, data based on recording a minute-by-minute snapshot of what teachers were doing when in class ("lesson lost to non-teaching activities") to estimate **instruction time as experienced by students**.

How Much Time Do Teachers Teach?

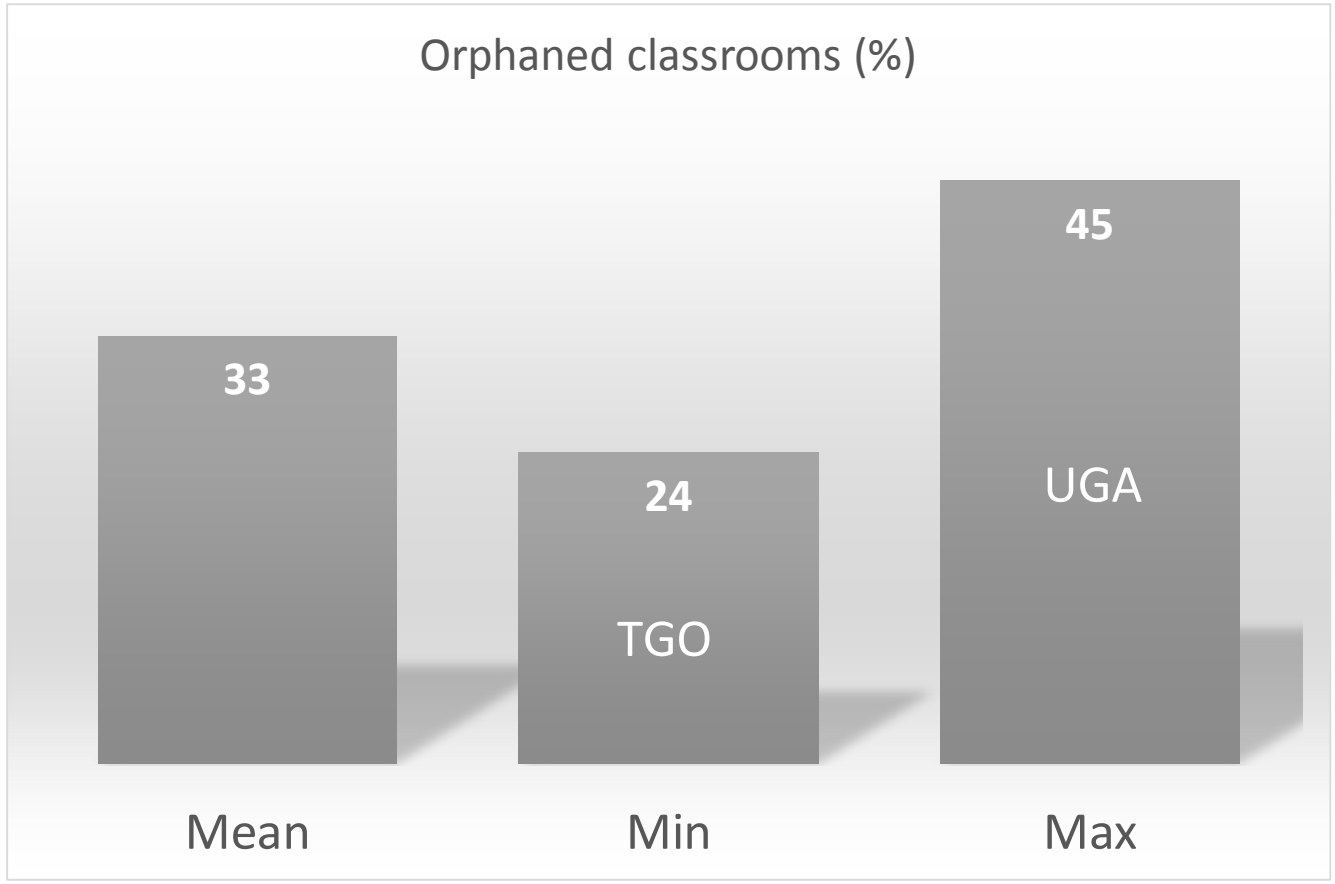
	All	Min	Max
Absence from class (%)	44	23 (NGA)	57 (UGA)
Absence from school (%)	23	15 (KEN)	45 (MOZ)
No. of teachers	16,543		

- Rank correlation coefficient=0.5 at the country level

How Much Time Do Teachers Teach?



How Much Time Do Teachers Teach?



What do teachers know?

What Do Teachers Know?

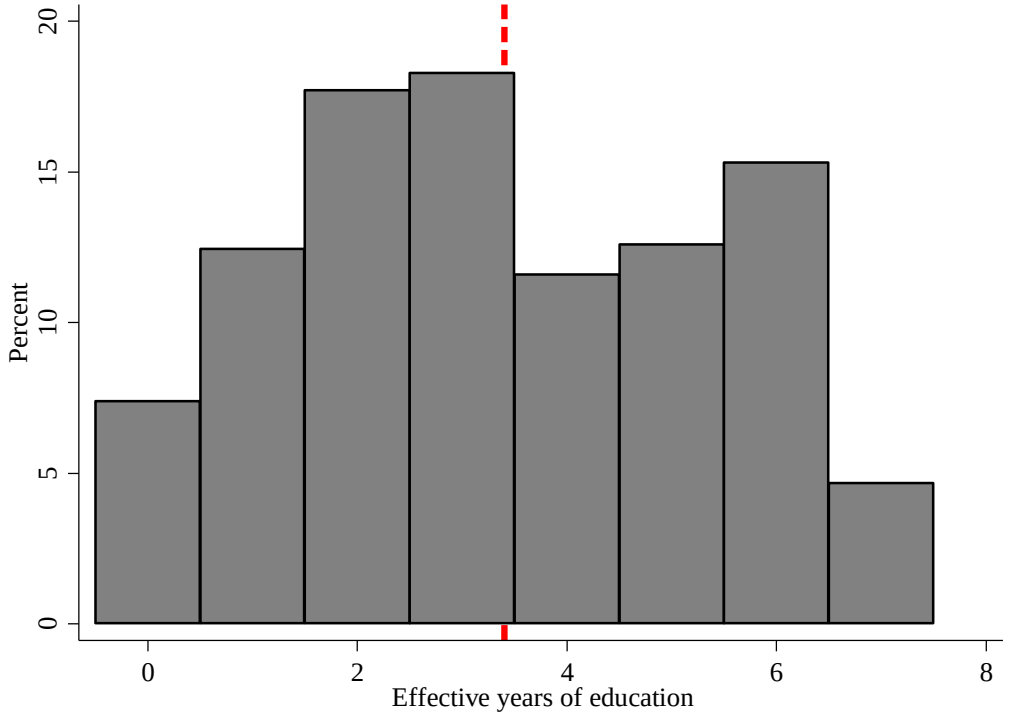
- Measurement: All language and maths teachers teaching Grade 4 in the current year (or Grade 3 in the previous year) were assessed
- Teachers here were asked to mark (or “grade”) mock student tests in language and in mathematics.
- Method of assessment has two potential advantages:
  - aims to assess teachers in a way that is consistent with their normal activities.
  - recognizes teachers as professionals.
- All questions on the teacher test were based on common items taken from the primary curricula of each country, covering material from Grade 1 level up to upper primary level.

