The Negative Consequences of Overambitious Curricula in Developing Countries

Lant Pritchett
Amanda Beatty

The Negative Consequences of Overambitious Curricula in Developing Countries

Abstract

Learning profiles, that track changes in student skills per year of schooling, often find shockingly low learning gains. Using data from three recent studies in South Asia and Africa, we show that a majority of students spend years of instruction with no progress on basics. We argue shallow learning profiles are in part the result of curricular paces moving much faster than the pace of learning. To demonstrate the consequences of a gap between the curriculum and student mastery, we construct a simple, formal model, which portrays learning as the result of a match between student skill and instructional levels, rather than the standard (if implicit) assumption that all children learn the same from the same instruction. A simulation shows that two countries with exactly the same *potential* learning could have massively divergent learning outcomes, just because of a gap between curricular and actual pace—and the country which goes *faster* has much *lower* cumulative learning. We also show that our simple simulation model of curricular gaps can replicate existing experimental findings, many of which are otherwise puzzling. Paradoxically, learning could go faster if curricula and teachers were to slow down.

JEL-Code: I210, I250, O150.

Lant Pritchett
Harvard Kennedy School
Cambridge, MA 02138 / USA
lant_pritchett@hks.harvard.edu

Amanda Beatty
Harvard University
Cambridge, MA / USA
amandabeatty@post.harvard.edu

December 10, 2012

The Negative Consequences of Overambitious Curricula[1]

> *If the curriculum were radically simplified, if the teacher's mission were squarely defined as making everyone master every bit of it, and if children were allowed to learn at their own pace, by repeating if necessary, the vast majority of children would get something from the years they spend in school.*
>
> Banerjee and Duflo, *Poor Economics* (2011)

The expansion in schooling around the world is one of the amazing successes of our time. More children are in school, longer, and completing more grades that ever before in history. The average years of schooling of the labor force in the developing world *more than tripled* from 1950 to 2010--from 2.0 to 7.2 years (Barro & Lee, 2011). The average worker in Bangladesh in 2010 completed more years of schooling than the typical worker in France in 1975. While the world may miss reaching the Millennium Development Goal of universal primary school completion by 2015—it won't be by much.[2] However, as the world nears a *schooling* goal, there is rising concern this expansion might not coincide with *education* or *learning* goals (Filmer, Hasan, & Pritchett, 2007) . There is cumulating evidence that the *learning profile*—the relationship between years of schooling and measures of student mastery—is far too flat.[3] Children are learning so little from each year of instruction that the completion of even basic schooling leaves children lacking necessary skills. Thus, some development and education organizations are reorienting their programming and goals from schooling to learning.[4] As they do so a key question is *why* there is so little learning and what can be done about it.

I) The puzzle of flat learning profiles

Many studies start from an assessment of student learning tested in a given grade and hence start from the *level* of student achievement. But every student's level of mastery is the just the accumulation of their lifetime learning. Once one begins to investigate the *dynamics* of learning in developing countries the immediate question is: "how can it possibly be that so many children are learning so very little during a year's worth of schooling?"

I.A) Three examples of learning profiles from India

The Andhra Pradesh Randomized Evaluation Study (APRESt), a large and sophisticated experimental study of the impact of educational practices on learning, is ongoing in four districts

---

[1] We would like to thank many, while implicating none. This paper grew out of insights gleaned from conversations with Rukmini Banerji and her tireless efforts to improve learning in India. Conversations with Luis Crouch, Jishnu Das, Annie Duflo, Deon Filmer, Asim Khwaja, Michael Kremer, Leigh Linden, Karthik Muralidharan and Sukhmani Sethi and comments from Justin Sandefur have been very helpful.

[2] Only 19 countries are unlikely to meet the primary completion goal, and 13 countries are off track from meeting the primary school gender parity goal by 2015 (World Bank, 2011).

[3] The learning profile is a purely *descriptive* graphic device—there is no implication of causation.

[4] E.g. the World Bank's 2020 sector strategy and DFID's 2010 strategy are *learning for all*, USAID's strategy is *opportunity through learning,* and AusAid is also adopting learning goals.

of Andhra Pradesh, India. Because one goal of the study is to track learning progress across grades, the *same* questions are asked of the approximately 100,000 children in grades 2 to 5. Figure 1 shows the grade learning profile, the fraction of children in each grade who answer correctly, for four selected items involving simple arithmetic operations[5].

Even for the most mechanical arithmetic operations there is amazingly slow progress from second to fifth grade. For a "level two" competency like "two digit addition without carry" 40 percent of children in grade 2 answer correctly and yet only 70 percent answer correctly by grade 5. This means that of the 60 percent of children who did not already master addition by grade 2 less than half gained the skill in *three full years of additional schooling.* There are three of every ten children who enrolled school, stayed in school until grade 5, and never learned even the simplest addition.

Figure 1: Learning profiles from Andhra Pradesh, India demonstrate only a minority of children learn basic addition skills between grades 2 and 5
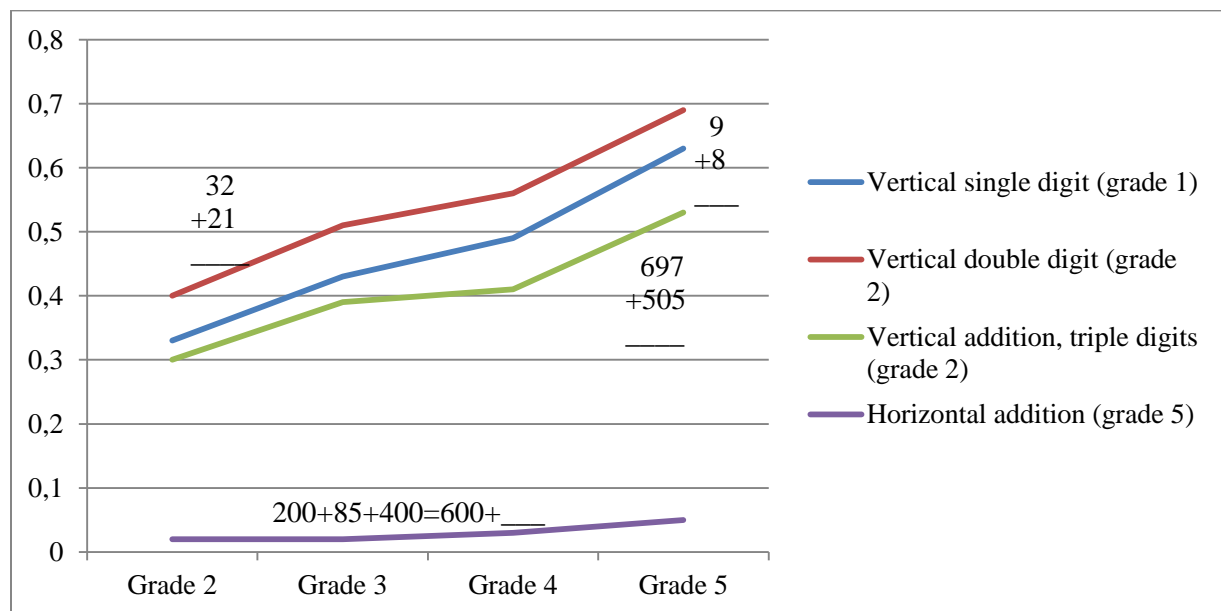
Source: APRESt data provided by Karthik Muralidharan

Moreover, the results are even worse when one looks at conceptual learning. Fewer children could add 9+8 than 32+21, as "carry" requires some knowledge of place. While half of students could correctly add three digit numbers when presented in a mechanical fashion (but only half) but few could use that skill to answer the simple question that required rearranging addition. Less than 10 percent could answer: 200+85+400=600+___.

A second estimate of a learning profile comes from Educational Initiatives (EI) (2010), which tests students' conceptual understanding. EI's study covered in-school children in grades 4, 6 and 8, in government schools in 18 states in India, representing about 74 percent of the Indian population (urban and rural). EI asked a common set of three questions in the same grades in

---

[5] This is not tracking the learning dynamics of individual students but is the description of performance of students across grades.

language and mathematics. For the language questions (e.g., adding a word to complete a sentence), the percent correct only increased from 51 percent to 57 percent from grades 4 to 8. In math, the average percent correct went up by only 7.1 percentage points per year.

As with the APRESt results, EI found that even this progress in some ways overstates the learning gains because students perform much better on mechanical than conceptual skills. For example answering a question like 43x2 (written vertically) was much easier for 4th graders, 67 percent of whom answered correctly, than a question that asks 3 x ? = 3+3+3+3 (solving for the question mark), which only 30 percent of students answered correctly. Even though the multiplication required is much easier, the question as posed requires conceptual understanding that multiplication is repeated addition.

The Annual Status of Education Report (ASER) testing methods and reporting process were developed and administered by the Indian education NGO Pratham and more recently by the ASER Centre. (So ASER is the name of the organization that administers tests and produces results, the name of its test, and the name of its report showcasing test results.) This testing exercise is valuable because it is large-scale (over 500,000 children tested), carried out across India each year, and it tests all children, not just those in school. This child-based information can be used to construct a learning profile by self-reported grade completed. By grade 5 (the end of primary school), only 1 in 5 children in Uttar Pradesh in 2007 could do a simple division problem (one digit into three digits), and only 40 percent could read a text at the "grade 2" level. Even by grade 8, only about 50 percent of students could do the simple division problem, and just above 70 percent could read a grade 2 text. Almost three in ten students who have persisted through *eight years* of schooling still cannot read a simple paragraph (ASER, 2008).

Figure 2: Learning profiles from Uttar Pradesh, India demonstrate that only 75-80 percent of children persisting thought eight years of schooling achieve basic reading and mathematics skills
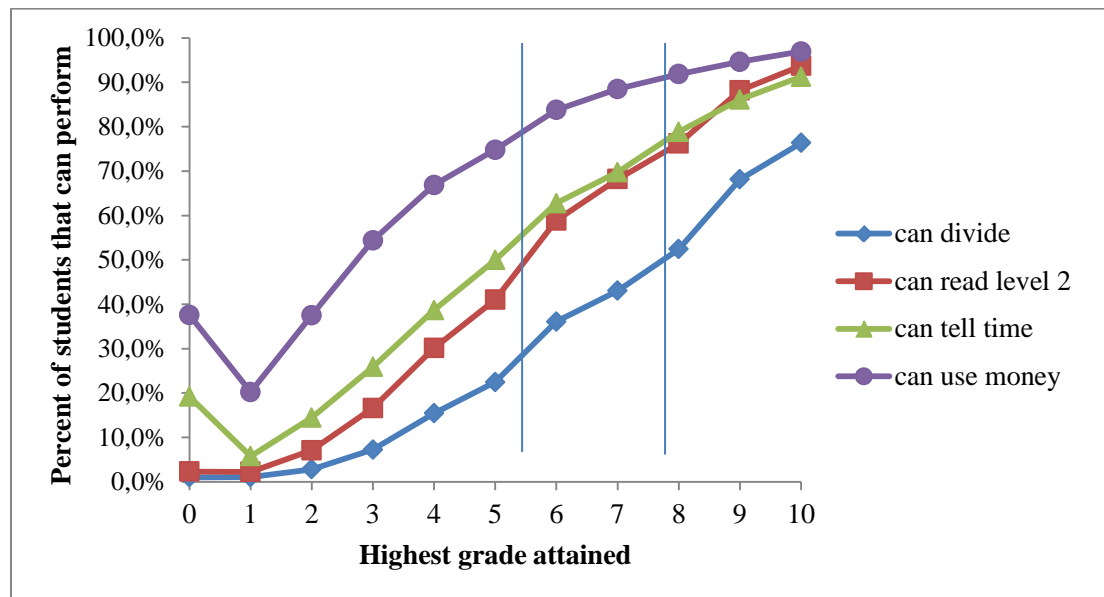
Source: Authors' calculations with ASER 2008 data

Even most studies that are able to give learning profiles are synthetic in that they compare not the same children over time but children in different grades. One very recent study Bhattacharjea, Wadhwa and Banerji 2001) tracked roughly 22,000 children across grades (grade 2 to grade 3 and grade 4 to grade 5) testing them in a set of language and mathematical capabilities. This shows the learning gains of actual children as they moved from grade to grade. This study also shows very slow progress. For instance 41.4 percent of children could write a dictated word in grade 2 and only 54.7 percent could do so in grade 3. As usual, there are huge variations across states. In Himachal Pradesh the number of children who could do a word problem involving one digit increased from 12.1 percent to 22.9 percent. In Rajasthan this increased from 7.1 to only 9.6 percent—an only 2.5 percentage point increase—and in Jharkhand the proportion actually *fell* from 14.8 percent in grade 2 to only 11.0 percent when the same children were tested in grade 3. On average, across the 11 items in mathematics asked (all relatively simple numerical or word problems) the percent of gain as a fraction of those who did not already know was only 12.7 percent, implying *seven of eight* children made no progress on the questions asked.

