The Lost Human Capital

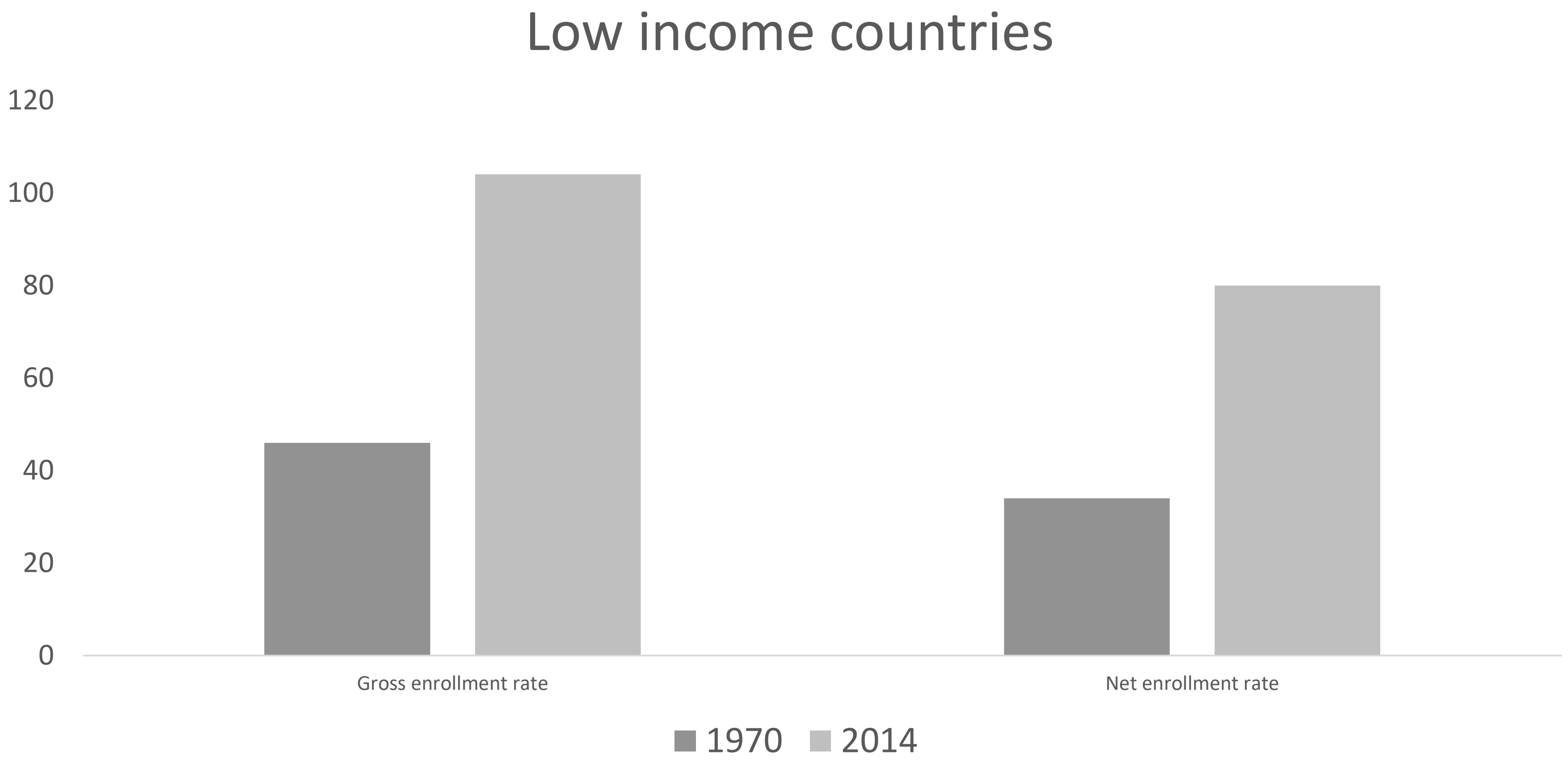
Teacher Knowledge and Student Achievement in Africa

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

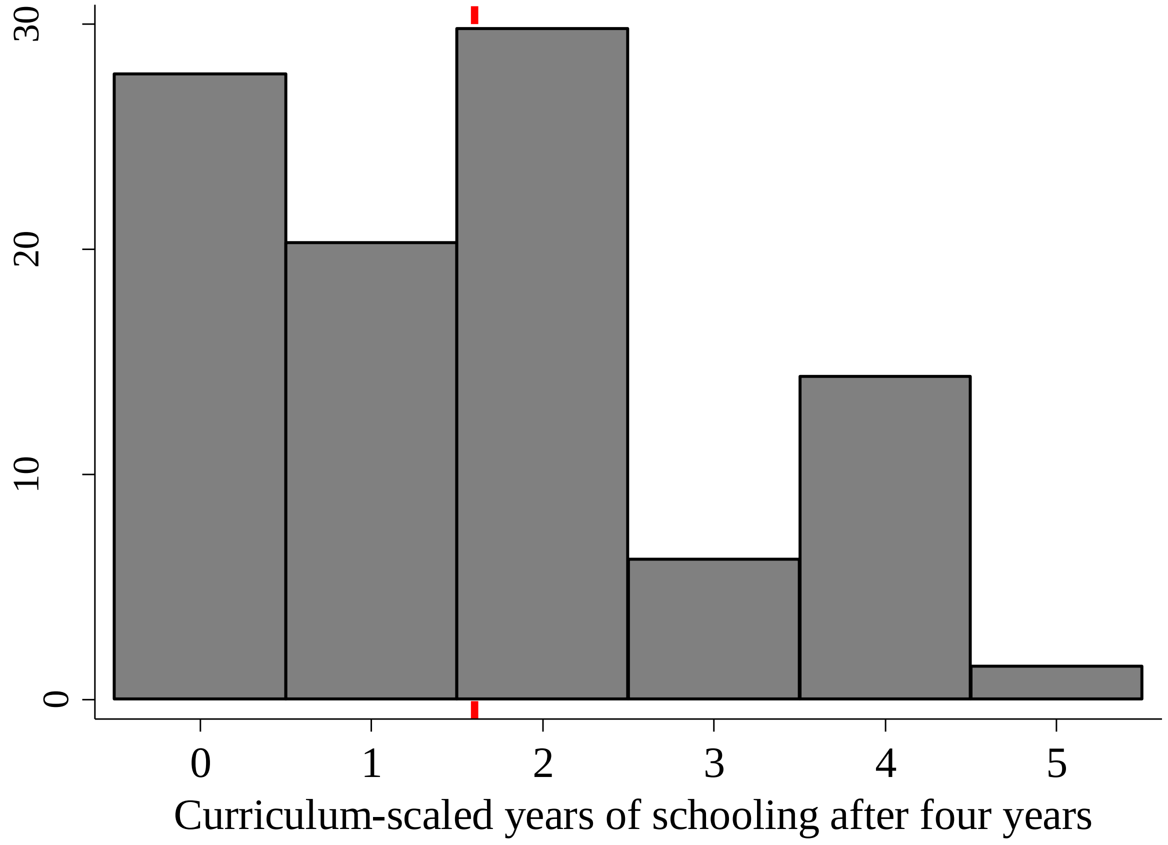
What do we know about education in LIC?