What Do Teachers Know? Outcome measures

- Use the raw scores to determine the grade level of proficiency of teachers and label it as effective years of education acquired
- Advantage of the transformation:
  - test covered test items from grade 1 up to the end of primary school and since the samples by construction are nationally representative.
  - transformed data points are informative in themselves.
  - compare students and teachers using the same scale.
  - extrapolate beyond effective years of schooling observed in the sample in a meaningful way.

More on the test

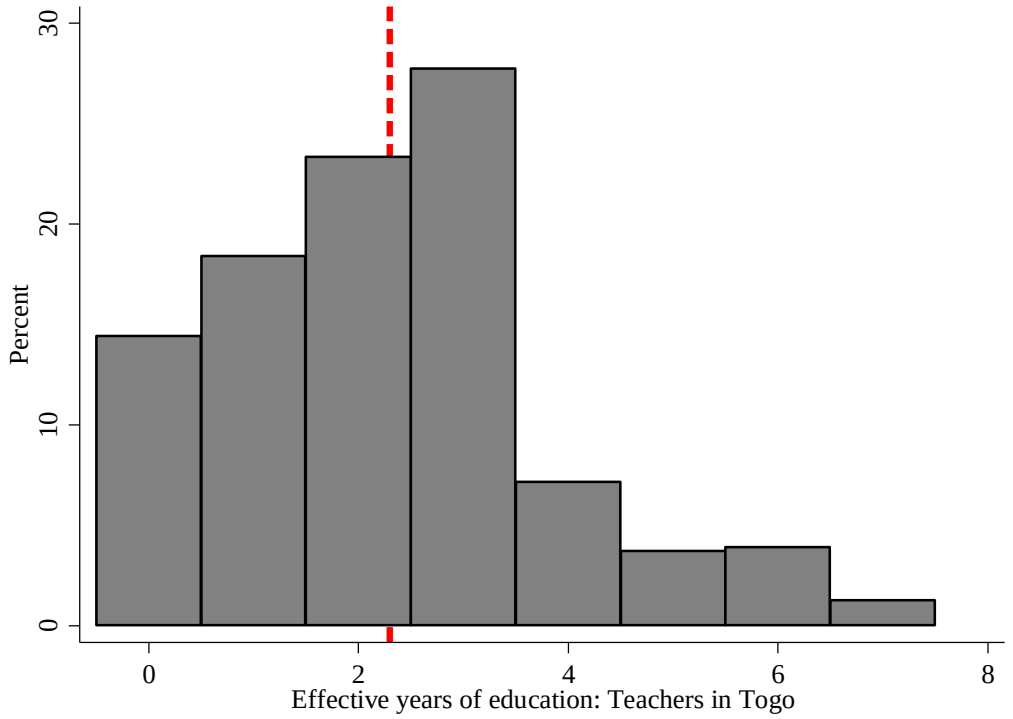
Effective years of education of teachers



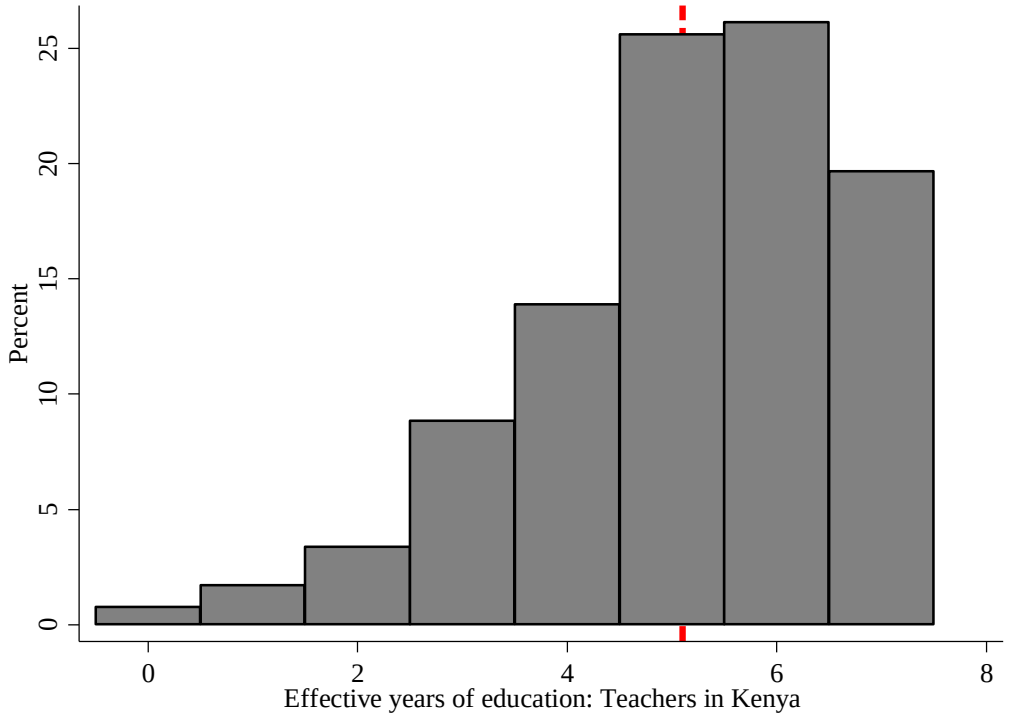
Alternative definition

What do teachers know: Human capital

Togo



Kenya



What do teachers know: Examples

Math	All	Min	Max
Can add double digits (%)	91	75 [TGO]	98 [KEN]
Can subtract double digits (%)	77	59 [NGA]	94 [TZA I]
Can multiply double digits (%)	68	44 [MOZ]	89 [SEN]
Can interpret data in a graph (%)	25	12 [TGO]	62 [KEN]
Can solve algebra (%)	35	3 [MOZ]	74 [KEN]
Can solve math story problem (%)	15	7 [SEN]	22 [TZA I]
No. of teachers	3957		

How well do teachers teach?

How Well Do Teachers Teach?