Table 1 illustrates learning progress per year of schooling using selected questions from these three studies. Different studies employing different strategies to measure different components of mathematics skill consistently demonstrate that there is just a 3 to 12 percentage point increase in student mastery of basic mathematics per year. The last column shows the percent of students who learned to answer the question in a typical additional year of instruction.[6] This

---

[6] Again, this is the net gain not tracking of individual students so does not capture the "churn" in the learning and forgetting of individual students.

ranges from only 4 percent (on the conceptual question of measurement) to 20 percent (on doing division).

Table 1: Studies across India demonstrate very little conceptual or procedural mathematics learning as children progress through school

	Skill/competency (examples from each assessment)	Average percentage point increase per year of schooling	Of students who didn't know skill, percent who learned in the next grade
APRESt (four districts of Andhra Pradesh)	Vertical single digit addition (w carry)	10.0	17.6
	Identify which shape is a triangle	5.3	7.0
	11 common questions, grades 2 to 5	**6.1**	**9.4 (≈ 1 in 11)**
ASER (nationwide, rural)	**Division (3 digit by 1 digit), grades 2 to 8**	**9.0**	**12.6 (≈1 in 8)**
	Reading grade 2 story, grades 2 to 8	**11.8**	**21.8 (≈1 in 5)**
ASER (five states, rural)	**11 questions in mathematics (numerical and word problems), from grade 4 to 5**	**9.5**	**12.7 (≈1 in 8)**
Education Initiatives (18 states)	19+X=32	7.5	16.9
	Measuring length with ruler (when object is placed at 1cm)[7]	2.9	3.9
	All items asked across grades (4 and 6, 6 and 8, or 4,6 and 8)	**4.9**	**8.5 (≈1 in 12)**

Sources and notes: APRESt: four districts in Andhra Pradesh. Average yearly gain from grades 2-5.
ASER nationwide: 2012. Grades 2-8. Sample is every district in India, 30 villages per district and 20 households per village. Average yearly gain grades 2-8 since division and reading are grade 2 competencies as stipulated by the national curriculum.
ASER five-state (Assam, Andhra Pradesh, Himachal Pradesh, Jharkhand, Rajasthan) report by Bhattacharjea, Wadhwa, and Banerji (2011) follows the *same* children across grades. Sample is three districts per state, 60 government primary schools per district.
EI: 18-state average for grades 4, 6, 8 with grades 5 and 7 imputed.

Looking at all questions for which we have learning profiles from the three studies, we find that the *learning*, measured as the net addition in percent correct, is only about 8 to 22 percent per year. That is, of children lacking the ability to answer these simple curricular-based questions

---

[7] One remarkable finding regarding this question is that ASER used it in a five-state study, yet for teachers. Only 82 percent of teachers willing to participate in the testing were able to answer this correctly. See Bhattacharjea, Wadhwa and Banerji 2011.

going into a typical grade, only about 1 in 8 demonstrate that skill after an additional year of schooling. This means that *seven out of eight* children made no progress on a typical item after an entire year of schooling. Even *at best*, with the simplest skills like reading a simple passage, four out of five children who go into a grade not able to read will finish the grade still unable to read.

These numbers are shocking. How can learning profiles be so flat? What is going on inside the classroom such that *nearly all students* who cannot read going into grade 4 don't learn to read while in grade 4? How can a child sit through an entire year of schooling not knowing how to do simple addition and not learn to add? What is going on in the schooling process that children can spend *thousands of hours* in school and yet not master even the basic competencies?

I.B) Learning profiles and cumulative achievement in the developing world

This problem of a flat learning profile is not unique to India. Several other countries have implemented ASER-like instruments to measure what fraction of children can do division or read a simple level 2 passage. Thus these countries implemented nearly identical tests and found the fraction who learn the expected skills in grades 2 to 5 is similarly low—only about 8 to 34 percent of those who don't have the skill gain it in a year of instruction between grades 2 and 5.

**Table 2: Similar assessments around the world show similarly flat learning profiles**

Country	Average percentage point increase per year of schooling, grades 2-5	Of students who didn't know skill, percent who learned in the next grade, grades 2-5
Reading (grade 2 story)		
India	13.2	17.1
Pakistan	14.1	18.7
Tanzania	14.8	21.5
Uganda (English)	7.7	8.2
Kenya	20.4	34.1
Division (one digit into three digits)		
India	8.3	9.3
Pakistan	9.9	11.6
Uganda	11.3	14.4
Kenya	18.6	28.4
Sources: Uwezo Tanzania 2011, Uwezo Uganda 2010, Uwezo Kenya 2011, ASER Pakistan 2011, ASER India 2012		
Notes: Tests in Tanzania and Kenya were in Kiswahili, and in Uganda, English. India tests were in local languages, and Pakistan in Urdu or Sindhi. All tests used the ASER standard of being able to read a grade 2-level text as the highest level of reading comprehension, with the exception of Uganda, which used being able to read *and understand* a grade 2-level text, as distinct from just being able to read, as the highest level of competency measure. Understanding requires correctly answering comprehension questions about the text. Table above reflects Uganda results for comprehension. For division, Uganda reports "ability to solve at least two numerical written down division sums of P2 difficulty." Uwezo Tanzania did not test division.		

Asking the same questions of children in different years of schooling to estimate the learning profile is relatively rare. Most empirical studies compare learning achievement of students by testing a single cross-section of students (by grade or age). But cumulative attainment is just the result of a life-long learning profile as the skills at any age or grade are the result of a child's initial endowment of capabilities, early childhood nutrition and care and experiences, and his/her lifetime of learning experiences, both in and out of school. Therefore low attainment of children who have completed schooling is just the cumulative result of a shallow learning profile.

Figure 3 illustrates such progression by comparing the (estimated) entire distribution of student performance in a typical OECD country, Denmark, with a typically scoring developing country, Indonesia. Seeing the entire distribution of student performance shows two main points.

- The difference between the "average" student in the two countries is around 100 points, which is massive and equates to 1.25 student standard deviations in Indonesian test scores.
- This difference in the entire distribution of performance implies that *the majority* of students in Indonesia are below a low threshold of learning. PISA established five "levels" of learning, and in reading, 58 percent of Indonesian students are below the most basic level, compared to only 16 percent of Danish students. The educational quality problem is not a problem of marginalized or a subset of students in Indonesia – it is a problem of the *majority*.

Figure 3: There are massive differences in learning between typical developed and developing countries: a majority of Indonesian students are below an international benchmark in reading
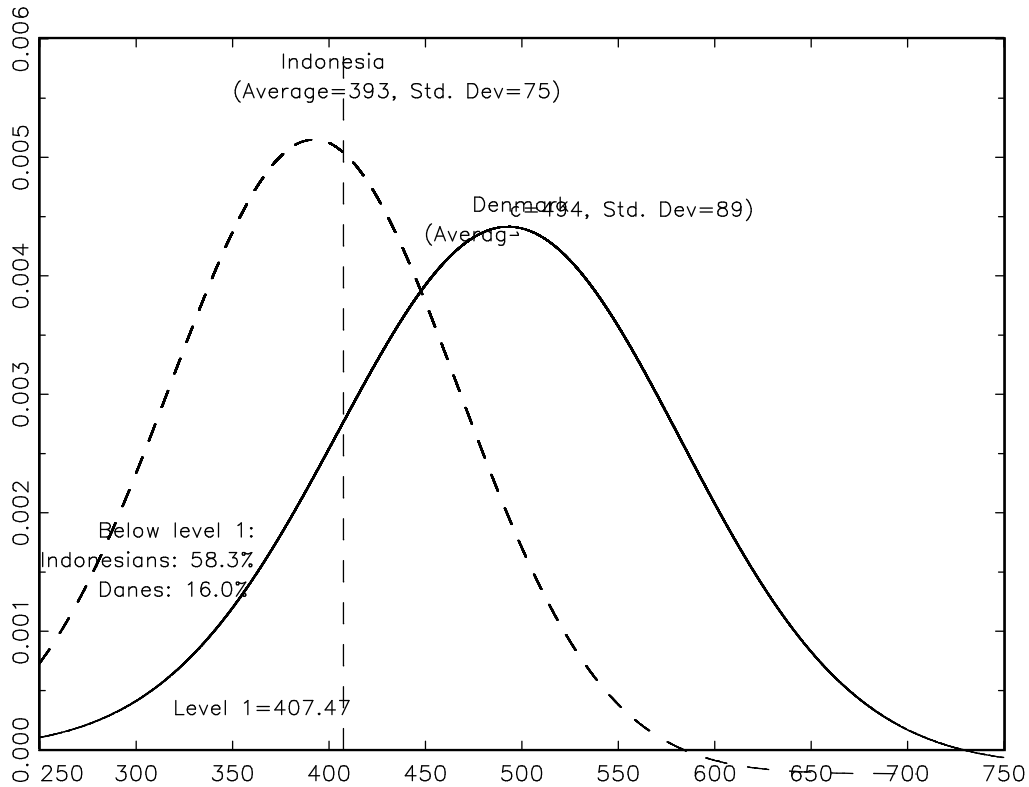
Source: PISA 2006 country means, standard deviations, and fractions below levels I are from http://pisacountry.acer.edu.au. Graphs are generated by simulating a normal distribution of results from the reported means and standard deviations.

When two countries end up with very different performances at age 15 it is either (a) due to a difference in the initial distribution of knowledge/skill; (b) due to the learning profile of a cohort; or (c) both. Figure 3 just shows that the cross-national comparison at a point in time (age 15 or grade 8) is a slice through an ongoing learning dynamic for each student. In Figure 4 we depict scenario (b): Indonesian and Danish students started out with the same initial knowledge or skill, but Indonesian students diverged to end up with much lower assessed performance in grade 8 (or age 15) due to a flatter learning profile.