- Two facts

Large increase in enrollment (primary school)



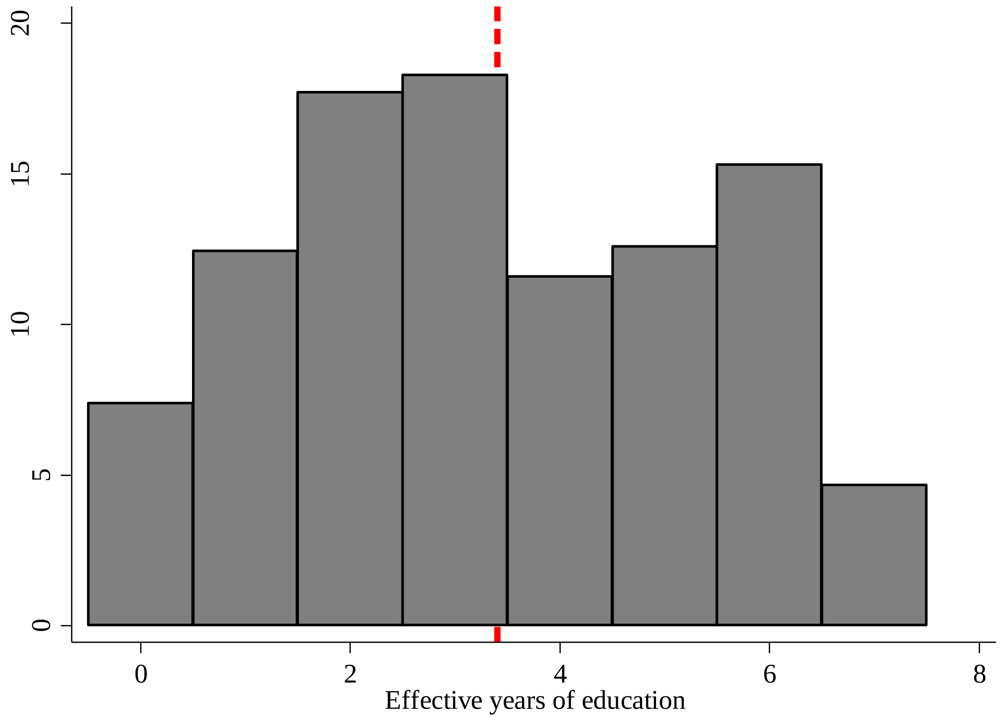
but without much learning



Why is learning so low? Teacher quality

- 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.
- But teacher quality is low in many sub-Saharan African countries (Bold et al, 2017).
  - Classroom absence across Sub-Saharan Africa is 44%.
  - Two thirds of teachers have knowledge equivalent to the students (4^{th} graders) they are teaching.
  - Hardly any teachers have sufficient pedagogical knowledge for teaching.
- Can low teacher quality and in particular subject knowledge explain low student learning?

Effective years of education of teachers



Alternative definition

The Data

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
- In total data on 2,600 schools, over 21,000 teachers and 24,000 students
- **Of particular importance for this paper: student achievement in language and mathematics in grade 4, matched to current and previous teacher knowledge in the same subjects**

Does it matter?

Statistical model for cognitive achievement

(1) $$y_{ijt,k} = F\left[T_{ij,k}(t), S_{ij,k}(t), P_{ij,k}(t), \omega_{ij}, \varepsilon_{ijt,k}\right]$$

$T_{ijt,k} = \left\{x_{ijt,k}, c_{ijt,k}, \bar{x}_{ijt}, \bar{c}_{ijt}\right\}$ is a vector of vector of teacher-supplied inputs

$x_{ijt,k}$ subject content knowledge

$c_{ijt,k}$ other subject-by-teacher characteristics/skills

$\bar{x}_{ijt}$ and $\bar{c}_{ijt}$ are corresponding subject-invariant terms

Statistical model for cognitive achievement

(1) $$y_{ijt,k} = F\left[T_{ij,k}(t), S_{ij,k}(t), P_{ij,k}(t), \omega_{ij}, \varepsilon_{ijt,k}\right]$$

$S_{ijt,k} = \left\{s_{ijt,k}, \bar{s}_{ijt}\right\}$ is a vector of school-supplied inputs

$s_{ijt,k}$ subject-specific inputs

$\bar{s}_{ijt}$ subject-specific inputs

Statistical model for cognitive achievement

(1) $$y_{ijt,k} = F\left[T_{ij,k}(t), S_{ij,k}(t), P_{ij,k}(t), \omega_{ij}, \varepsilon_{ijt,k}\right]$$

$P_{ijt,k} = \left\{p_{ijt,k}, \bar{p}_{ijt}\right\}$ is a vector of parent-supplied inputs

$p_{ijt,k}$ subject-specific inputs

$\bar{p}_{ijt}$ subject-specific inputs

Statistical model for cognitive achievement

(1) $$y_{ijt,k} = F\left[T_{ij,k}(t), S_{ij,k}(t), P_{ij,k}(t), \omega_{ij}\right]$$

ω_{ij} individual's innate ability (or motivation)

Identifying the structural parameters

- Linearizing the production function (1) and express as difference across subjects
  - All subject-invariant unobserved heterogeneity at the student, school and parent level is removed
- Restrict attention to students who were taught by class teachers
  - Any teacher-specific, subject-invariant heterogeneity in grade 4 and 3 removed
- Assuming age-independent contemporaneous effects and constant fade-out, we can rewrite the production function as

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

Estimating a range for α and γ

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

Assuming $\text{cov}(\Delta x_t, \epsilon) = 0$, can estimate range for α, γ

OLS estimator of β_4 and β_3 in (3)

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

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

Can show that the sum of $\hat{\beta}_4$+$\hat{\beta}_3$ provides a lower bound for the total cumulative effect of teacher knowledge.

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

Estimating a range for α and γ

Range is determined by assumptions on correlation between observed and unobserved teacher knowledge:

1) $\rho_{t,t'} \geq 0$
2) $\rho_{t,t'}$ is decreasing in $|t - t'|$.
3) $\rho_{t,t-1}$ is decreasing in t.

Identification: Approach

Is assumption that cov$(\Delta x_t, \epsilon) = 0$ warranted?

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 identified if students, parents, schools do not sort/respond to subject-differences in teacher content knowledge.