- Good teaching also requires that teachers know how to translate their subject knowledge into effective pedagogy and then apply this in the classroom.
- Teachers must also know how to assess student capabilities and react appropriate.
- Three outcome measures
  - Pedagogical knowledge
  - Assessing students
  - Skills and practices in the classroom

How Well Do Teachers Teach?

Pedagogy	All	Min	Max
Preparing a lesson, %	33	15 (NGA)	58 (TZA)
Formulate questions to check understanding, %	23	5 (NGA)	55 (KEN)
Formulate questions to apply to other contexts, %	7	3 (NGA)	15 (TZA)
Assessing students' abilities, %	19	8 (NGA)	39 (KEN)
Evaluating students' progress, %	12	5 (NGA)	26 (KEN)
Ask a mix of lower and higher order questions, %	31	14 (MOZ)	44 (UGA)
No. of teachers	5,181		

Sum up: What do teachers do and know?

- If “adequate” teaching is characterized as being taught by teachers with at least basic pedagogical knowledge and minimum subject knowledge in language and mathematics for almost the full scheduled teaching day,

  then essentially no public primary school students in these countries are offer adequate quality education.

Selection, training,
emuneration of tea

Does it matter?

Assess the effect of teacher knowledge on student learning

- Present a statistical model of cumulative knowledge acquisition, accounting for imperfect persistence in learning btw grades
- Exploit the fact that we have test score data for both the current and previous year's teacher to identify two structural parameters: **contemporaneous effect** and **extent of fade out** of the teachers' impact in earlier grades.
- Exploit within-student within-teacher variation to identify these effects empirically
- Combine structural parameters to estimate the **cumulative effect** of teacher knowledge on student achievement and calculate the learning achievements in a series of **counterfactual experiments**

Statistical model for cognitive achievement

General education production function

(1) $$y_{ijt,k} = F\left[x_{ij,k}(t), . , . , . , \varepsilon_{ijt,k}\right]$$

where $y_{ijt,k}$ is student i's achievement in school j after t years of schooling (or in grade t) in subject k.

$x_{ijt,k}(t)$ teacher subject content knowledge (assume: $\text{cov}(x, \varepsilon) = 0$)

Linearizing (1) and taking first difference (across subjects), assuming that the effects of teacher content knowledge (properly normalized) on student learning are subject-invariant

A simple cumulative model of student learning

(2) $$\Delta y_4 = \beta_0 + \beta_4 \Delta x_4 + \beta_3 \Delta x_3 + \beta_2 \Delta x_2 + \beta_1 \Delta x_1 + \Delta \varepsilon_4$$

Three core assumptions (as in VA literature, Todd & Wolpin (2003))

1. The stationarity of the test score distribution over time: $\text{var}(\Delta x_t) = \text{var}(\Delta x_{t'}) = \text{var}(\Delta x)$
2. The contemporaneous effect (α_t) of teacher knowledge is independent of the age at which it is applied $\Longrightarrow \alpha_t = \alpha$
3. The effect of teacher content knowledge declines (geometrically) with distance $\Longrightarrow$ if α is the effect on y_4 of grade 4 teacher, then $\gamma\alpha$ is the effect on y_4 of gr. 3 teacher, where γ "degree of persistence"

A simple cumulative model of student learning

Rewrite the production function as

(2') $$\Delta y = \alpha_0 + \alpha\Delta x_4 + \alpha\gamma\Delta x_3 + \alpha\gamma^2\Delta x_2 + \alpha\gamma^3\Delta x_1 + \epsilon$$

Two structural parameters: α, γ and CE = $\alpha \sum_{t=1}^{4} \gamma^{4-t}$

Model we can estimate

(3) $$\Delta y = \beta_0 + \beta_4\Delta x_4 + \beta_3\Delta x_3 + \mu$$

Even assuming $\text{cov}(\Delta x_t, \epsilon) = 0$, we cannot recover the structural coefficients directly from the reduced form coefficients β_4 and β_3

OLS estimator of β_4 and β_3 in (4)

(4) $$\text{plim}\,\hat{\beta}_4 = \alpha + \alpha\gamma^2\left(\frac{\rho_{24}-\rho_{23}\rho}{1-\rho^2}\right) + \alpha\gamma^3\left(\frac{\rho_{14}-\rho_{13}\rho}{1-\rho^2}\right)$$

(5) $$\text{plim}\,\hat{\beta}_3 = \alpha\gamma + \alpha\gamma^2\left(\frac{\rho_{23}-\rho_{24}\rho}{1-\rho^2}\right) + \alpha\gamma^3\left(\frac{\rho_{13}-\rho_{14}\rho}{1-\rho^2}\right)$$

(7) $$\Delta x_4 = \rho_0 + \rho_{43}\Delta x_3 + v_{4,3}$$

Allow the correlation coefficients, $\rho_{t,t'}$, to vary freely in a mildly restricted space and estimate the full distribution of possible effects (α, γ and $\alpha\sum_{t=1}^{4}\gamma^{4-t}$)

OLS estimator of β_4 and β_3 in (4)

(4) $$\text{plim}\,\hat{\beta}_4 = \alpha + \alpha\gamma^2\left(\frac{\rho_{24}-\rho_{23}\rho}{1-\rho^2}\right) + \alpha\gamma^3\left(\frac{\rho_{14}-\rho_{13}\rho}{1-\rho^2}\right)$$

(5) $$\text{plim}\,\hat{\beta}_3 = \alpha\gamma + \alpha\gamma^2\left(\frac{\rho_{23}-\rho_{24}\rho}{1-\rho^2}\right) + \alpha\gamma^3\left(\frac{\rho_{13}-\rho_{14}\rho}{1-\rho^2}\right)$$

(7) $$\Delta x_4 = \rho_0 + \rho_{43}\Delta x_3 + v_{4,3}$$

Make three restrictions on $\rho_{t,t'}$ motivated by the typical pattern of transitions of teachers and their grades through primary school coupled with the assumption that these transition patterns are the main drivers of correlations across grades.