Figure 4: Flat learning profiles can result in the lower empirically
observed performance in cross−sectional comparisons
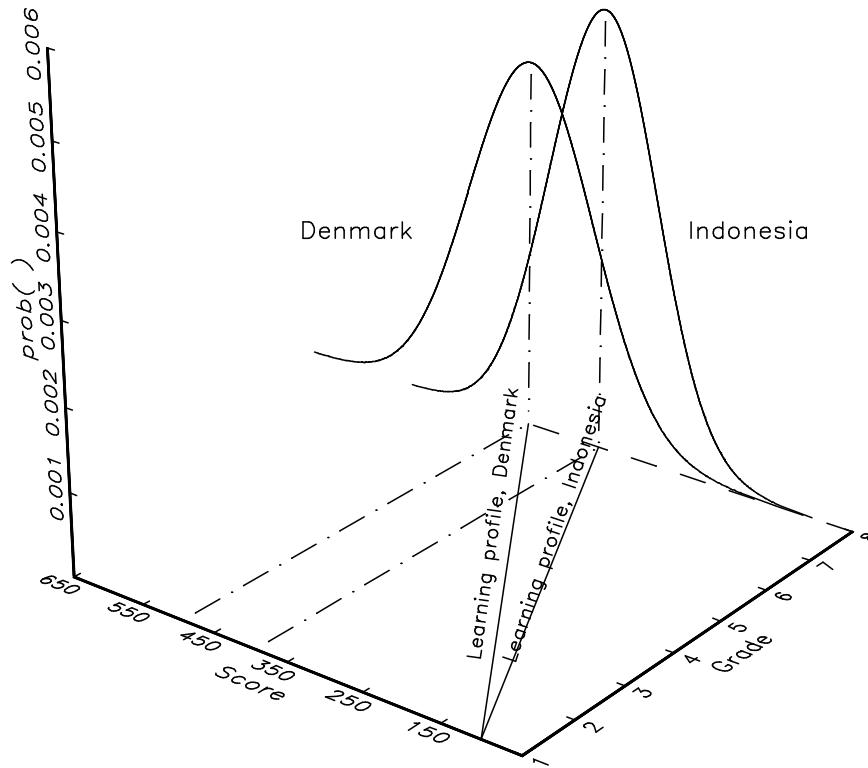
Source: Authors' simulations to fit actual distributions assuming equal starting distribution, using PISA 2006 from http://pisacountry.acer.edu.au.

Therefore the shallow learning profiles observed directly in the India studies are consistent with the very low average performance observed in countries from Latin America to South Asia to Africa to parts of Asia. Hanushek and Woessmann (2009) have combined all the available international tests into an estimate of "cognitive skills," summarized in Figure 5. The "typical" developing country (labeled "Dev") is around 400, the weaker countries are in the 300s (two standard deviations below the OECD average (by construction) of 500, and even the best developing country is well below the OECD mean and about equivalent to the worst high-performing east Asian (HPEA) country. A priority research question is why cumulative learning outcomes are so low and what can be done about it.

Figure 5: the typical developing country outcome is at a full student standard deviation behind the OECD average

Data source: Hanushek and Woessmann 2009
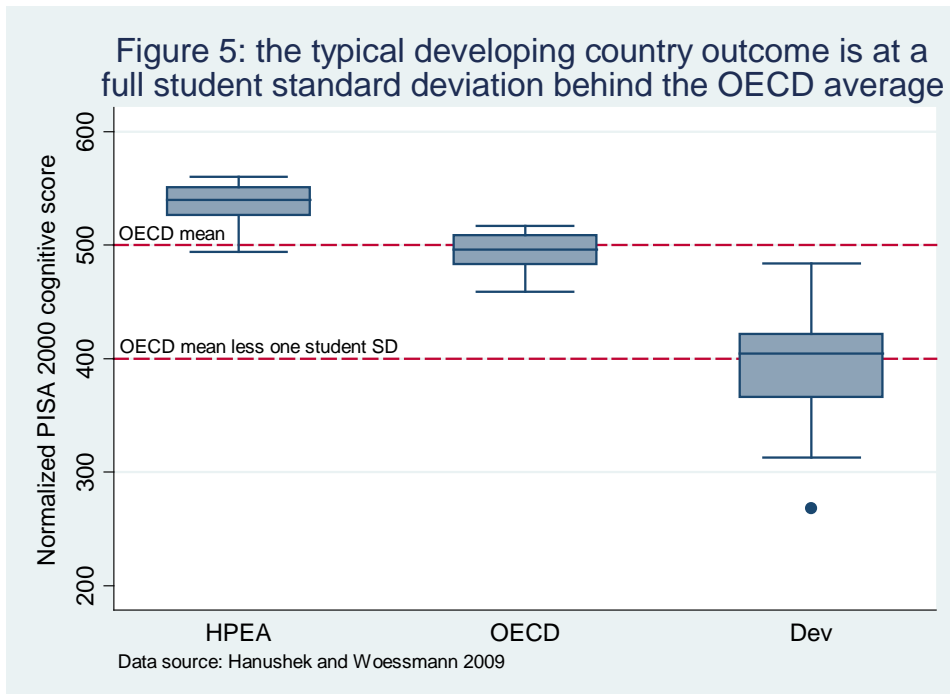
II)   The gap between the learning and curricular pace

> *"…the teacher still feels accountable for completing the prescribed syllabus…. Children who fall behind stay behind stay behind as the rest of the class moves ahead. Most primary school teachers take the least risky route of concentrating on those children who…can cope with the syllabus."*
>
> *Rukmini Banerji, 2000*

We believe we can make progress on the puzzle of why learning profiles are so flat by standing the problem as usually posed on its head. The usual question is "why are students so far behind the curriculum?" but the more telling question is "why is the curriculum so far ahead of the students?"[8]

Curriculum mismatches around the world

In many of the same countries described above with appallingly low levels of learning, we also see students lagging behind curricular expectations. For example, according to the curriculum, Pakistani students are expected to master single digit addition by kindergarten or at latest first

---

[8] Throughout the paper we use the term curriculum to broadly describe what teachers are instructed or mandated to teach, what they plan to teach, and what they actually do in the classroom. The problems we describe might not only be due to teachers following a prescribed curriculum but, also due to any combination of curricula, syllabi, lesson plans or instruction. We recognize that even if teachers are incentivized and instructed to teach an overambitious curriculum, they may not do so in practice; or that they may be guided to employ "student-centered learning" yet still adhere to the guidelines of the system's curriculum or lesson plans. Our main point is that the level of instruction is overambitious relative to student skill, irrespective of who is driving the parameters of instruction – curriculum is just shorthand for this level of instruction.

grade. But the Learning and Educational Attainment in Punjab Schools (LEAPS) study in Pakistan found only 65 percent of third graders in Punjab, Pakistan could perform single digit addition (Das, Pandey, & Zajonc, 2006). By third grade, the Pakistani curriculum expects children to be able to add fractions, a task that requires both mastery of addition and multiplication and the more sophisticated understanding of "common denominator." Not surprisingly, only 19 percent of third grade students could add fractions. The LEAPS study reports: "typically, children are three to four grade-levels below the curriculum". This means if teachers are teaching the curriculum, they are using material "three and four grade levels" above the ability of their students.

Similar results were found in India. Even in one of India's highest achieving states, Himachal Pradesh, in reading, nearly only a third of students starting grade 3 were able to perform what was expected of them by the end of grade 1. Rajasthan is one of the lower performing states in a five-state study by ASER, yet the expectations of the curriculum are even higher than in Himachal Pradesh. In Rajasthan, the curriculum specifies that in grade 1, children are supposed to be able to read a chapter written in story form, yet by the end of grade 3, less than 40 percent of students were able to read even simple two letter words (Bhattacharjea, Wadhwa, & Banerji, 2011).

Table 3 compares actual achievement with what students are *expected* to learn in the curriculum. The Indian government has set "minimum learning levels" (MLLs) in each subject (and this is supposedly a "minimum" standard). In mathematics, there are five learning areas: 1) whole numbers and numerals, 2) adding, subtracting, multiplying and dividing whole numbers, 3) simple problems of daily life relating to units of money, length, mass (weight), capacity, area and time, 4) fractions, decimals and percentage, and 5) geometrical shapes and spatial relationships.[9] Each MLL competency area has multiple items by grade level. In grade 5 there are 14 separate items in the category "ability to use fractions, decimals and percentage."

The gap between the actual student abilities in rural Andhra Pradesh and the "minimal" expectations of the curriculum could not be more stark. Less than half of children can say that the figure is $1/4^{th}$ shaded, only a quarter can multiply a simple fraction by a whole number, and less than 1 in 10 can understand a sentence with a fraction in it. Yet the curriculum expects, at a "minimum," that children do percentages, multiply and divide fractions, do problems with length, weight, etc, and. using decimals. And these are only the 14 items expected in *one* learning area.

---

[9] See http://www.education.nic.in/cd50years/r/2S/99/2S990401.htm

Table 3: Very few grade 5 students in Andhra Pradesh, India can even approach mastering what is expected of them in the curriculum area *ability to use fractions, decimals and percentage*

What fifth graders in Andhra Pradesh know		What the "minimum" levels of learning say children in grade 5 should know about fractions, percent and decimals
Test question	**Percent correct**	
What part of the figure is shaded?	48	4.5.1. (e.g. area 4 of Mathematics, grade 5, item 1) Arranges simple proper fractions in ascending or descending sequence with denominators not exceeding 10 4.5.2. Reduces simple fractions to lowest terms 4.5.3. Adds and subtracts fractions and mixed
Which figure has 1/5 apples	38	numbers with denominator not exceeding 10 4.5.4. Solves daily life problems involving comparing, addition and subtraction of fractions and mixed numbers with denominator not exceeding 10 4.5.5. Adds and subtracts mentally in daily life
Which is larger: $2/3^{rd}$ of 4500 or $1/6^{th}$ of 6000?	25	problems some combinations of fractions which occur frequently (e.g. (1/2+ 1/4= 3/4) 4.5.6. Multiplies and divides 2 fractions with
Rama ate $1/3^{rd}$ and Rita ate $1/4^{th}$ of a piece of a chocolate. Who ate the larger piece of chocolate?	9	denominators up to 10 and express the answer in its lowest terms 4.5.7 Adds and subtracts decimals up to three decimal places 4.5.8 Expresses units of length, weight and capacity in decimals up to 3 decimal places. 4.5.9. Multiplies and divides a decimal number up to 3 decimal places by a single digit number 4.5.10 Solves daily life problems involving length, weight, capacity, etc. involving comparing, addition, subtraction, multiplication, and division of decimals up to 3 places. 4.5.11 Converts fractions and decimals into percentage and percentage into fraction in lowest terms and decimal 4.5.12 Finds required percentage of a given number or measure 4.5.13 Solves simple daily life problems involving application of percentage 4.5.14 Converts mentally frequently used percentages into fractions and vice versa (e.g. 50%=1/2, 1/4=25%, etc.)
Sources: APRESt data and http://www.education.nic.in/cd50years/r/2S/99/2S990404.htm		

We are not the first to notice the negative learning consequences of overambitious curriculum. Banerji (2000) explains: "The government has an 'all pass' policy in primary school; children who have attended school regularly are automatically promoted to the next grade, regardless the level of learning.…As a result, the number of children who need extra help and attention to stay at grade level continues to grow as the children move through primary school…Without serious and committed remedial help at school, the children who begin to lag behind just get further behind and can never really catch up." As we will highlight in section III, the NGO Pratham in India for instance has designed interventions like Read India to remediate these deficiencies.