Within-student within-teacher variation

$$\epsilon_{ij} = \Delta\omega_{ij} + \sum_{t=1}^{4}\left[\theta_t^S \Delta s_{ijt} + \theta_t^P \Delta p_{ijt} + \theta_t^T \Delta c_{ijt|f_k=f_{k'}}\right] + \sum_{t=1}^{2} \theta_t^{T\prime} \Delta C_{ijt|f_k \neq f_{k'}} + \Delta\varepsilon_{ijt}$$

$\Delta\omega_{ij} \neq 0$ only if students have subject specific abilities/motivations.

Identifying assumption: rules out that students systematically sort, based on these subject-specific abilities, into schools with subject-specific teacher knowledge.

Example: students with relatively higher motivation for math sort into schools with relatively more knowledgeable math teachers

Within-student within-teacher variation

$$\epsilon_{ij}$$

$$= \Delta\omega_{ij} + \sum_{t=1}^{4}\left[\theta_t^S \Delta s_{ijt} + \theta_t^P \Delta p_{ijt} + \theta_t^T \Delta c_{ijt|f_k=f_{k'}}\right] + \sum_{t=1}^{2} \theta_t^{T'} \Delta C_{ijt|f_k \neq f_{k'}} + \Delta\varepsilon_{ijt}$$

Allow for parents (or school) to respond to their children's low maths skills by providing additional teaching or other inputs, but they cannot do this to compensate for insufficient teacher mathematics knowledge.

In the context of lower primary schooling in Africa, these assumptions appear reasonable

- Araujo et al. (2016) find that while parents recognize better teachers, they do not change their behaviors to take account of differences in teacher quality.

Within-student within-teacher variation

$$\epsilon_{ij}$$
$$= \Delta\omega_{ij} + \sum_{t=1}^{4}\left[\theta_t^S \Delta s_{ijt} + \theta_t^P \Delta p_{ijt} + \theta_t^T \Delta c_{ijt|f_k=f_{k'}}\right] + \sum_{t=1}^{2} \theta_t^{T'} \Delta C_{ijt|f_k \neq f_{k'}} + \Delta\varepsilon_{ijt}$$

A teacher, teaching both subjects, may be more motivated to teach a subject she masters relatively well

Put more effort into teaching if she is less knowledgeable of the subject

To the extent these additional subject-specific traits are systematically correlated with teacher subject-specific content knowledge, α_t needs to be reinterpreted slightly more broadly

The impact of teacher content knowledge and other unmeasured teacher subject-specific teaching traits correlated with it

Findings

Reduced Form Results

	(1)	(2)	(3)	(4)	(5)
Dep. variable			Student test score		
Content knowledge of current teacher	0.175***	0.082***	0.068**	0.060***	0.034
	(.013)	(.013)	(.017)	(.022)	(.027)
Content knowledge of prior teacher				0.038*	0.049*
				(.021)	(.027)
Constant	0.504***	-0.030***	-0.184***	-0.069***	-0.185***
	(.030)	(.006)	(.011)	(.009)	(.016)
Lower bound (total effect)				0.099***	0.083***
Number of schools	1,974	1,503	1,503	1,503	626
Number of students	16,922	10,324	10,324	10,324	4,503
Country FE	x				
Student FE		x	x	x	X
Same teacher			x		X

Reduced Form Results

	(1)	(2)	(3)	(4)	(5)
Dep. variable			Student test score		
Content knowledge of current teacher	0.175^{***}	0.082^{***}	0.068^{**}	0.060^{***}	0.034
	(.013)	(.013)	(.017)	(.022)	(.027)
Content knowledge of prior teacher				0.038^{*}	0.049^{*}
				(.021)	(.027)
Constant	0.504^{***}	-0.030^{***}	-0.184^{***}	-0.069^{***}	-0.185^{***}
	(.030)	(.006)	(.011)	(.009)	(.016)
Lower bound (total effect)				0.099^{***}	0.083^{***}
Number of schools	1,974	1,503	1,503	1,503	626
Number of students	16,922	10,324	10,324	10,324	4,503
Country FE	x				
Student FE		x	x	x	X
Same teacher			x		X

Reduced Form Results

	(1)	(2)	(3)	(4)	(5)
Dep. variable			Student test score		
Content knowledge of	0.175***	0.082***	0.068**	0.060***	0.034
current teacher	(.013)	(.013)	(.017)	(.022)	(.027)
Content knowledge of				0.038*	0.049*
prior teacher				(.021)	(.027)
Constant	0.504***	-0.030***	-0.184***	-0.069***	-0.185***
	(.030)	(.006)	(.011)	(.009)	(.016)
Lower bound (total effect)				0.099***	0.083***
Number of schools	1,974	1,503	1,503	1,503	626
Number of students	16,922	10,324	10,324	10,324	4,503
Country FE	x				
Student FE		x	x	x	X
Same teacher			x		X