Estimating structural parameters

- Assume:
  - $\rho_{t,t'} \geq 0$
  - $\rho_{t,t'}$ is decreasing in $|t - t'|$
  - $\rho_{t,t-1}$ is decreasing in t ($\Longrightarrow \rho_{t,t-1} \geq \rho$ for $t < 4$)
- Given a specific parameterization of the $\rho_{t,t'}$, we are left with 3 unknowns (α, γ, and ρ) and 3 equations

Identification: Approach

1. Linearizing the production function (1) and express as difference across subjects

   $\Longrightarrow$ All subject-invariant unobserved heterogeneity at the student, school and parent level is removed

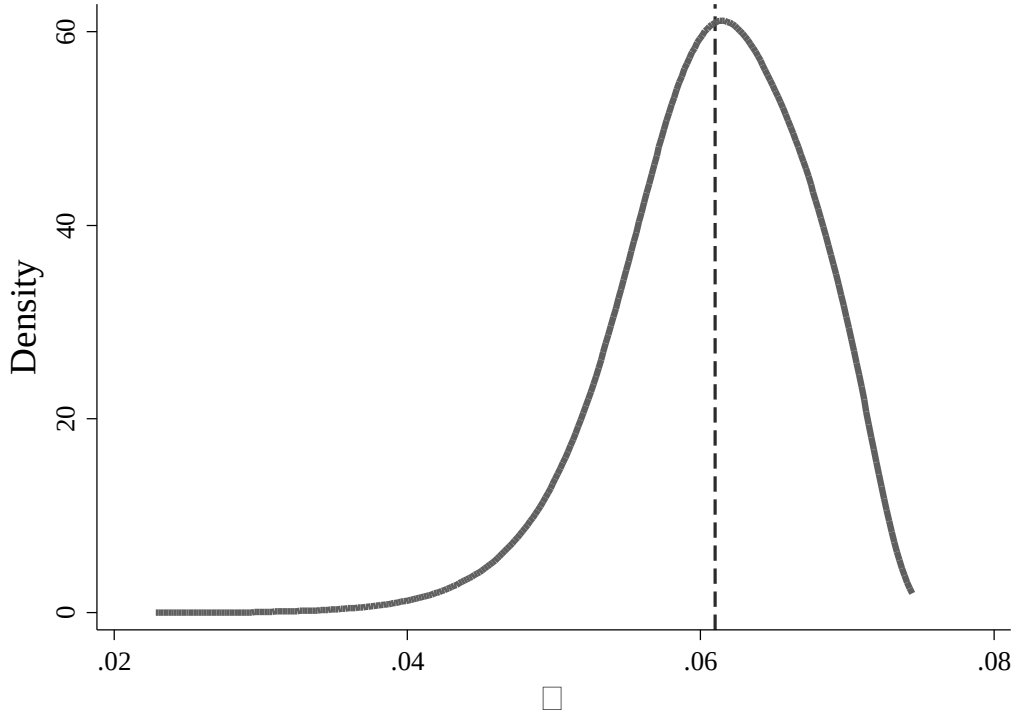
2. Restrict attention to students who were taught by class teachers in gr. 3 and 4

   $\Longrightarrow$ Any teacher-specific, subject-invariant heterogeneity in gr. 4 and 3 removed

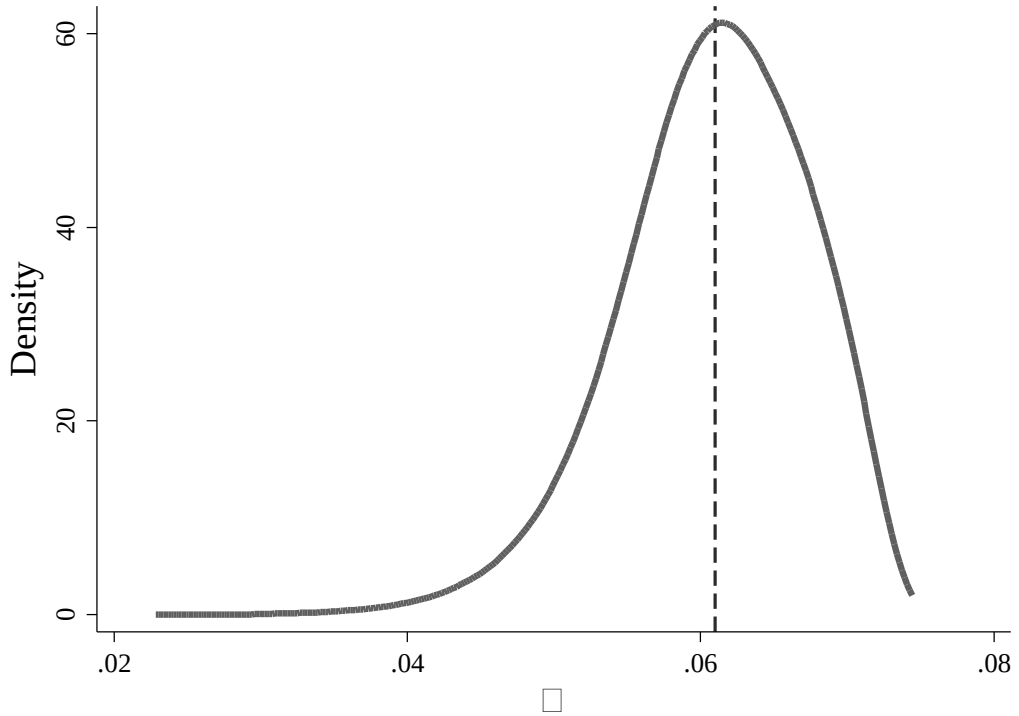
Parameters are identify if students, parents, teachers, schools do not sort/respond to subject-differences in teacher content knowledge.