Education experts consistently point out that many countries curriculum are wildly overambitious in terms of both coverage and pace (Piper, 2009; Crouch & Korda, 2009; Korda & Piper , 2011; World Bank, 2005; Bhattacharjea, Wadhwa, & Banerji, 2011; Dubeck, Jukes, & Okello, 2012) .

If the official school curriculum covers too much, goes too fast and is too hard compared to the initial skill of the students and the ability of the schools to teach this can produce disastrous results. An overambitious curriculum causes more and more students get left behind early and stay behind forever. If children do not acquire reading and writing skills early, then textually based teaching in higher grades is pointless.[10] If children don't acquire simple arithmetic concepts—like place and common denominator—then more sophisticated operations like adding fractions is impossible. If children don't acquire basic reasoning skills—like filling in a word to complete a meaningful sentence--then asking for creativity or critical analysis later is impossible. This is how it can be that children enter—and leave—third grade without being able to read or to do addition: the curriculum has moved on.

Modeling the mismatch

In this section we construct a simple simulation model of student learning, instruction and grade progression, illustrating the consequences of a gap between learning potential and curricular pace, which we call the *curricular gap*. This simulation model shows that a divide between the pace of learning—the increment to student skills from year to year – and the curricular pace can, *in and of itself*, cause a flat learning profile. That is, we use a simple simulation model to illustrate that two countries, both of which start with exactly the same initial student ability and exactly the same *potential* learning can end up with completely different cumulative learning achievement—and that it is the country that tries to go too fast that ends up far behind. Our simple simulations show that curricular gaps alone can generate the entire observed cross-national variation in grade 8 observed outcomes in tests like TIMSS or PISA that ranges from 300 to 500—even under the assumption that all countries having *exactly the same* initial level of student ability and the same potential learning function.

As a method of investigation, a simulation model does not prove anything, but has four advantages. First, by building a simple artificial world we know completely and precisely how that artificial world it works, what causes what, and how—something that is never true of real world observation or experiment. Second, a simulation can tease out the magnitudes: how big can effects possibly (even plausibly) be. Third, simulations illustrate what *can* cause what.

---

[10] For example, even in the USA, a study finds that students who aren't reading proficiently by grade 3 are four times more likely than proficient readers to drop out before finishing grade 12 (Hernandez, 2011).

Finally, a simulation that can with parsimonious inputs generate a range of results, can show how seemingly disparate empirical phenomena—like large effects of remediation, early drop-out, equivalence of contract with regular teachers, positive impacts from tracking—all *could* be results of the same underlying cause.

II.A) A simple simulation model of learning and curricular pace

In order to build a simulation model of school instruction and learning, we first construct a *potential pedagogical function* (PPF). Think of any instructional experience, whether it is a tennis lesson, a piano lesson, a lecture on quantum physics, learning addition in second grade, or a film on Greek history. A student enters an instructional episode with a given skill set or capability in a set of domains. How much does a student actually learn—that is, expand his capability set--from an instructional episode?

To minimally characterize a given student's learning from an instructional episode we need four elements.

First, what is the *maximum gain* of any student, which is the *height* of the pedagogical function. The main difference between our simulation and most of the existing literature is that explicitly or implicitly most of the empirical assumes that this is the *only* relevant characteristic of a PPF. That is, they assume either (a) that all children who share an instructional experience (e.g. in the same school, in the same classroom, exposed to the same teacher) learn the same amount or (b) that child learning is specific to the child due to some underlying ability to learn like a general measure of "intelligence." This is illustrated in Figure 6a.

However, common sense tells us that a one parameter (height only, like in Figure 6a) PPF cannot be right. Instruction can be either too easy or too hard. Take a lecture on quantum physics. If the lecture is centered on cutting edge research, then a graduate student might benefit, an undergraduate in physics might benefit a small amount, and we (the authors) would not get anything out of it at all, lacking the background in math or physics. In contrast, a lecture geared to a general audience would benefit us, but benefit a graduate student less and bore a professor. More prosaically think of piano lessons. Instruction on reading notes (Every Good Boy Does Fine and all that) benefits the beginner but not the virtuoso and a master class on the appropriate dynamics of the Moonlight Sonata refines the virtuoso but hardly benefits a child who cannot read the notes.[11]

The second feature of a PPF is the *shape*. Does learning fall off symmetrically and linearly from a unique maximum? The PPF would be triangular (Figure 6b). Does every child learn the same within a range but outside that range children learning nothing? The PPF would be rectangular (Figure 6c). Doe learning decline exponentially from the maximum? The PPF could have a "normal" distribution (Figure 6d). There is no reason the shape need be symmetric and learning could decline more sharply for those above than below the maximum (Figure 6e).
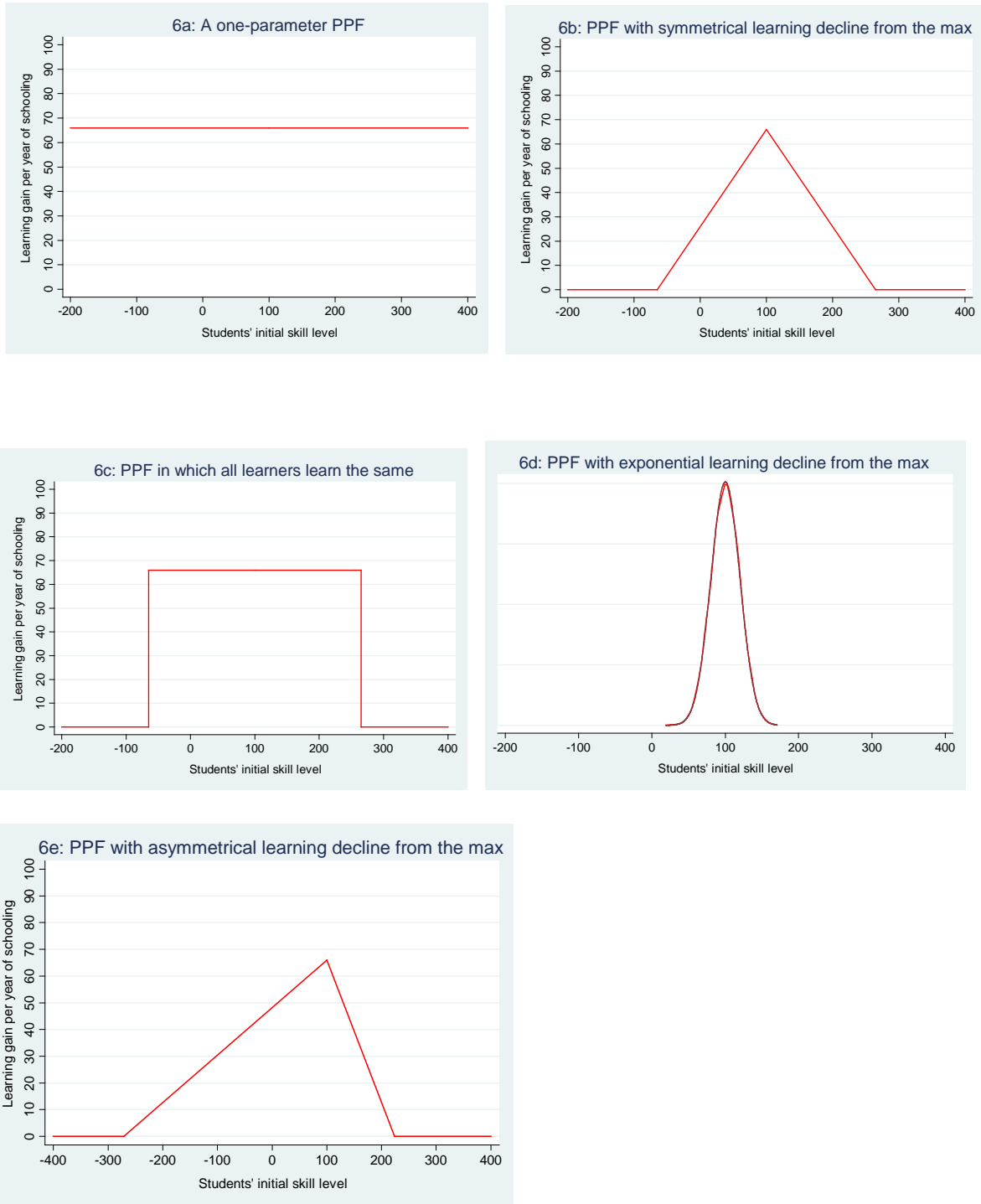
Third is the *range* of the pedagogical function (determined by the height and shape) which is the extent of the student skill over which the student learns, or the spectrum of skills over which episodes are applicable or relevant and allow for learning. Below we also refer to this as the *width*.

---

[11] Or benefit, much to our mothers' disappointment, the authors.

Fourth, with a "height only" PPF (as in Figure 6a) location is irrelevant, but with a PPF with a shape the *location* matters. To what level of student ability or capability is the instruction pitched. Particularly for any symmetric representation we can refer to the level of student ability on which the PPF is *centered*.

To be fair, most educational researchers did not assume that a "height only" PPF really had infinite extent. Rather they assumed (implicitly usually) that "height only" was a useful approximation because (a) instruction was centered, (b) the shape was quite flat and/or (c) the range of the PPF was large relative to the distribution of student initial capability. The very design of schooling into grades with a sequenced curriculum is intended to reproduce those conditions to achieve cost efficiency so that students can be taught in large(ish) groups. We illustrate the consequences for learning if instruction is not centered on actual student ability and the range of student ability is large.

Figure 6: Illustrations of possible shapes of potential pedagogical functions (PPFs) demonstrate pace and range of student learning

6a: A one-parameter PPF

6b: PPF with symmetrical learning decline from the max

6c: PPF in which all learners learn the same

6d: PPF with exponential learning decline from the max

6e: PPF with asymmetrical learning decline from the max

Source: Authors' simulations

In our subsequent simulations, we use symmetric triangular potential pedagogical functions because this makes the math of the simulations easy (learning declines linearly away from the

maximum) and makes the graphs easier to interpret (because PPFs will be triangles and distributions of student learning will start out as a "normal" distribution). We are not asserting the pedagogical functions *are* in fact triangular and nothing about the main findings of our simulations hinges on using this particular shape of the PPF.

Actual student learning is the interaction of the PPF with the initial distribution of student skill. The learning of any particular student "i" of initial skill level "s^i" is expressed in a simple piece-wise linear equation that is a function of three parameters of the triangular PPF: "c" (the center in the skills dimension), "m" (the maximum of the PPF, or in this case the height of the triangle, which is at c), and "r" the range.

$$Eq. 1)\ L\left(s^i; PPF^N(max = h, shape = ST, range = r, center = c)\right) = \begin{cases} = 0 \ if \ s^i < c - \dfrac{r}{2} \\ = 0 \ if \ s^i > c + \dfrac{r}{2} \\ = h - \dfrac{h}{\left(\frac{r}{2}\right)} * |s^i - c| \end{cases}$$

This implies that *average* learning of any group of students from exposure to the same PPF depends on the *gap* between the center of the PPF and the center of the distribution of student skill. This is illustrated in figures 7a and 7b, and in table 4.

Figure 7a shows the interaction of normal distribution of student skills and a triangular PPF centered on the mean of student skills of 340. We have calibrated the PPF and initial distribution in PISA/TIMSS units of OECD mean 500 and student standard deviation 100, so that centered curriculum reproduces (roughly) the PISA/TIMSS means and standard deviations.