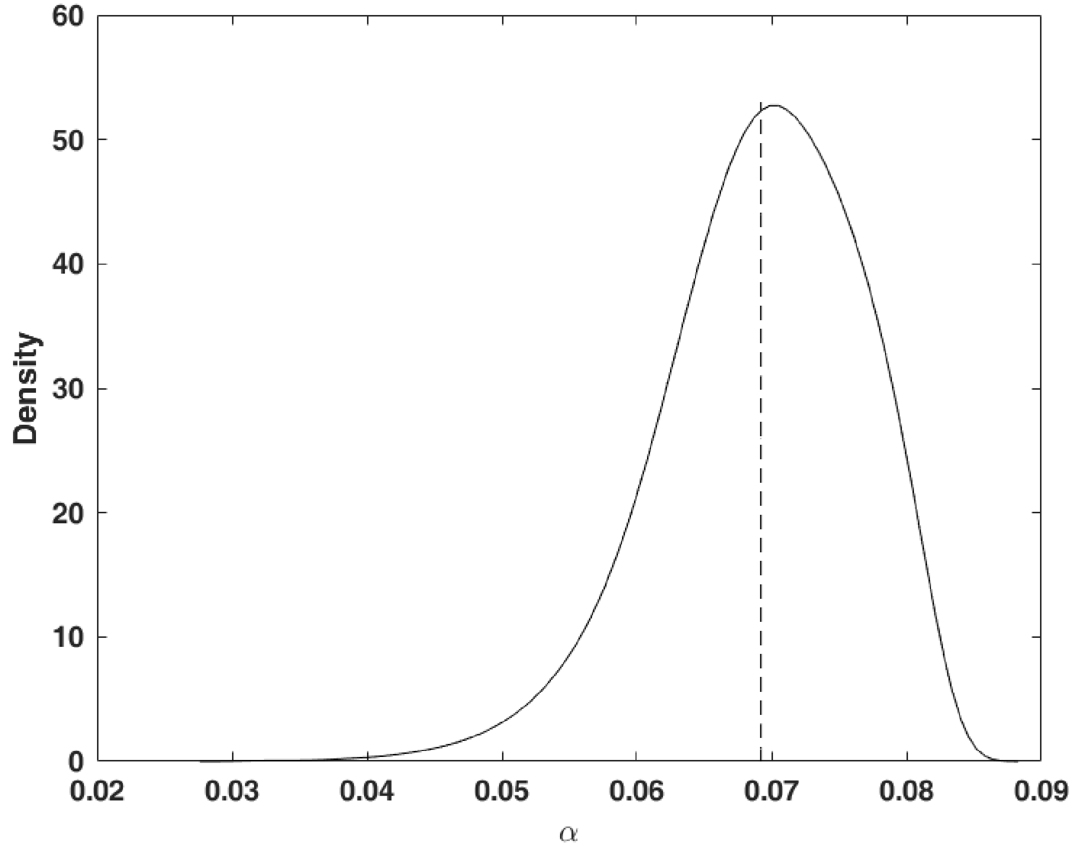
Reduced Form Results

	(1)	(2)	(3)	(4)	(5)
Dep. variable			Student test score		
Content knowledge of current teacher	0.175***	0.082***	0.068**	0.060***	0.034
	(.013)	(.013)	(.017)	(.022)	(.027)
Content knowledge of prior teacher				0.038*	0.049*
				(.021)	(.027)
Constant	0.504***	-0.030***	-0.184***	-0.069***	-0.185***
	(.030)	(.006)	(.011)	(.009)	(.016)
Lower bound (total effect)				0.099***	0.083***
Number of schools	1,974	1,503	1,503	1,503	626
Number of students	16,922	10,324	10,324	10,324	4,503
Country FE	x				
Student FE		x	x	x	X
Same teacher			x		X

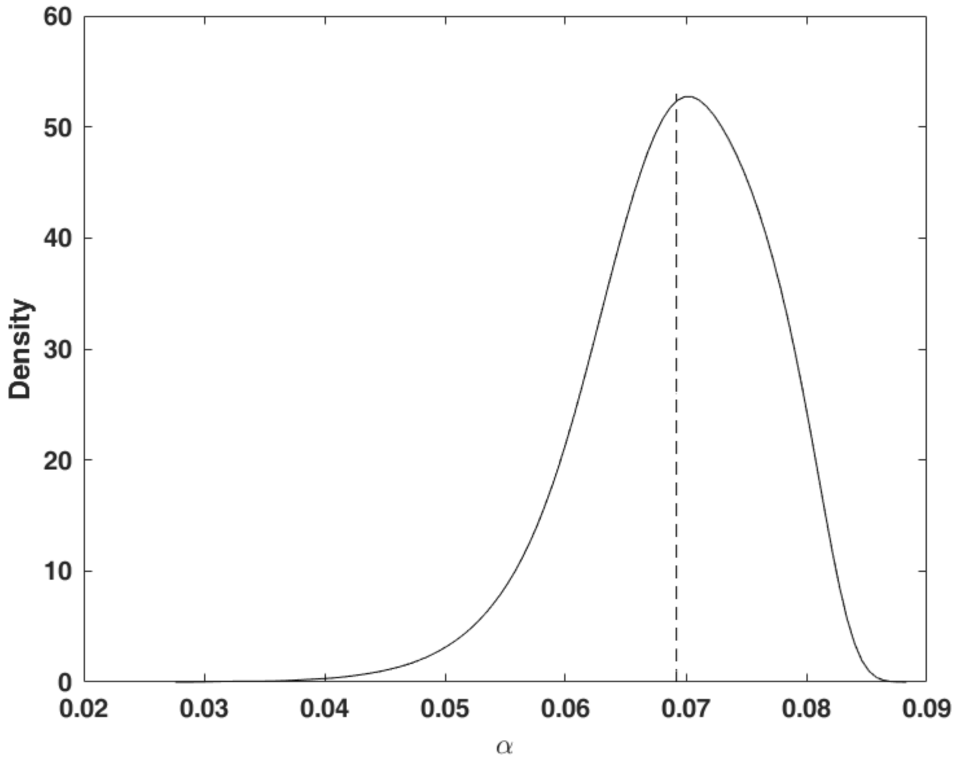
Robustness

- Identification assumptions:
  1. no other factors (at teacher level or otherwise) that drive both student and teacher subject differences in knowledge.
  2. no sorting by students and teachers on the basis of subject differences.