Findings

Probability density functions of the estimated α

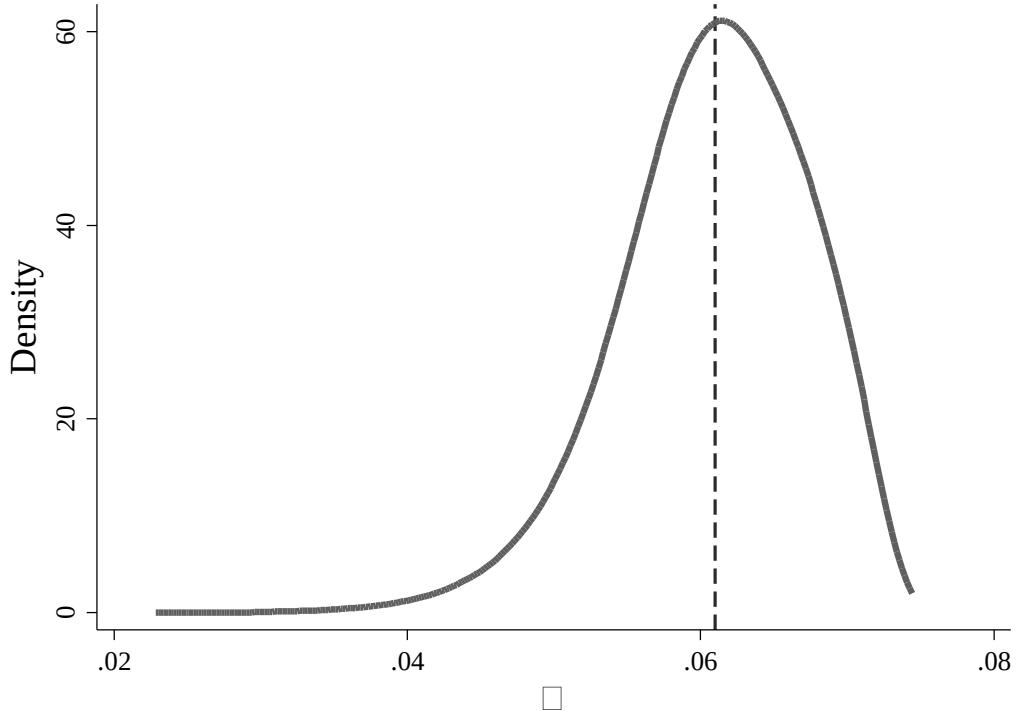


Probability density functions of the estimated α



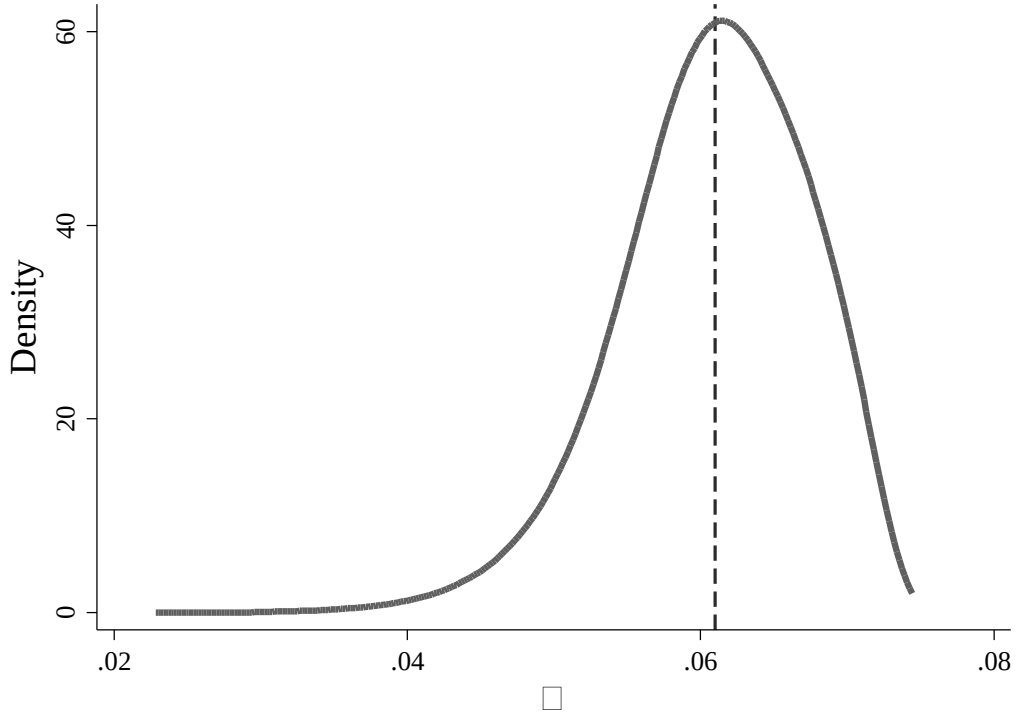
- At α^{med}: Teacher with 1 more year of effective education $\Longrightarrow$ increase student learning by ¾ of a month after one year

Probability density functions of the estimated α



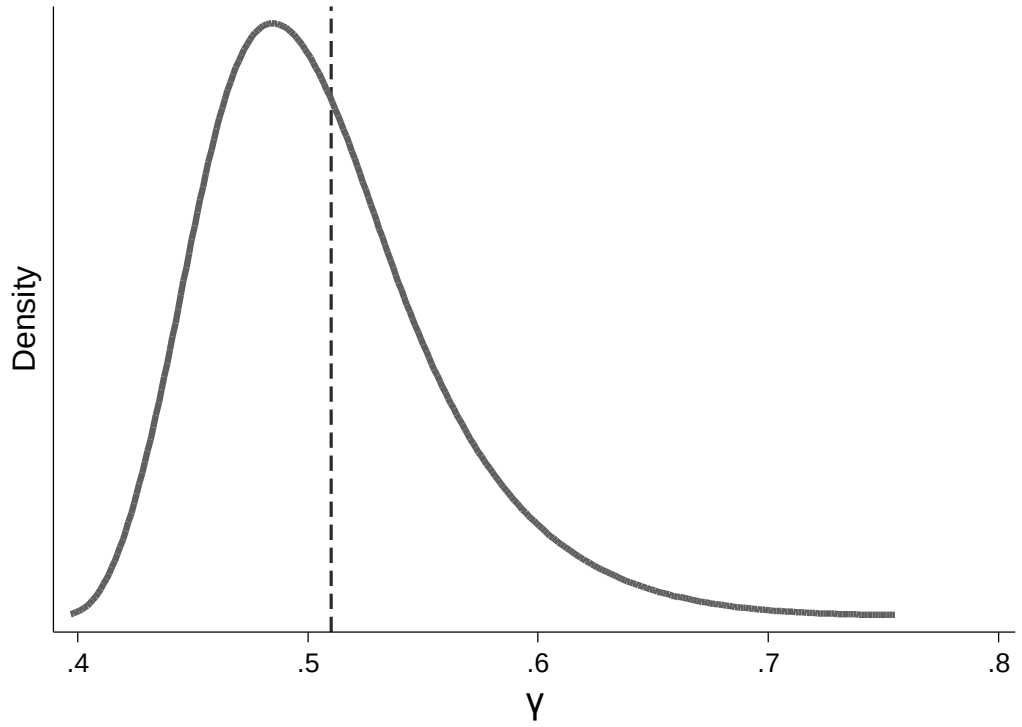
- A 1 SD increase in effective years of education for a teacher increases student learning by 0.09 SD