When the PPF(66,ST,330,340) is centered on a assumed student ability of N(340,20) then children across the range of abilities all learn, but somewhat different amounts. (340 is roughly the mean for grade 5 students, using the PISA/TIMSS calibration discussed above.) Students two standard deviations above or below the mean, student A at 300 or student E at 380, learn 50—so that after the year of instruction they are at 350 and 430. The student at the mean learns 66 and so will emerge at 406.

Table 4: Shifting the center of instruction (the skill level of the student who learns the most) away from the center of the actual distribution of student skill lowers learning gains, particularly for lagging students

Student	Hypothetical distribution of initial student ability, mean 340, std. dev. 20	Learning gain when instruction is centered on actual student mean (340)	Learning gain when instruction is centered on higher level than actual student average (460)
A (average less two std. dev)	300	50	2
B (average less one std. dev)	320	58	10
C (average)	340	66	18
D (average plus one std. dev.)	360	58	26
E (average plus two std. dev.)	380	50	34

Figure 7b illustrates that students with *exactly the same* initial ability, exposed to an overambitious curriculum centered on a student of skill 460--PPF(66,ST,330,460), will learn less as the material is too hard. Student A now learns next to nothing at all (2 units) and even the average student, C, learns only 18 units, compared to the potential of 66 if instruction were centered on her ability.

Figure 7: Identical distributions of student skill can result in very different learning outcomes, depending on where the level of instruction is centered
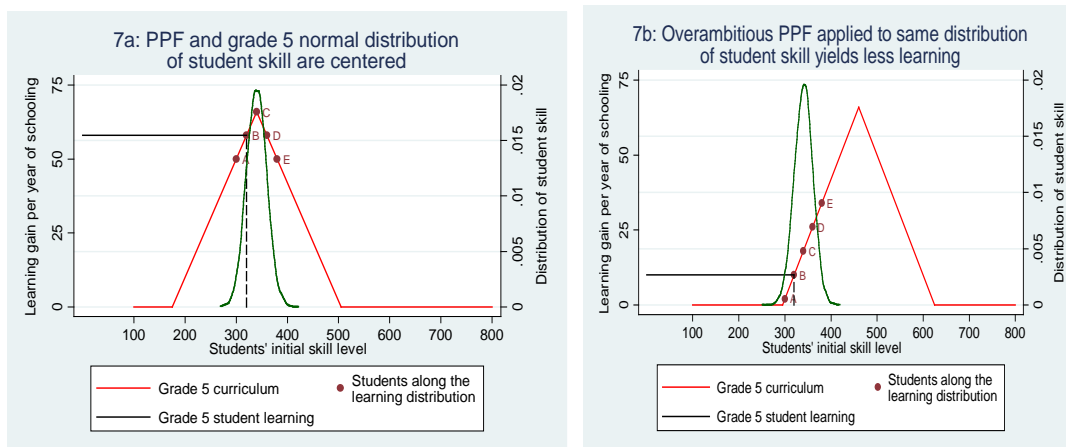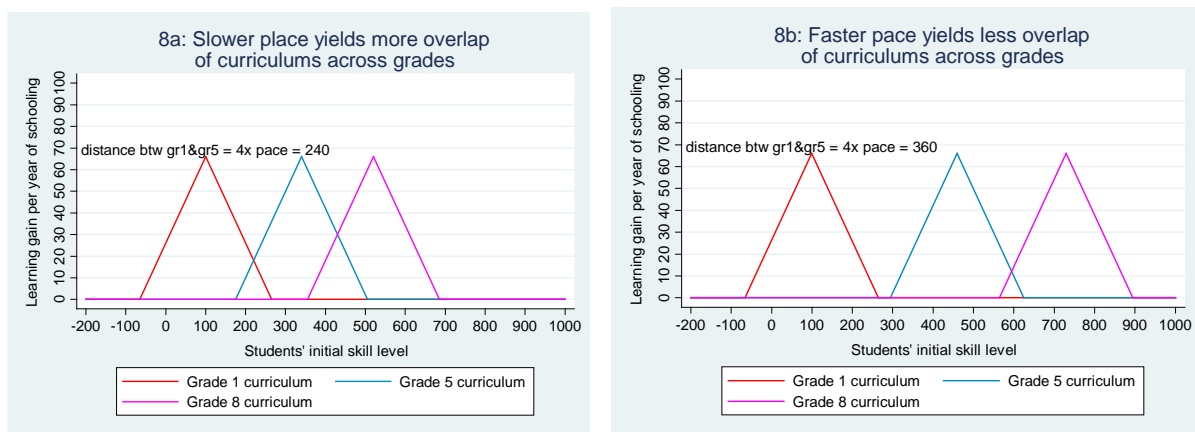

Source: Authors' imagination

Schooling is a sequence of instructional episodes. Let us treat a "year of schooling" as instructional episode on a single dimension of "skill." A school curriculum specifies what is to be taught in various grades. This implies a pace with which the *curriculum* moves through the material—how well students are expected to read and write, what arithmetic or mathematical

skills they should conceptually master, what they should understand about the natural world, how their critical thinking skills should progress.[12] (Curricular goals also include other dimensions like values, practices, attitudes which are also an integral part of schooling goals.)

A faster curricular pace means the PPFs for various grades move more rapidly ahead in the skill levels. Figure 8 shows two scenarios with two different curricular paces. Scenario 8a demonstrates that with a pace of 60, the 'distance' between the grade 1 and grade 5 is 240 points, meaning that the curriculum expects to shift learning by 240 points between these grades; whereas 8b has a pace of 90 and 360 point difference between grade 1 and grade 5.

Figure 8: Different curricula have different levels of expected learning or paces from grade to grade

Source: Authors' imagination

Here is the first and essentially only tricky thing to understand in this paper. The *curricular pace* is a *policy determined* shift in the skill around which the official instruction is centered. Graphically this is the *horizontal displacement* of the PPF from year to year. The *interaction* of the PPF and initial student skill describes how much the skill of each student *actually moves* from year to year. This is the horizontal displacement of actual student skill, or, in a graph with grades on the horizontal axis and learning on the vertical this is the learning profile.

Because the curricular pace is a *policy* decision there is no intrinsic connection between curricular pace and actual pace. The *curricular pace* can be slower or faster than the *PPF determined actual pace.*

II.B) Results with our simulation model of learning

Our simulation model has just these three elements:

- a potential pedagogical function (PPF) that specifies student learning per grade for each student (in our simulations, the function is symmetric and triangular),
- a curricular pace that specifies how far the PPF shifts from grade to grade,
- an initial normal distribution of student skill in grade 1.

---

[12] Throughout these simulations, we assume that students move uniformly through grades, or automatic promotion.

With those elements we can simulate the trajectory of skills of each individual student and hence, by simulating the learning trajectories of a large number of individual students, we can generate the evolution of the distribution of skill across a population of students through the entire course of basic education from grade 1 to grade 8.

Our first goal is to show the consequences in this simulated world of different choices for *curricular pace.* The parameters of the simulation are chosen such that *centered instruction—* defined as a match between the curricular pace and the actual learning pace—moves the average student performance from 100 in grade 1 to 500 in grade 8, and produces a standard deviation in grade 8 equal to those empirically observed for a typical OECD country, say Denmark. The PPF in Figure 9a with maximum height of 66 points gained per year, total width/range of 330 (five times height) and triangular shape, generates an average gain in grade 1 of 60 points a year, when initial student skill is normally distributed around a mean of 100 with a standard deviation of 20. With this PPF and a curricular pace of 60 (equal to the first year's actual gain), this simulation generates a grade 8 learning profile similar to an OECD country.[13] The equation for the PPF for grade N with height, shape, with, and center is:

$$Eq. 2)\ PPF^N = PPF(66, ST, 330, 100 + (N-1)*CP)$$

The results displayed in Figure 9b are based on exactly the same simulation with just one change—the curricular pace has *increased* to 100. That is, the center of the PPF now moves ahead in the skill dimension by 100 points each year. One might think that making the curriculum more ambitious could only lead to good things, but in this simulation it is disastrous. The *average* score of students in grade 8 *falls* from 500 to only 247. Exactly the same PPF that generated average scores of 500 also generates scores of grade 8 students below 300—only because the *curricular pace* increased.

---

[13] The principal limitation of the simple simulation with a symmetric triangular PPF is that it generates a "standard deviation" that is similar to Denmark's but not the "normal" looking distribution empirically observed. Rather, given the mechanics of the learning specified in which more advanced students regress to the mean, the resulting distribution is concentrated near the average, but with a long left-tail of students who are gradually left behind. The result is a very "non-normal" shape as there is skewness with a long-left tail. We think we can fix this without sacrificing other features of the simulation (say, by using a asymmetric triangle or introducing "student ability") but have not yet done so.

Figure 9: A curricular pace that moves much faster than actual learning possible with the available PPF results in much lower learning
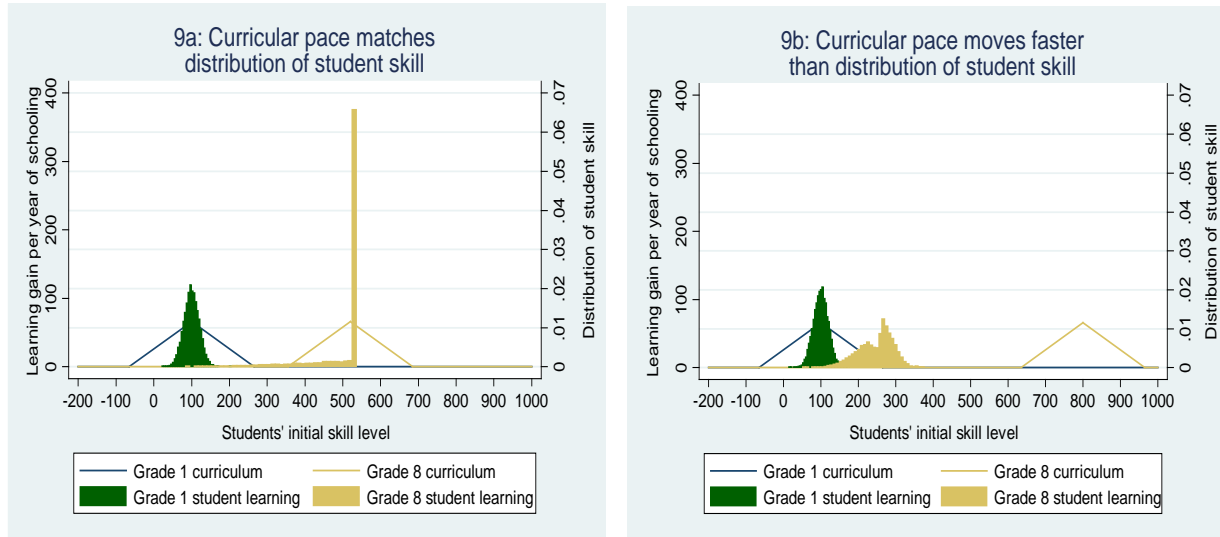
Source: Authors' simulations

The mechanics of this are simple. If the PPF shifts from grade to grade by more than the actual distribution of student learning from year to year, then more and more students are left behind because they learn little or nothing in a given year. In the next year, given that they are even further behind, from having learned so little, they learn less (or nothing). If the curricular pace is faster still, then even the *average* student gets left behind. When the pace is very fast compared to the PPF then *everyone* eventually gets left behind.

We can simulate learning grade to grade with exactly the same PPF and exactly the same distribution of student initial skill —with the only change being curricular pace of 60, 70, 80 or 90.