Probability density functions of the estimated α

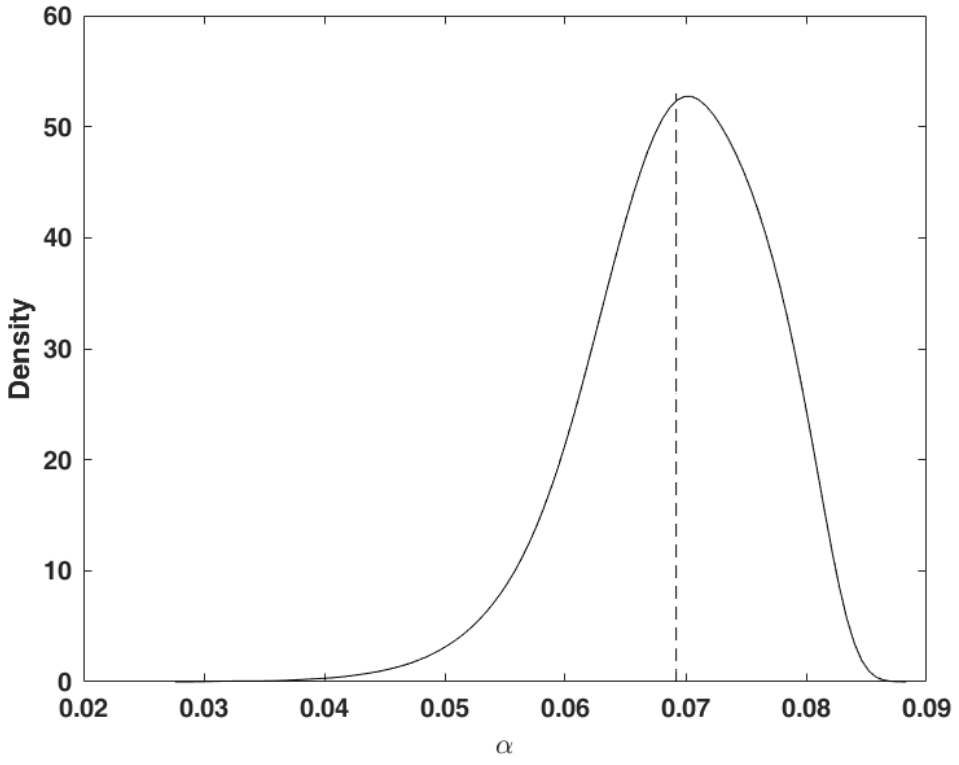


Probability density functions of the estimated α -- IRT measure



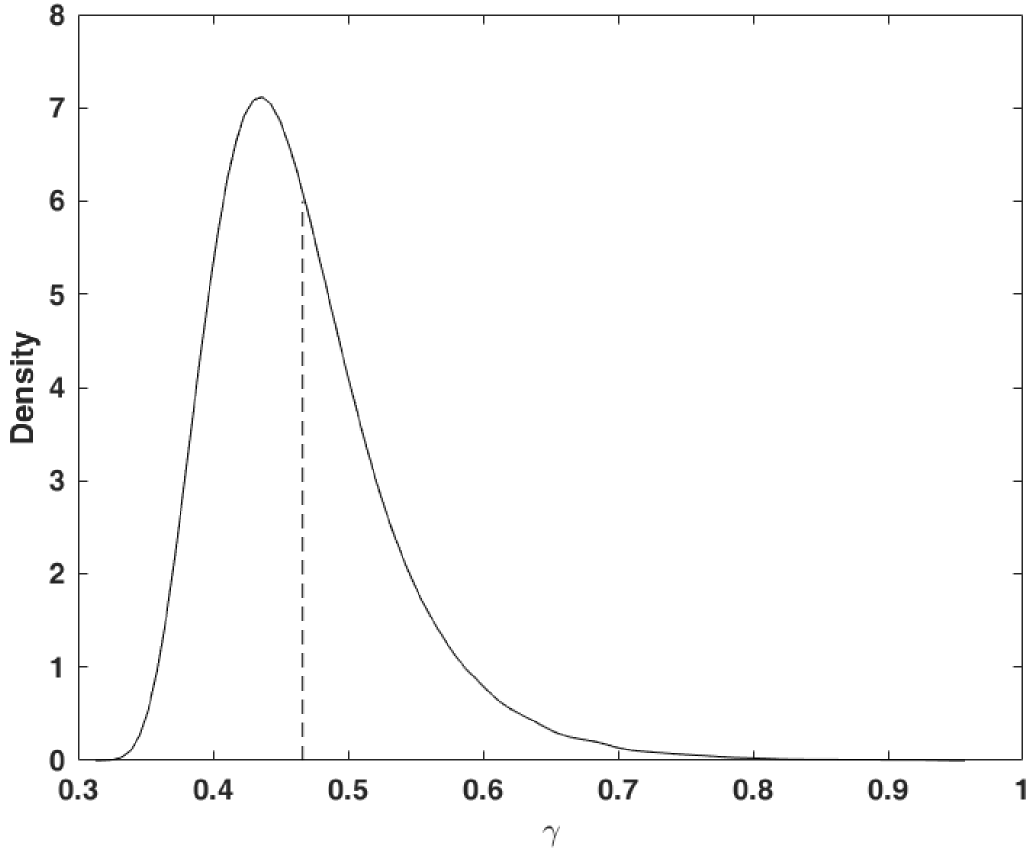
- At α^{med}: A 1 SD increase in effective years of education for a teacher increases student learning by 0.07 SD

Probability density functions of the estimated α -- IRT measure

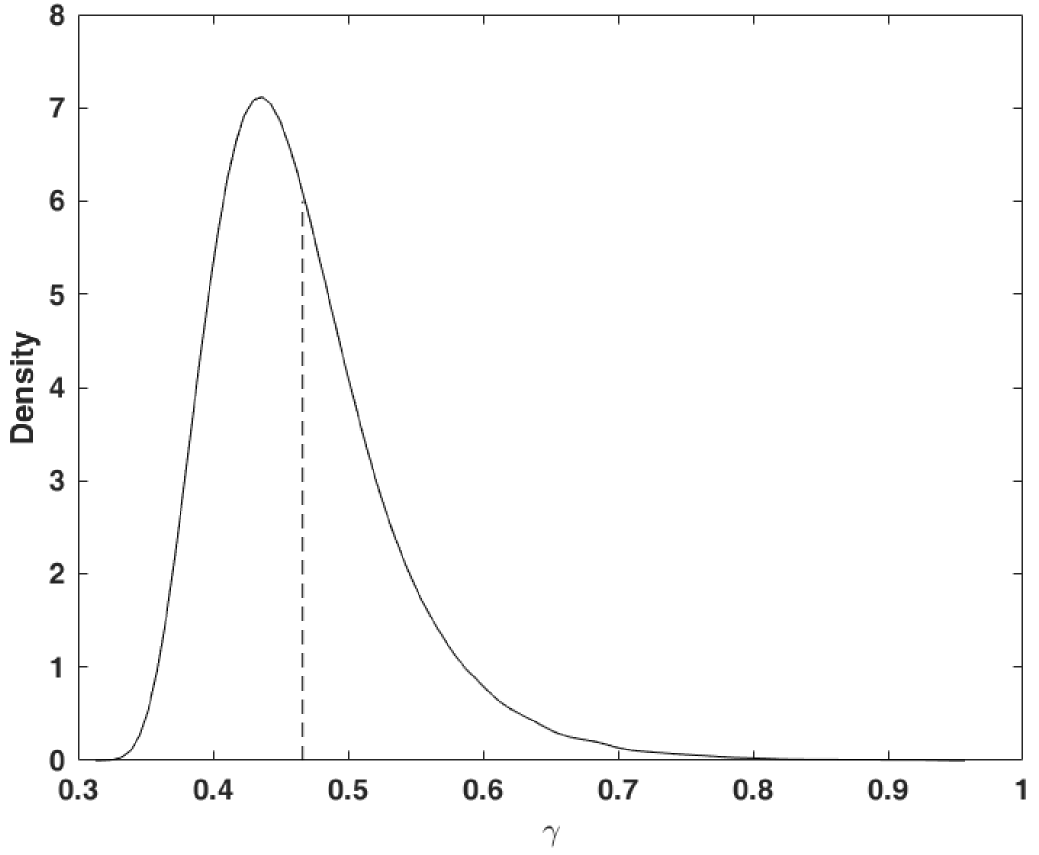


- Effect size = 0.07 SD
- Cf with VA literature: Effect sizes ranging from 0.1-0.2SD (Rockoff, 2004; Rivkin et al., 2005; Aaronson et al., 2007; Chetty et al., 2014; Araujo et al., 2016; and Bau and Das, 2017).
- 1 SD increase in teacher test scores raise student test scores by 0.07 SD (Bau and Das,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