Probability density functions of the estimated α

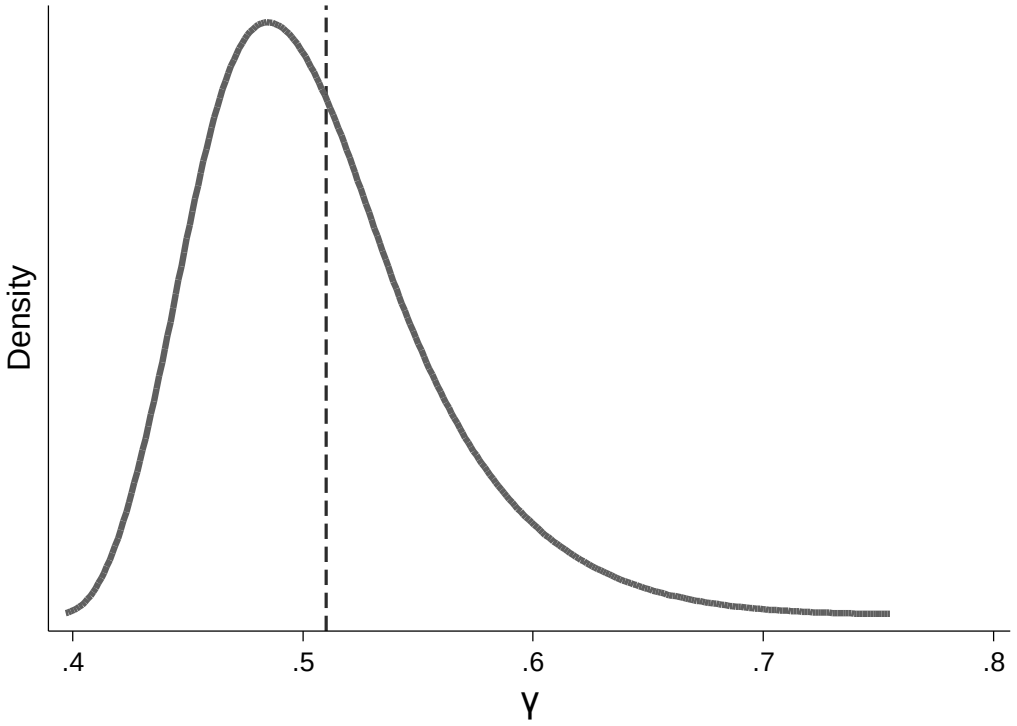


- Effect size = 0.09 SD
- Cf with VA literature: Effect sizes ranging from 0.1-0.2SD (Rockoff, 2004; Rivkin et al., 2005; Aaronson et al., 2007; Chetty et al., 2014; Araujo et al., 2016; and Bau and Das, 2017).
- 1 SD increase in teacher test scores raise student test scores by 0.07 SD (Bau and Das, 2017)

Probability density functions of the estimated γ

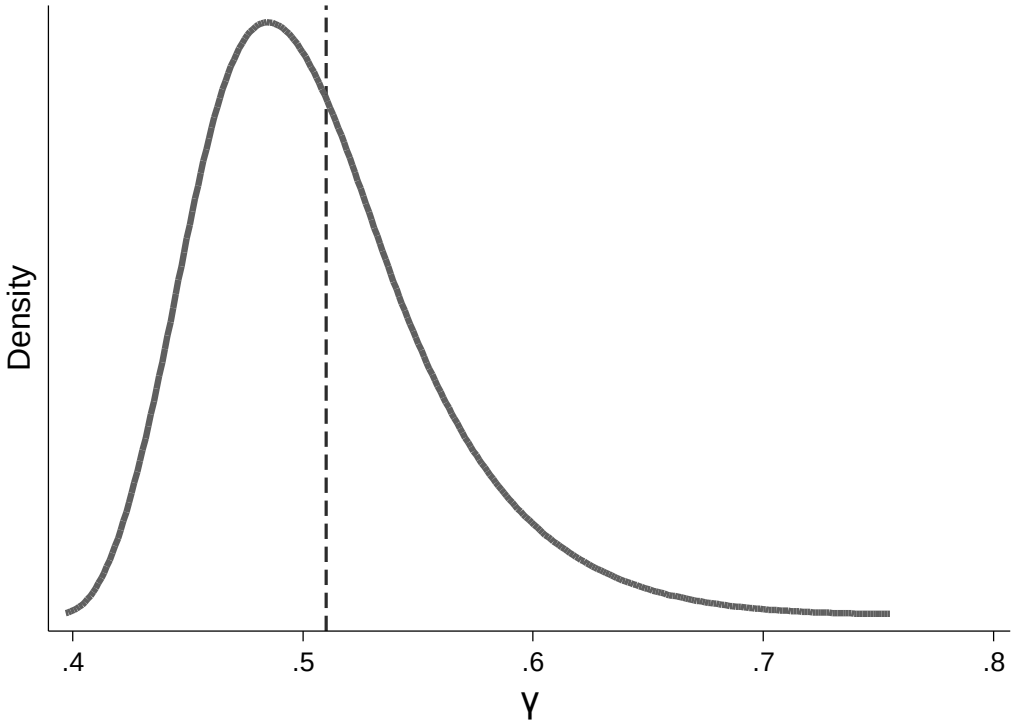


Probability density functions of the estimated γ



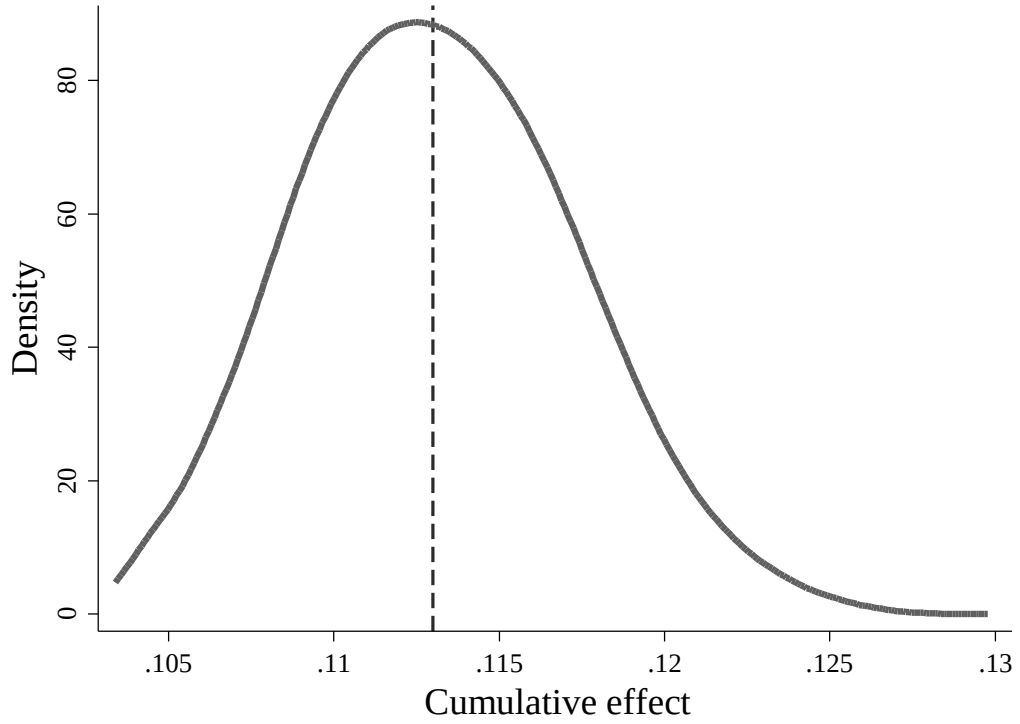
- Approx. 50% of the short-run effect persists between grades