Figure 10 shows the results of the fraction of students *not learning anything* by grade. At curricular pace of 60, almost no one stops learning before grade 8. But if the curricular pace increases to only 70, 20 percent of students learn nothing by grade 8, as they have fallen outside of the range of the PPF. Further increases in the curricular pace cause increasingly rapid rises in the fraction of students not learning; at CP=80, only 20 percent are learning anything in grade 8, and at CP=90 essentially no one learns anything past grade 6.

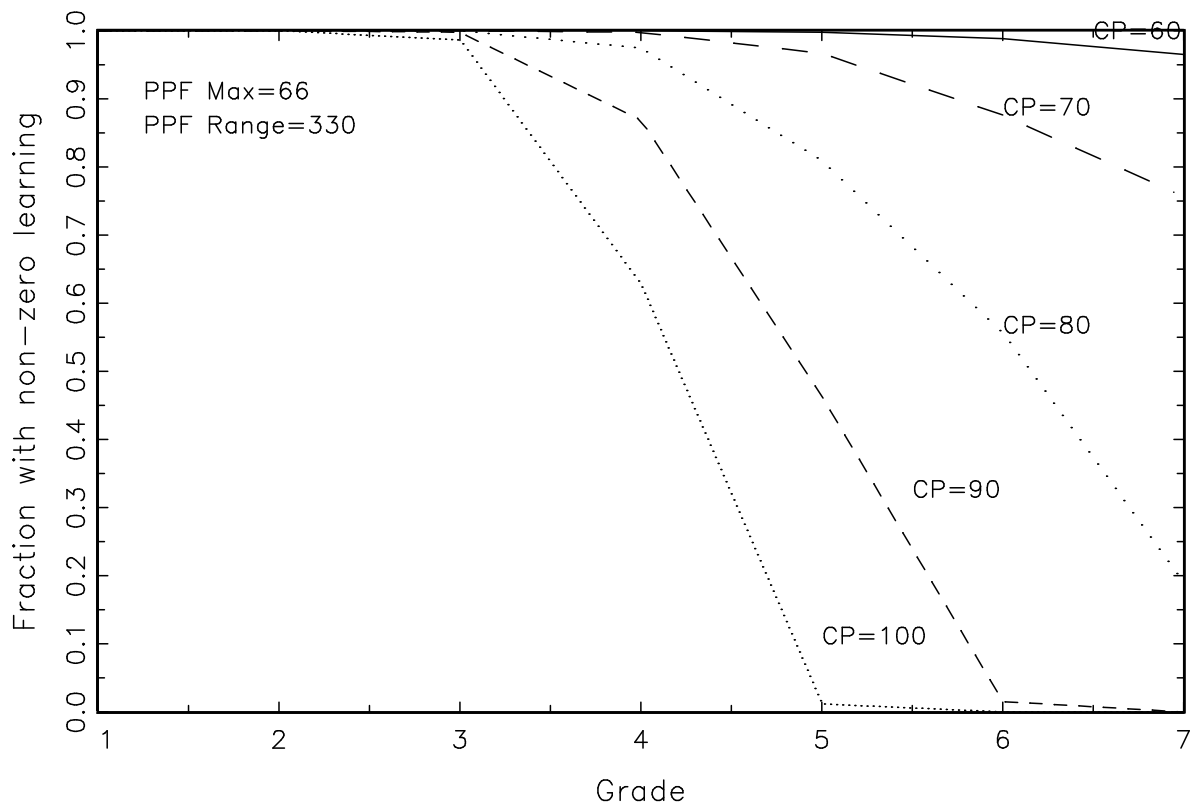Figure 10: Fraction of kids learning anything falls as curricular pace (CP) increases even with the same PPF

Figure 11 shows the resulting learning profiles from these same simulations. By construction, CP=60 produces a fairly smooth pace of actual learning and in grade 8 the average score is 500.[14] If the curriculum is more ambitious and CP=80, then average learning in grade 8 falls to 338—intuitively because more and more students are further and further from where the PPF says the most learning happens. By the time CP=90, the learning profile is essentially flat from fifth grade on as the curriculum has moved so far beyond student skill.

---

[14] The average learning per year actually falls gradually because the standard deviation of student learning increases (as is realistic) and since actual learning is a weighted average of the PPF where the weights are the student population this means more of the triangle has students in it (that is, the typical student is further from the mean—which is what "standard deviation" means) and hence the average of the same PPF falls.

Figure 11: The Learning Profile Flattens as Curricular Pace Increases
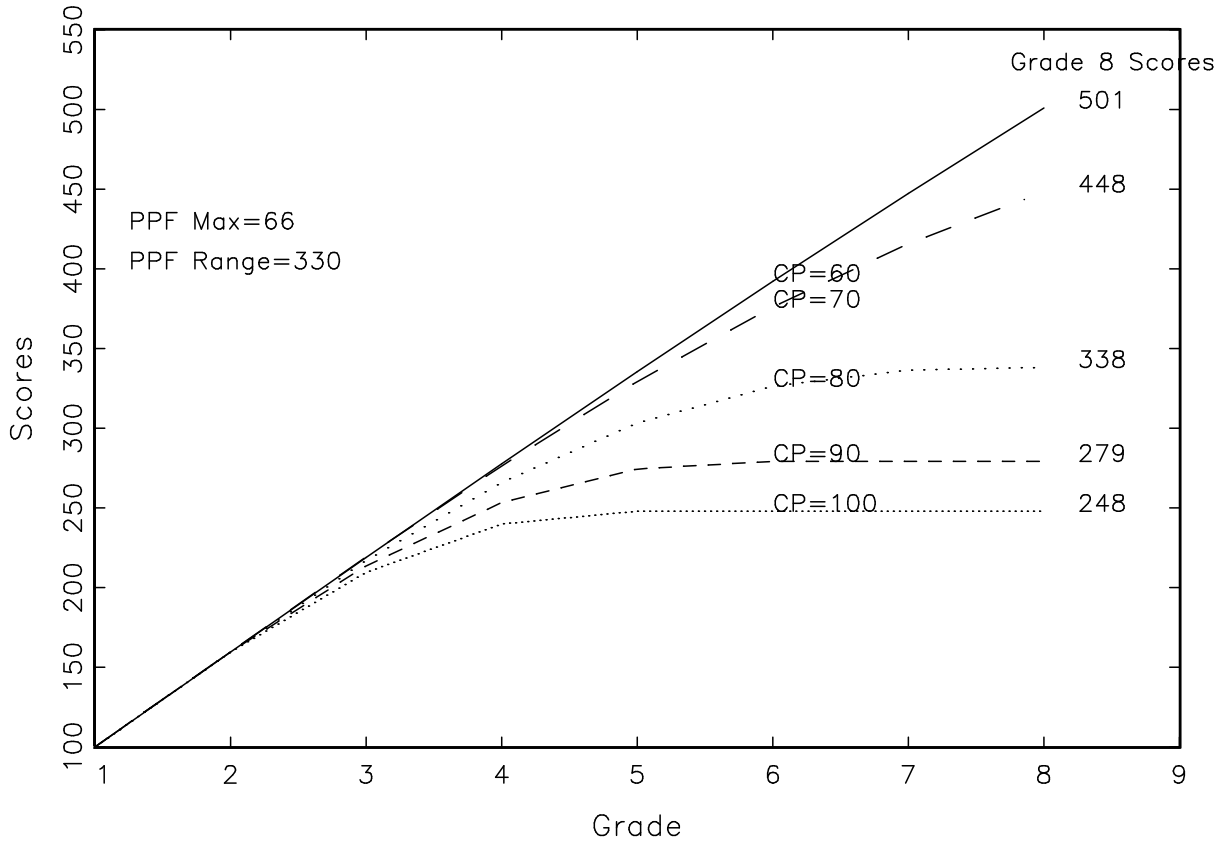
All of the elements of the simulation are illustrated in Figure 12, which is a complex graph to understand, but the investment is worthwhile, as understanding Figure 12 will change your life.

The PPF at the overambitious curricular pace=90 is illustrated for grades 5 and 8. This is a triangle with height 66 (maximum learning per year) and width of 330 that is centered on 460 (average skill) in grade 5 (460=100+90*(5-1)), and centered on 727 in grade 8 (730=100+90*(8-1)), so $PPF^5(66,ST,330,458)$ and $PPF^8(66,ST,330,727)$. The line of curricular pace when CP=90 in Figure 12 shows the evolution of the center of the PPF^N across grades.

The entire distribution of actual student skill that results from the sequence of learning in each grade from the PPFs advancing at CP=90 is shown for grades 5 and 8. The actual learning profile (ALP) of the average is just the line plotting the means of those distributions (this is labeled ALP(60,90) as it is the learning profile when the "achievable" learning in grade 1 is 60 with a centered PPF but the curricular pace is 90). This is the same learning profile as that shown in two dimensions in Figure 11, flattening out around grade 5 and leading to a grade 8 score of only 279.[15]

---

[15] Also, figure 9 is just a two dimensional compression, a projection onto the y-z plane (or, Figure 9 is what you would see from a distance looking in the x direction standing on the y-axis).

Figure 12: A curricular pace more rapid than potential learning causes an actual learning profile (ALP)flatter than the potential (PLP)
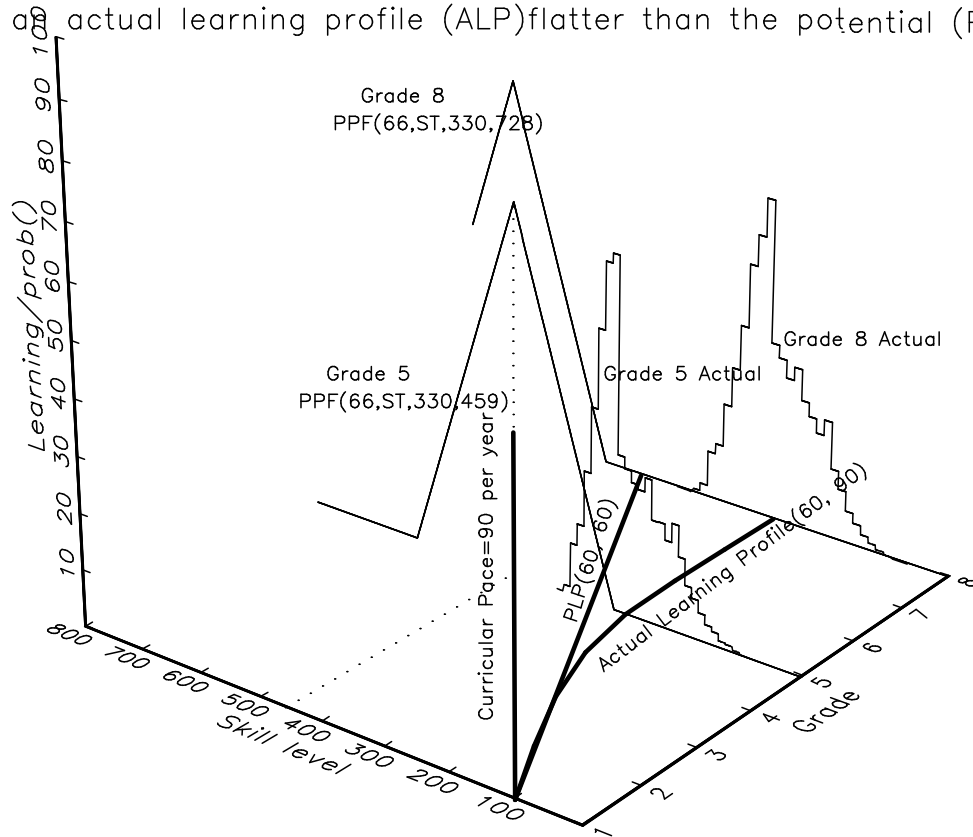
The advantage of the showing the z dimension is that we can also show the probability distribution of student skill levels.[16] Comparing the PPF with the distribution of student learning, we can see that more than half of students in grade 5 are at a skill level where the learning from the PPF is zero, as looking at the distribution of grade 5 the peak is roughly over where the PPF (the triangle) hits zero. This is the same fraction not learning as shown in Figure 10 for the line CP=90 at grade 5. By grade 8, the PPF (the triangle) and the actual distribution of student skill (the histogram) do not overlap, meaning no one learns anything and the learning profile is therefore completely flat.