Probability density functions of the estimated γ

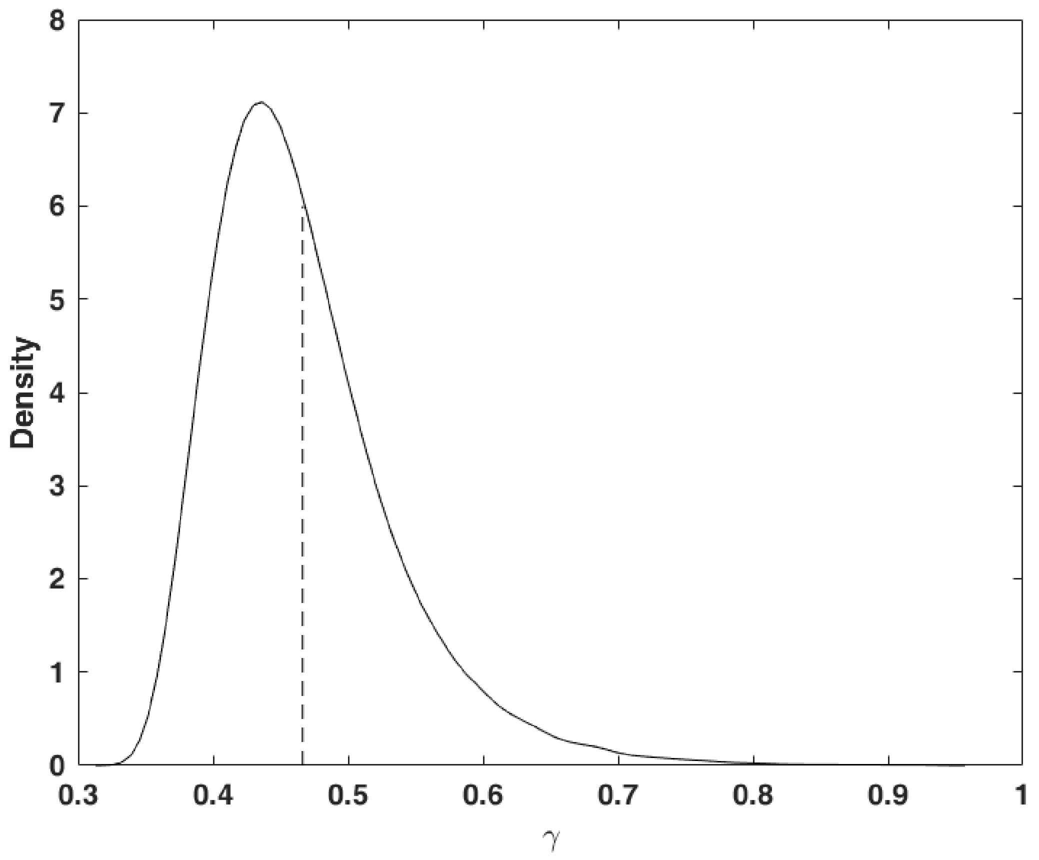


Probability density functions of the estimated γ



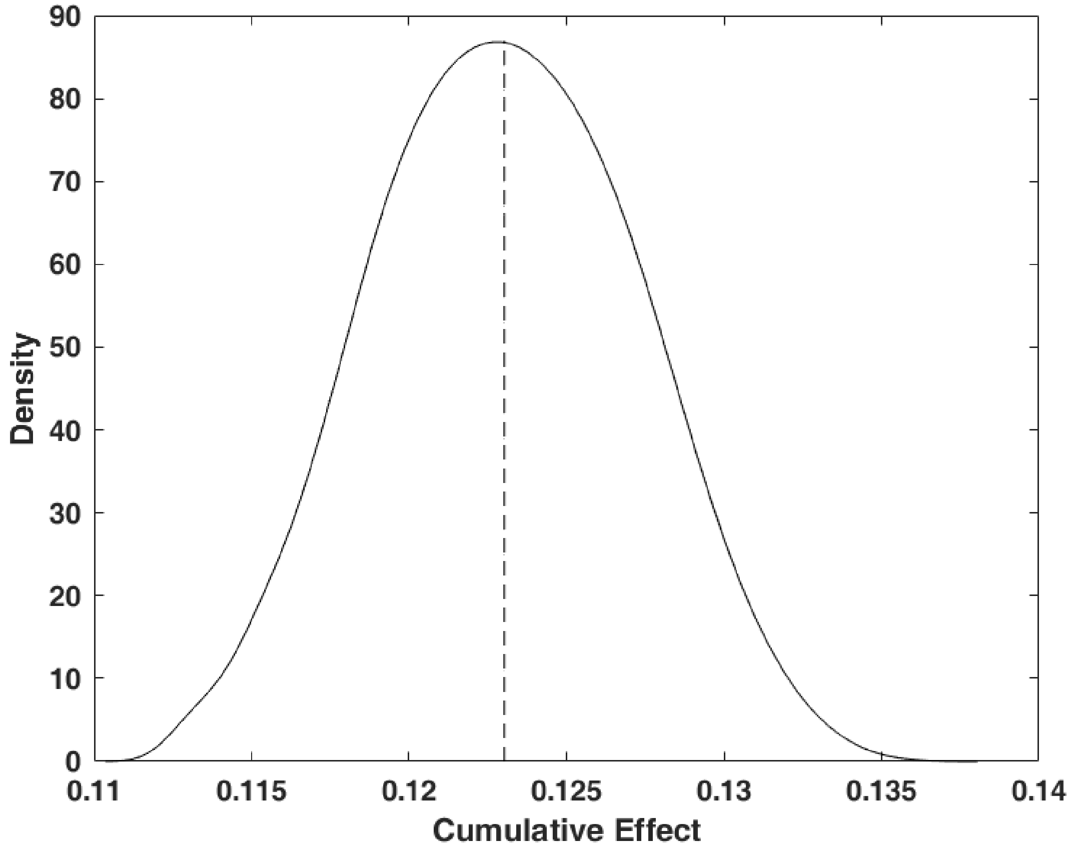
- Approx. 50% of the short-run effect persists between grades