Probability density functions of the estimated γ



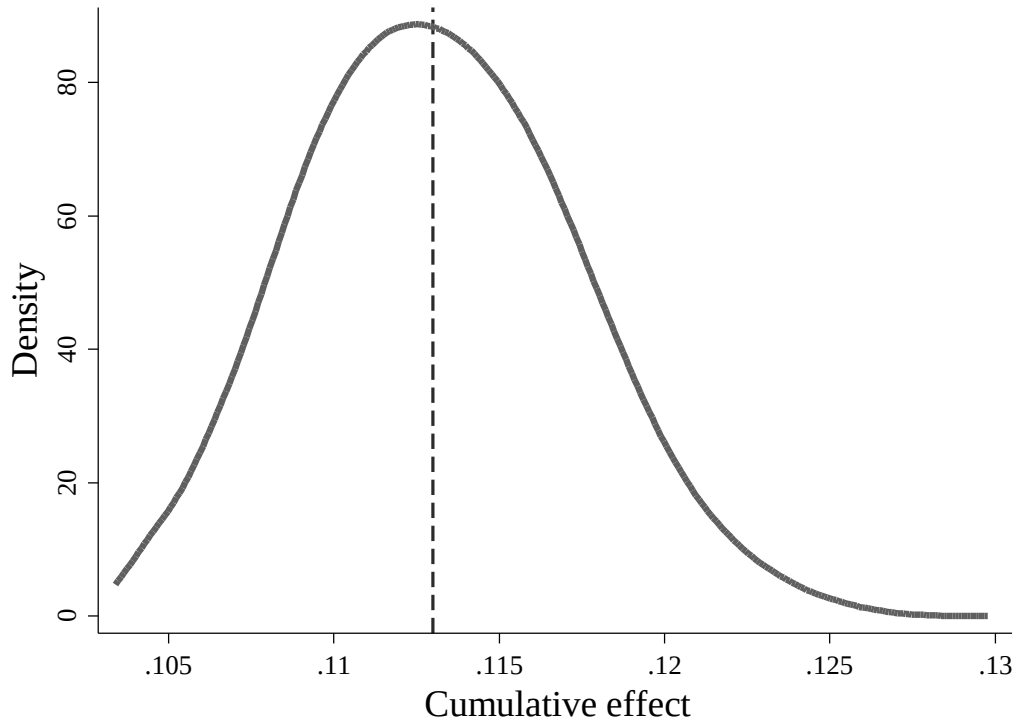
- Approx. 50% of the short-run effect persists between grades
- Consistent with what has been reported using data from Pakistan and the US (Kane and Staiger 2008; Jacob, Lefgren, and Sims 2010; Rothstein 2010; and Andrabi et al., 2011

Probability density functions of the estimated CE



$$CE = \alpha \sum_{t=1}^{4} \gamma^{4-t}$$

Probability density functions of the estimated CE



- Being taught, throughout lower primary, by a teacher with 1 more year of effective education would increase student learning by almost a month and a half after four years.