Figure 12 also shows the potential learning profile (PLP) if the PPFs are more centered, which is the actual learning profile when the curricular pace=60 (so PPF^N(66,ST,330,100+60*(N-1))) and is labeled PLP(60,60). This learning profile reaches a score of 500 in grade 8. *All* of the deviation between the potential learning profile and the actual learning profile—and hence all of the gap between students who reach 500 on average and students who only reach 279 in grade 8 - -is because of the accelerated pace of the curriculum.

---

[16] Since we cannot have two z-axes, the shapes of the probability distributions are re-scaled to the z-axis by making their maximum point of the assumed normal distribution in grade 1 equal to the height of the PPF. The actual distribution shown does not have the shape of a normal distribution, as the repeated multiplicative application of a triangular learning function to a initially normal distribution cannot mathematically produce a normal distribution.

II.C) Other features of the simulation

We have shown that a simple simulation model of accelerated curricular pace can generate huge differences in learning outcomes *even if* students start at exactly the same place and *even if* the *potential* of learning is exactly the same. We do not argue that these *are* exactly the same across countries. Because of cumulative nutritional insults or lack of early stimulation children or negative stimulation due to stress, children from different backgrounds could begin formal schooling less school ready. Moreover, due to lack of inputs of all types—from learning materials to poorly constructed classrooms to less knowledgeable teachers—the potential pedagogical function could be lower in some countries than others. These differences exist and interact with curricular pace in ways that reinforce each other. Two quick illustrations.

Low(er) initial skill combined with accelerated pace: first, if the students arrive at first grade with, on average, a lower level of skill, which in this case could be called "school readiness," then the problems of an overambitious curriculum are exacerbated as more students are left behind, sooner. Table 5 shows the results of doing exactly the same simulations as above, changing only the initial skill or school preparedness level from 100 down to 80 or to 60. Even using the same curricular place of 60 (PPF height of 66), but starting 20 or 40 initial skill points behind 100, results in the simulations in a grade 8 cumulative learning 80 points or over 200 points lower. The problem of accelerated pace is exacerbated earlier—so that even small increases in the curricular pace cause large falls in cumulated learning outcomes. For example, for students that start with skill levels of 60, even just shifting the curricular pace from 60 to 70 causes a 32 percent drop in grade 8 scores. So the *combined* effect of lower initial skill levels with an inappropriately accelerated curriculum is much larger than just the low skill level alone.

Table 5: Students lack of "school readiness" exacerbates the problems of overambitious curricula as even small increases in curricular pace produce large falls in learning

Curricular Pace (CP)	Initial average skill level of grade 1 students					
	100 (base case)		80		60	
	Grade 8 learning achievement	Percentage fall due to raising CP	Grade 8 learning achievement	Percentage fall due to raising CP	Grade 8 learning achievement	Percentage fall due to raising CP
60 (Centered instruction)	**501**		417		287	
70	447	11	328	21	217	24
80	337	25	255	22	180	17
90	279	17	218	15	160	11
100	248	11	198	9	147	8

Source: Authors' simulations using PPF^N(66,ST,330,100+CP*(N-1) and initial student skill distribution of mean 100,80, or 60, standard deviation 20.

Low(er) educational potential combined with not decelerating pace: second, countries also could have different potential pedagogical functions—not be able to teach as much per year or have fewer resources, even when aligned. Our point is that this interacts with curricular pace, so that countries with lower PPF are obviously capable only of producing lower cumulative outcomes, but maintaining an overambitious curricular pace makes those losses much worse than they need

to be. For instance, suppose the maximum learning gain per year (or the height of the triangle in the PPF) shifts downward by 10 points, from 66 to 56 (and gets correspondingly narrowed to maintain the same shape of width=5*height). If simultaneously the curriculum also slows by 10 points per year, then as shown in Table 6, the loss is roughly 70 points by grade 8 from 500 to 433. But if the curriculum does not slow then there is an additional 50 points of loss, from the achievable 433 to 381. So the *combined* effect of lower school quality without slowing curricular pace is much larger than just the lower quality.

Table 6: Losses from lower potential combines with inappropriate curricular pace to produce even greater losses in learning

Curricular Pace	Grade 8 score with height of 66	Grade 8 score with height of 56	Losses
50	483	433	Loss from just lower PPF when curricular pace is allowed to slow to 50= **67** (500-433)
60	500	381	Combined learning loss from lower PPF **and** not slowing curricular pace by staying at 60= **119** (500-381)
			Loss from not slowing the curricular pace when actual pace of learning slows=52 (433-381)

Source: Authors' simulations

III) Understanding the empirical evidence on learning

To this point we have shown that *if* the curricular pace is too fast, this *can* cause flat learning profiles and low cumulative attainment, even if all the other conditions for learning are present. This is not, in and of itself, evidence that an excessive rapid curricular pace *is in fact* the cause of low learning attainment in any given context. In this section we argue that evidence from experimental and non-experimental studies are consistent with a large curricular pace mismatch.

The best evidence that there is a large mismatch between curricula and actual student ability is that there just is. That is, whenever countries test mastery of the curriculum, the typical student fares badly, often very badly. As in the Pakistan case cited above, the typical student is "three to four years" behind the curriculum – by the end of the third grade, only a small majority had mastered the first grade mathematics and language competencies expected of them (Das, Pandey, & Zajonc, 2006). Or the Uwezo study in Tanzania found that while multiplication is a key component of the grade 2 curriculum, only eight percent of grade 2 students could multiply, and over half could not yet add. Even by grade 7, 32 percent of children could not multiply. Progress was similar in reading – of children completing grade 7, only half could read a grade 2-level story in English (Uwezo, 2011). These are just examples to illustrate that all of the evidence that students are far behind the curriculum is at the same time evidence that, if teachers are teaching the curriculum, they are teaching at a level far above actual student ability.

In this section we show that the single concept of a curricular gap, induced by an excess of curricular pace over actual learning progress, helps unify and understand a large number of recently emerging findings about learning. Our simple simulation model, that adds some realism

to a description of learning and one new "parameter"—the curricular gap—to the formalization of "education quality," can easily replicate many otherwise puzzling findings.

As emphasized above, we are not creating new evidence; rather we are using a simple formal representation of a learning function to encompass the evidence and concepts already suggested by other economists, educationists and practitioners such as Banerji (2000), Glewwe, Kremer and Moulin (2009), Banerjee and Duflo (2011), and Crouch and Korda (2009). The primary themes highlighted in their research are (a) curricula and in turn teaching are overambitious and cater to the higher-ability rather than the average students, in part because many systems were designed to cater to an elite class of students; (b) teachers' incentives often support the most advanced students since teachers are evaluated on their ability to teach the prescribed, overly-ambitious curriculum; and (c) students passed from grade to grade, despite not learning, receive little remediation to catch up, and thus the proportion of students who lag behind grows as grades progress.

III.A) "Production function" studies—providing more inputs

The "curricular gap" adds a twist to understanding the voluminous empirical work estimating the "education production functions" that estimate the relationship between measured learning outcomes and schooling inputs. The puzzle that emerges from this literature is that, even when inputs are very low and hence we would expect an expansion of inputs to produce large learning gains, many empirical studies find surprisingly small impacts of expanded inputs (e.g. (Hanushek, 1986; Glewwe, Hanushek, Humpage, & Ravina, 2011; Kremer M. R., 1995).[17] Even experimental evidence of the learning impact of the availability of new budget controlled at the school level for inputs in India finds very modest impacts (Muralidharan & Sundararaman, 2011).

How much would we expect learning to increase from an increased application of inputs (e.g., from more textbooks, better classrooms, better trained teachers, smaller class sizes, etc.)? A crude "chain rule" is that the change in learning will be the change in inputs times how much inputs raise the height of the PPF times how much the height of the PPF raises learning (equation 3).

$$Eq. 3) \ \Delta Learning = \left(\frac{\Delta Learning}{\Delta h}\right)\left(\frac{\Delta h}{\Delta Inputs}\right) \Delta Inputs$$

The first term comes directly from the PPF (expected learning gain per year); and for the symmetrical triangular PPF, the derivative of learning with respect to the height of the PPF is specific to each student (Eq. 4) as it depends on the gap between the initial skill level of the student and the center of the PPF.[18] In this special case of a symmetrical triangular PPF, the learning a student gains from a given increase in the height of the PPF declines linearly. The gap

---

[17] This literature often hinges on which impacts are "statistically significant" which is really not the issue, the real issue is the *magnitude* of the learning impacts of inputs for which "statistical significance" is only a proxy as the relationship between magnitudes and statistical significance depends on the precision with which the impacts are measured.

[18] This formula is specific to one functional form of the PPF, but the general notion that the learning gain depends on the distance is common.

between student initial skill and the center of the PPF, with a slope that depends on how steep the PPF is (how quickly the relevance of the material declines), in turn depends on the width, or range, of the triangular PPF.

$$Eq.\,4)\quad \frac{dL(h,ST,r,|s^i-c|)}{dh} = \begin{cases} = 0\ if\ |s^i-c| > \frac{r}{2} \\ = 1 - \frac{|s^i-c|}{\frac{r}{2}} \end{cases}$$

Intuitively, if $|s^i-c|=0$ then the learning gain per gain in PPF height equals one—if the learning gain increases by 10 points per year at the maximum gain due to more books, better teachers, then the student on whom the instruction is centered gains 10 points. Also intuitively students who are too far behind or too far ahead gain nothing.

Figure 13 illustrates the decline in student learning for the width used in our simulations (330). A student 100 points (an OECD grade 8 student standard deviation on a PISA/TIMSS like instrument) from the center of the PPF would learn only 40 percent as much as the student for which the location of the PPF was ideal.