Probability density functions of the estimated γ



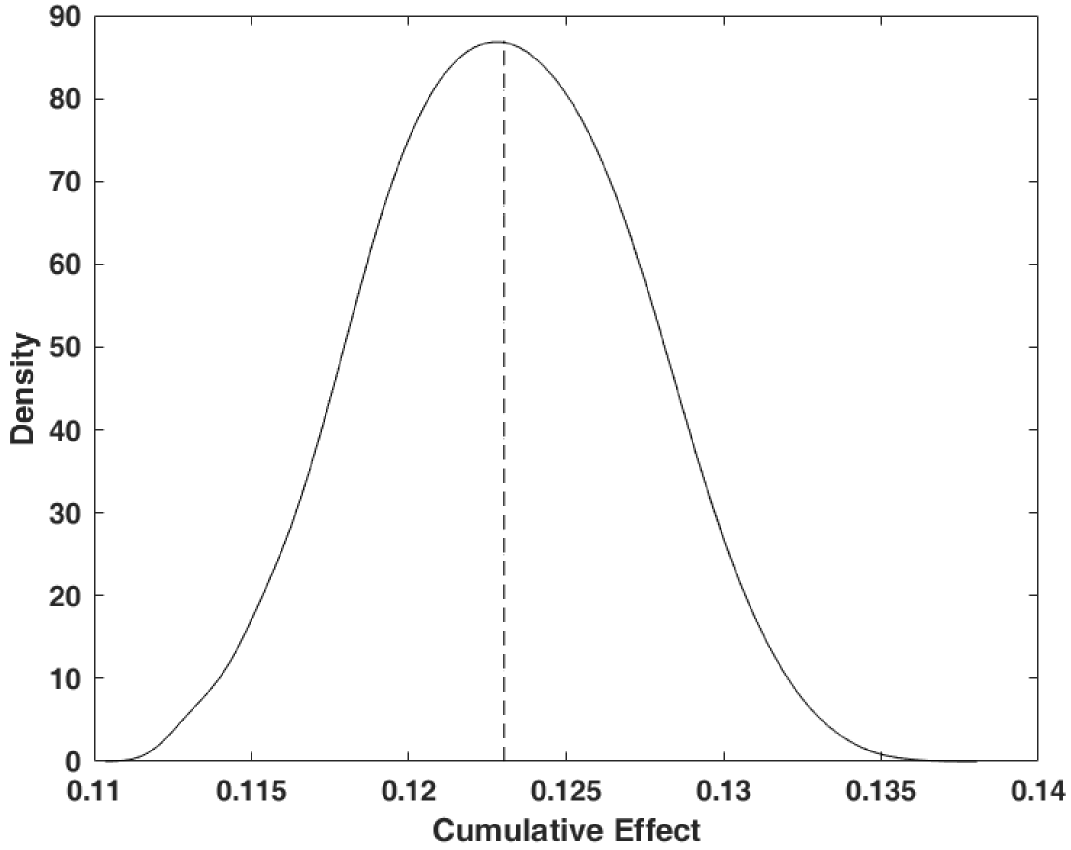
- Approx. 50% of the short-run effect persists between grades
- Consistent with what has been reported using data from Pakistan and the US (Kane and Staiger 2008; Jacob, Lefgren, and Sims 2010; Rothstein 2010; and Andrabi et al., 2011

Probability density functions of the estimated cumulative effect of teacher knowledge



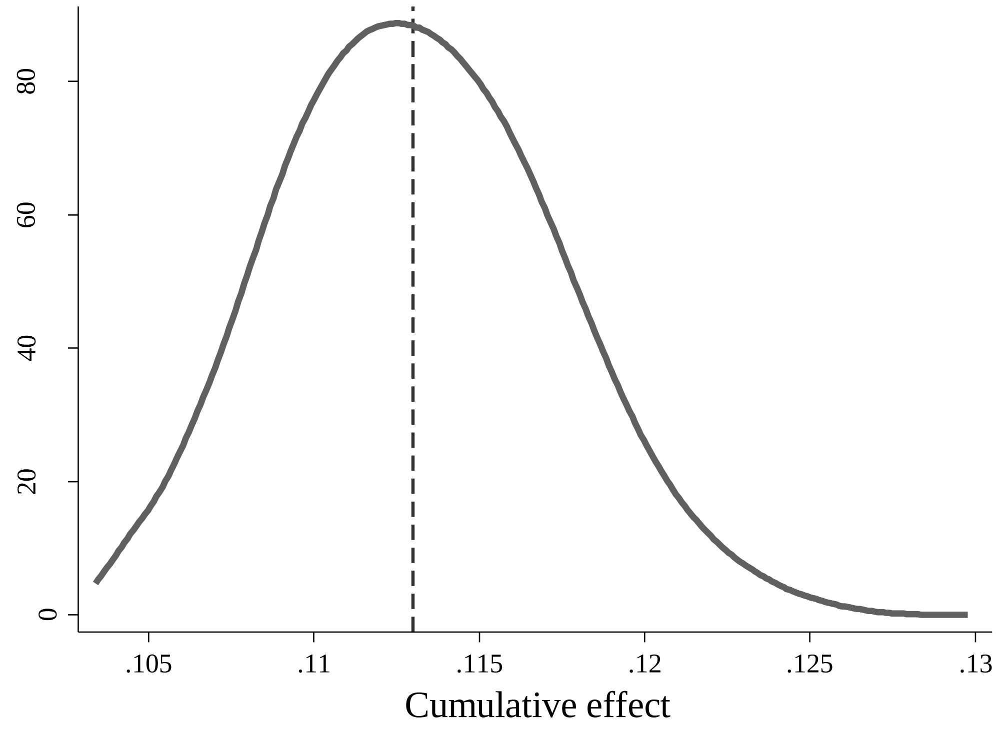
$$CE = \alpha \sum_{t=1}^{4} \gamma^{4-t}$$

Probability density functions of the estimated cumulative effect of teacher knowledge



- Being taught, throughout lower primary, by a teacher with 1SD more subject knowledge increases student learning by .12SD

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 a month and a half after four years.