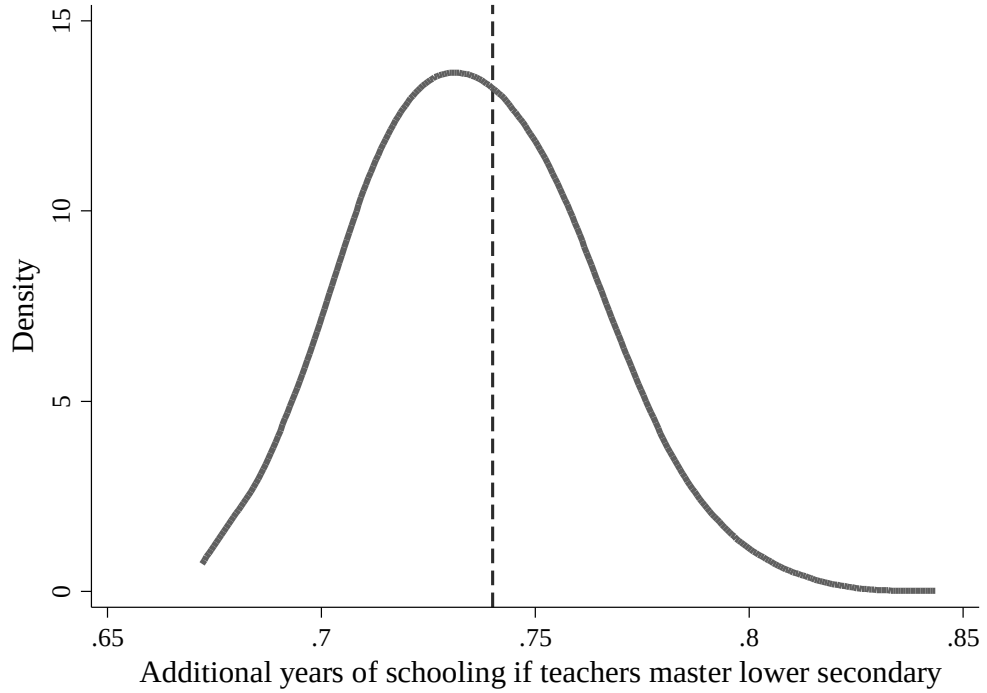
Counterfactual policy experiments

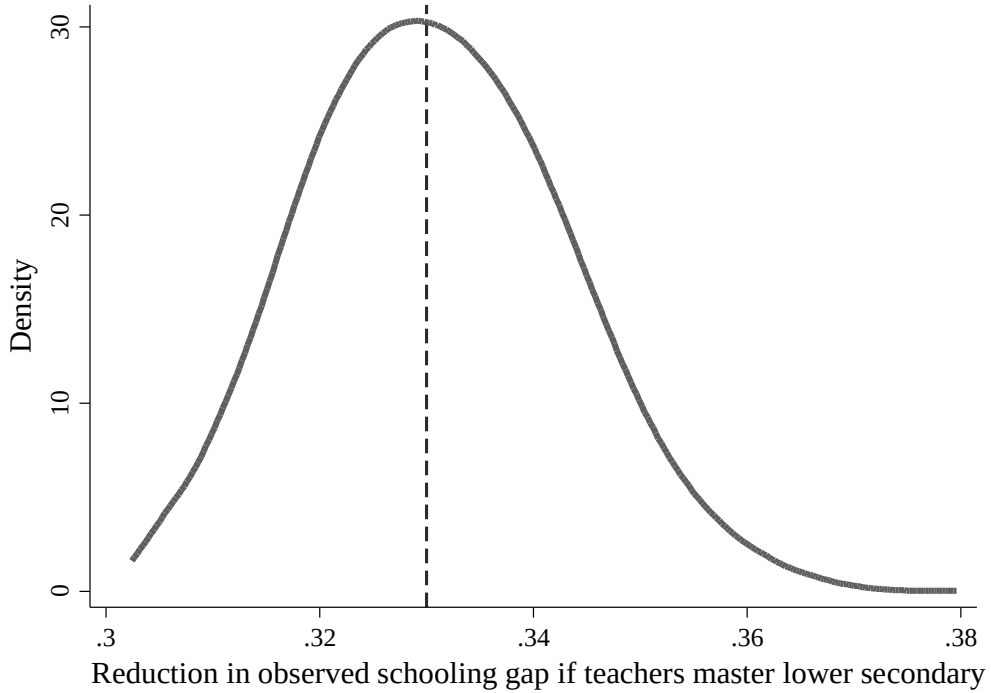
- Given that students lag behind 2.2 years of effective schooling already after four years of primary school and their teachers do not master the primary curriculum, what do these results imply for policy reforms designed to combat the learning crisis?
- **Policy experiment 1 – Accounting exercise**

Taught by teachers with minimum knowledge

- How many effective years of schooling would students accumulate after four years if teachers' effective years of education rose to the lower secondary level (minimum official requirement)?
- Policy experiment is equivalent to an increase of 6.5 years of teachers' effective years of education relative to the current average of 3.5 years.

Taught by teachers with minimum knowledge





Taught by teachers with minimum knowledge

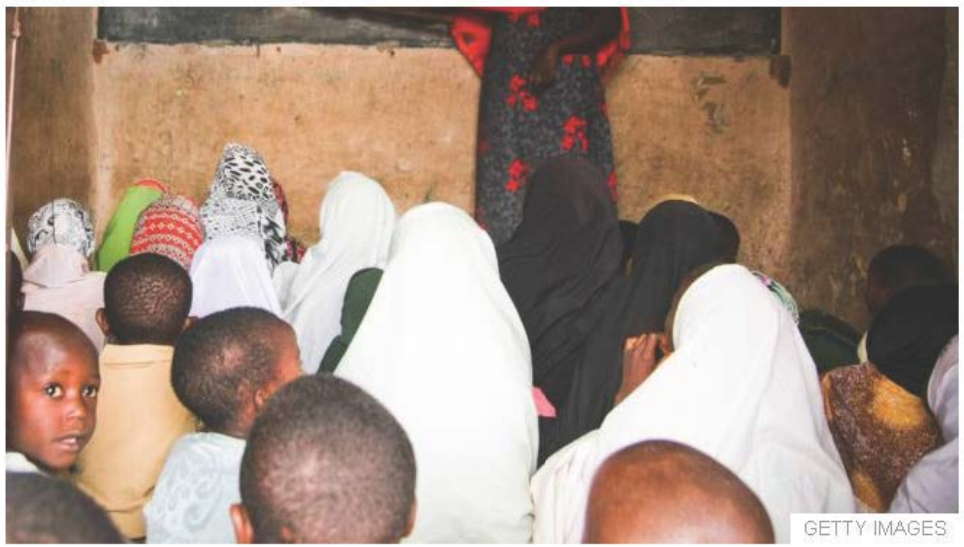
- Hold other dimensions of teacher quality constant and at a low level
  - Low teacher content knowledge may be the most important factor in accounting for the fact that children learn little from attending school
  - Reforms that focus purely on teacher knowledge and training would require teachers in Sub-Saharan Africa to complete (effective) education exceeding university level in order to completely close the gap in student learning that has opened up by grade 4

BBC

Home | News | Sport | Weather | Shop | Earth | Travel

Nigeria schools: Kaduna primary teachers fail pupils' exam

10 October 2017 | Africa



GETTY IMAGES

Thousands of primary school teachers in Nigeria's northern Kaduna state are to be sacked after failing the exams they set for their six-year-old pupils.

State governor Nasir El-Rufai said 21,780 teachers, two-thirds of the total, had failed to score 75% or higher on assessments given to pupils.

He said 25,000 new teachers would be recruited to replace them.

Mr El-Rufai made the comments at a meeting with World Bank representatives in the state capital, Kaduna.

"The hiring of teachers in the past was politicised and we intend to change that by bringing in young and qualified primary school teachers to restore the dignity of education in the state," he was quoted as saying by Nigeria's Daily Trust newspaper.

Policy experiment 2: The Nigerian experiment

BBC

Home | News | Sport | Weather | Shop | Earth | Travel

Nigeria schools: Kaduna primary teachers fail pupils' exam

10 October 2017 | Africa

Thousands of primary school teachers in Nigeria's northern Kaduna state are to be sacked after failing the exams they set for their six-year-old pupils.

State governor Nasir El-Rufai said 21,780 teachers, two-thirds of the total, had failed to score 75% or higher on assessments given to pupils.

He said 25,000 new teachers would be recruited to replace them.

Mr El-Rufai made the comments at a meeting with World Bank representatives in the state capital, Kaduna.

"The hiring of teachers in the past was politicised and we intend to change that by bringing in young and qualified primary school teachers to restore the dignity of education in the state," he was quoted as saying by Nigeria's Daily Trust newspaper.

Policy experiment 2

- Fire the 50% worst performing teachers

A. Incentive the remaining teachers to teach 100% of schedule time

B. Increase class sizes

What do these findings imply for education systems and education policy reform and for research?

Discussion

- Easy at a general level to list what governments “should” do to improve service performance in the education sector
  - teacher training programs should seek to attract talented candidates and prepare them to teach the curriculum effectively.
  - after teachers are hired, the need is for effective incentive schemes that ensure high effort and continued upgrading of knowledge and skills.
- But reforms aimed at systematically raising the quality of the teaching body along these lines should be viewed as more of a longer run solution.
  - huge improvement in the delivery of high quality education in countries such as South Korea and Singapore resulted from system-wide efforts over several decades

Discussion

- Millions of children in low income countries cannot afford to wait for system-wide reforms to be identified and implemented.
- While planning for longer-term solutions, it is also important to consider shorter-term improvements.

Summing up

- Provide a complementary approach to the VA approach to estimate the importance of teacher content knowledge
- We estimate two key structural parameters which can be used to run counterfactual policy simulations
- Show that had all students been taught by teachers deemed to master the lower secondary curriculum, the observed gap in effective education after four years would have been reduced by one third

Thank You!