Policy simulations

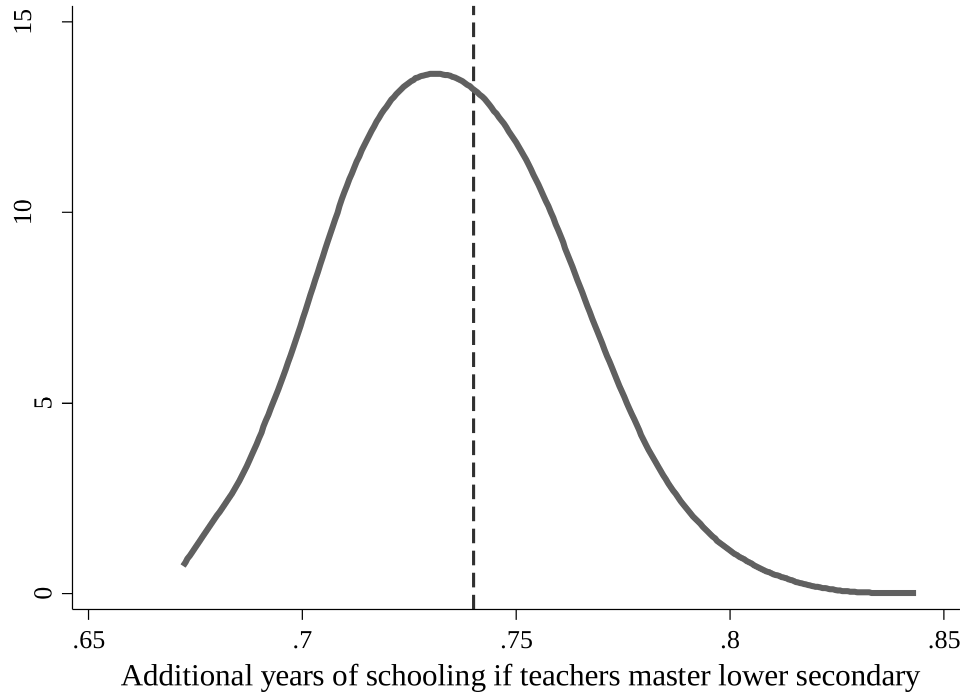
Policy experiments

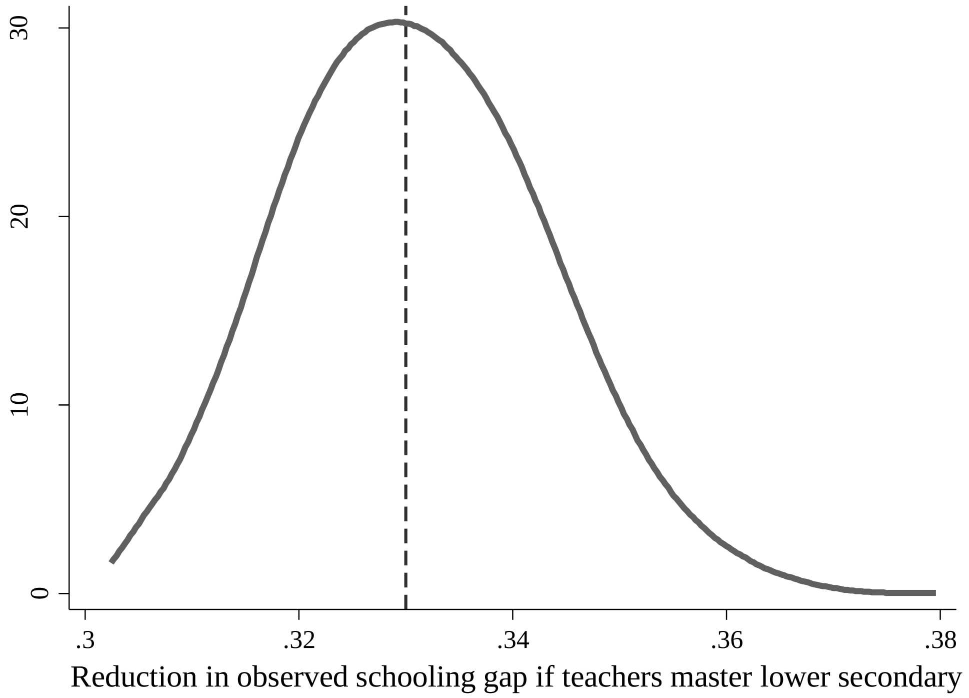
- Use structural estimates for three policy experiments
  - Development Accounting: how much does shortfall in teacher knowledge contribute to low student achievement.
  - Misallocation: how much learning is lost because students are not allocated to the best teachers.
  - Long-term reform: how much would learning increase if all newly hired teachers were properly trained and present on the job.

Policy experiment 1 – Development Accounting

- 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.

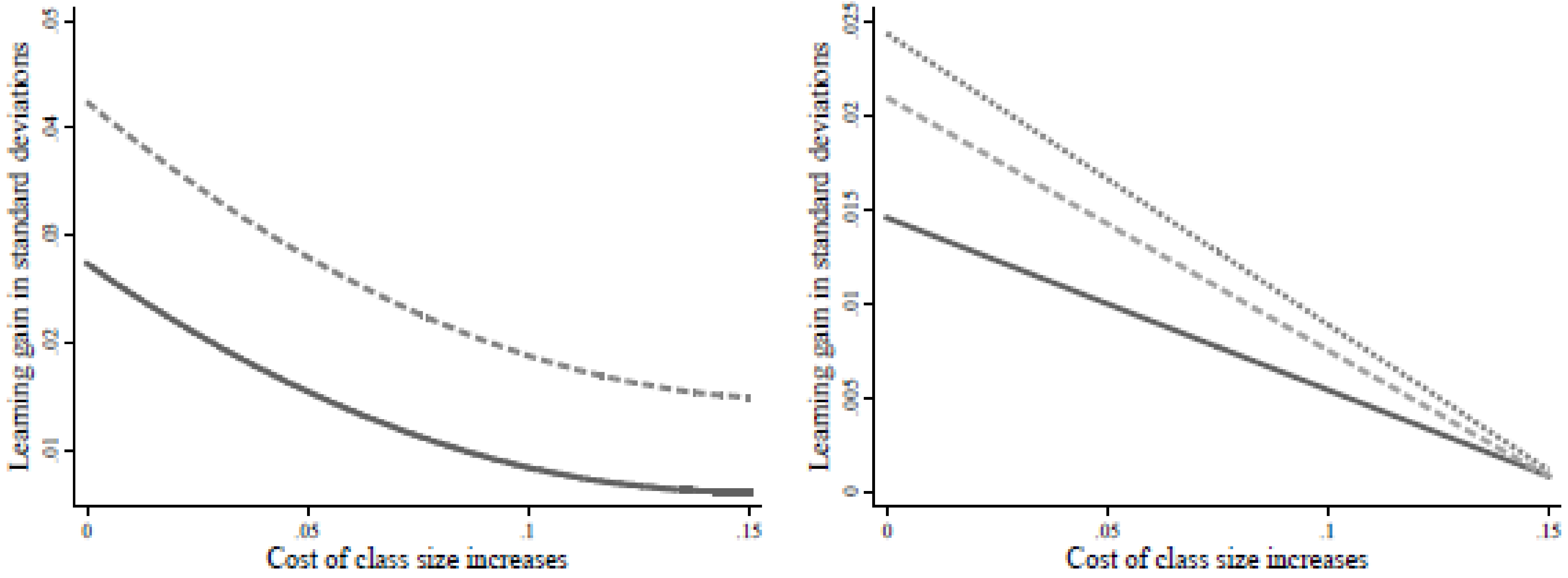
Taught by teachers with minimum knowledge





Policy experiment 2 – Reducing Misallocation

- What is the effect of moving students from the worst performing teachers to those with relatively better content knowledge
- Effects: teacher knowledge
  - Students who move from worst teachers to best teachers are exposed to higher teacher knowledge.
- Effects: class size
  - Students who move from worst to best teachers are exposed to higher class size.
  - Students who are already taught by better teachers are exposed to higher class sizes.
  - Students who remain with worse teachers are exposed to smaller classes.



Panel A: Reallocation at the district level

Panel B: Reallocation at the school level

Policy experiment 3 – Long-run reform

- All newly hired teachers across continent have mandated knowledge and teach mandated hours.

- After 10 years, students would accumulate 16% more learning
- After 30 years, student learning would have almost doubled.

Conclusion

Conclusion

- Teachers’ content knowledge, or lack thereof, is an important part of the reason why primary school students in Sub-Saharan Africa lag behind

➢ Research on how to improve teacher content knowledge should be a priority

- Reform that focuses on new teacher effort and knowledge can double learning, but only over long-term horizon.

➢ Important to experiment with shorter-term approaches, such as programs to supplement current teachers with additional instructors, to leverage computer-aided learning programs, or to support teachers with scripted lesson plans.

Thank You!