Figure 13: The gain to a student from an improvement in quality depends on the student skill relative to the center of the PPF (illustrated for PPF(66,ST,330,c)
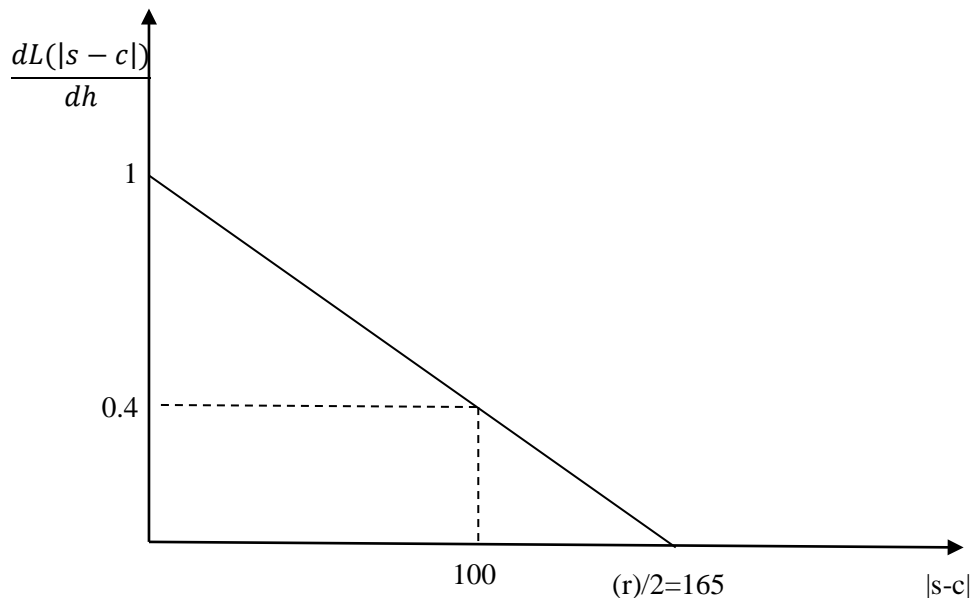
Many discussions of the "quality of education" treat "quality" exclusively as the *height* of the potential pedagogical function—what would make for greater *capacity* for learning—without the extra parameters of the shape, width/range, and centering. Considering height alone assumes that every student learns the same from a year's instruction (or that learning is a function only of student "ability"). A function that produces uniform learning per year is a is a special case, one

that would only be a useful approximation to our more complete representation of the PPF if (a) the width/range of the distribution of student ability is narrow compared to the range of the PPF, (b) the PPF is not very steep near the actual student skill and (c) the PPF is centered on student ability. However, in many settings none of those things is true.

Making explicit the idea of a "curricular gap" affects the interpretation of all of the empirical literature relating inputs to learning. One interpretation of the low association in the USA as documented first by Coleman (1966) and emphasized by Hanushek (1986, 1995)—and extended to cross-national data by Gundlach, Woessmann and Gmelin (2001) —is that at high levels of spending the marginal product of additional resources on increasing the height of the PPF was small. However, it is hard to believe the marginal product is low in poor countries, like India, Pakistan and Tanzania, where non-teaching inputs are low. Another interpretation is that the efficacy with which resources are used is low.

Kremer (2003) proposes a curricular gap as one explanation of low impact of budget resources on learning: "[g]iven the existing education system in Kenya, which like many developing countries has a curriculum focused on the strongest students, limited teacher incentives, and suboptimal budget allocation, simply providing more resources may have a limited impact on school quality." Similarly, Banerjee and Duflo (2011) observe, "[b]ecause in many developing countries, both the curriculum and the teaching are designed for the elite rather than for the regular children who attend school, attempts to improve the functioning of the schools by providing extra inputs have generally been disappointing."

Our simulation model formalizes this argument. We show a curricular gap adds an additional term to the empirically estimated marginal product in equation 2. Even if spending raises the *potential* learning is increased substantially if the PPF is not centered on actual student ability this will produce little increase learning. An obvious promising next step in this research would be to combine the usual meta-analysis of impacts of inputs (e.g. class size, budgets) with measures of the curricular gap.

III.A.1) Heterogeneous impact of inputs by initial ability

In a very influential study using RCT (randomized control trial) methods, Glewwe, Kremer and Moulin (2009) estimated the impact of providing additional standard government textbooks to grades 3-7 in English and math in 1996,[19] in 25 randomly chosen schools, with 100 schools used as controls and other interventions.[20] The empirical finding that first emerged was that textbooks had absolutely no impact on the learning of the average student. This was stunning because it put rigorous research methods squarely at odds with the widely accepted conventional wisdom.

Only gradually did the resolution of the puzzle emerge. Looking at the learning impact of additional textbooks across students' initial performance before receiving the textbooks, the researchers realized that textbooks had a statistically significant impact; but only on students in

---

[19] Notice the 13 year lag between the experiment and final publication of the academic paper, which explains how the paper has been so influential in spite of being recently published.

[20] Here we say *additional* since in grades 3, 4 and 5, 1 in 6 students had English and math textbooks. All English and math books were provided in early 1996, except for math texts for grades 4 and 6, which were provided in early 1997. In 1997, 25 more schools out of the 75 controls received grants, 43 percent of which was spent on textbooks. The other 50 controls received grants in two cohorts of 25 schools each, in 1998 and 2000.

the upper quintile of scores before the extra textbooks were provided.[21] The most persuasive explanation for this heterogeneous treatment effect was that most children had trouble even reading the textbooks, which were in English, often a third language for students in this area of western Kenya. The authors used the curricular gap to explain the results: "The finding that provision of Kenya's official government textbooks increased test scores only for above average students is arguably part of a much larger mismatch between the official curriculum and the needs of many students… In many developing countries— where students and teachers are frequently absent and scores on international standardized tests are very low—using curricula and textbooks that cover material at the same pace as in the developed world and assume that students know material from previous grade levels yields a serious mismatch."

Our formal model of uncentered instruction can generate simulated results with exactly this differential learning gain from improving the quality of inputs. Table 7 reports the results of increasing the height of the PPF in grade 4 by 10 points per year from our base case level of 66 to 76. The impact of this improvement, and on which students, depends on the curricular gap. If the CP=60, the learning of students initially in quintile I and quintile V is roughly equal. But if CP=80, *exactly the same change in the PPF* produces an effect size gain of only 0.05 for quintile I learners but 0.21 for quintile V learners.

Table 7: As curricular pace increases relative to the PPF, the impact of improving quality in upper grades is more concentrated on the better students

Curricular pace	Simulated effect size (learning gain as fraction of student standard deviation) of raising PPF in grade 4 by 10 points				
	Quintile I (worst)	Quintile II	Quintile III	Quintile IV	Quintile V (best)
60	0.26	0.35	0.34	0.34	0.33
70	0.12	0.21	0.25	0.27	0.28
80	**0.05**	0.12	0.17	0.19	**0.21**
90	0.00	0.05	0.09	0.11	0.14
100	0.00	0.00	0.03	0.05	0.08

Source: Authors' simulations, base case is $PPF^N(66, ST, 330, 100+CP*(N-1))$

Figure 14 shows how quality improvements have *differential* impacts depending on initial student skill.[22] The top panel of Figure 14 shows a massive increase in the "quality of schooling" as the height of the PPF increases by 50 points. *If* this is centered on actual student ability, the gain to quintile I and quintile V students is exactly the same, 38 points (due to the assumed symmetry of the PPF).   *Exactly the same* improvement in the "quality of education"--

---

[21] The "effect sizes" (gain as a fraction of a student standard deviation) for the students in the fourth and fifth quintiles after one year of exposure were 0.14 and 0.22 , significant at the 10 percent and 5 percent significance levels, respectively.

[22] The difference between this figure and figure 7 is that figure 14 shows the impact of quality improvements like inputs, ie raising the height of the triangle; while figure 7 shows the difference in learning when the level of instruction is centered or not. Figure 14 demonstrates that inputs don't do much in an environment in which instruction isn't centered, such as in figure 7b.

increasing the PPF by 50 points at its maximum--has different, and differentiated, effects if the curriculum is too hard for some students. We simulated grade 4 in which the pace of the curriculum was 30 points faster since grade 1 (90 per year versus 60), so the PPF is shifted right by a cumulative 90 points. In this case, there is almost no impact on the weaker students. Quintile I gains just 11 points, while quintile V students get roughly the same impact as the quintile V students in the slower pace scenario above (34 points). But in this scenario they gain less than the possible maximum because they are behind, not ahead of the teaching.

Figure 14: Exactly the same improvement in the "quality of education" has different impacts depending on how centered the curriculum is on actual student skill—with uncentered curriculum the advanced students benefit disproportionately
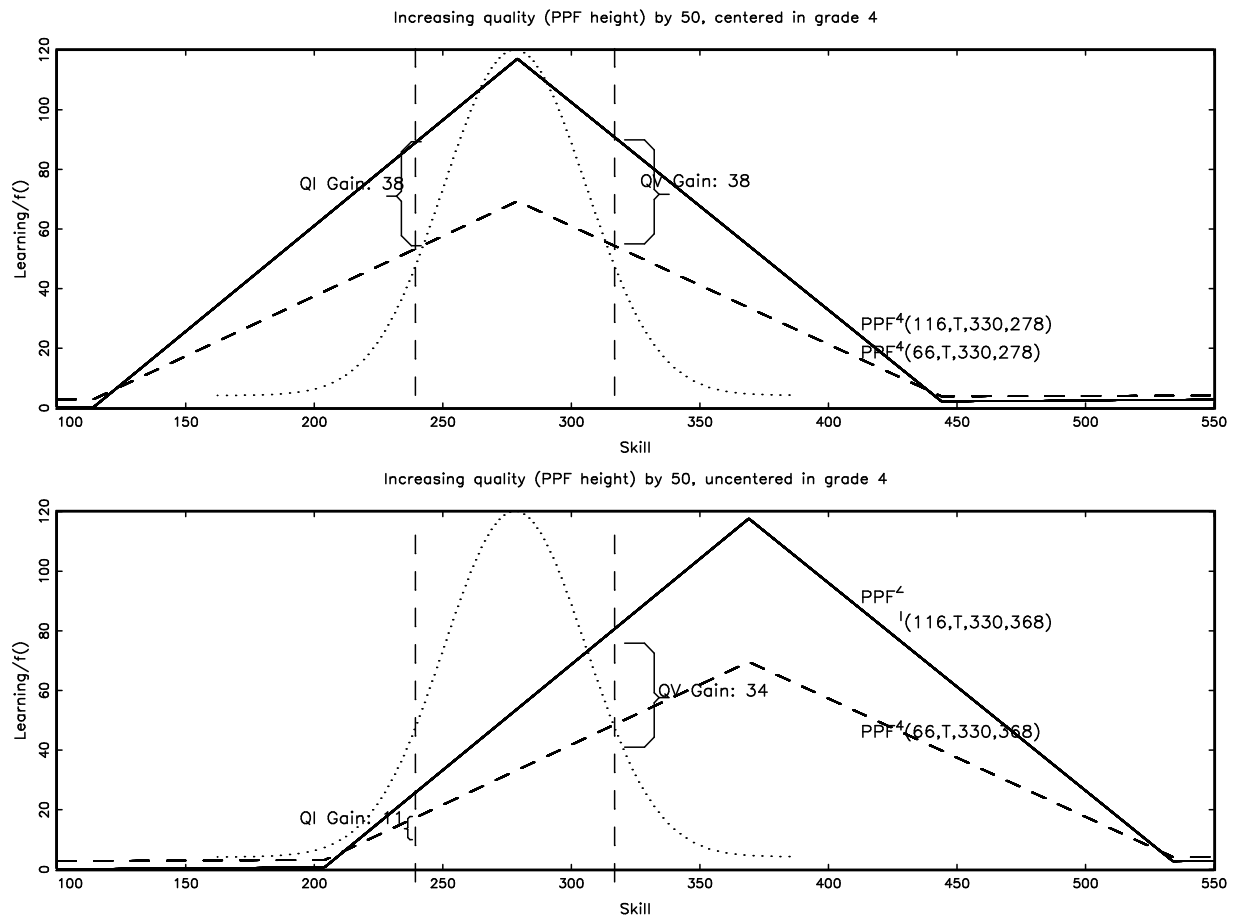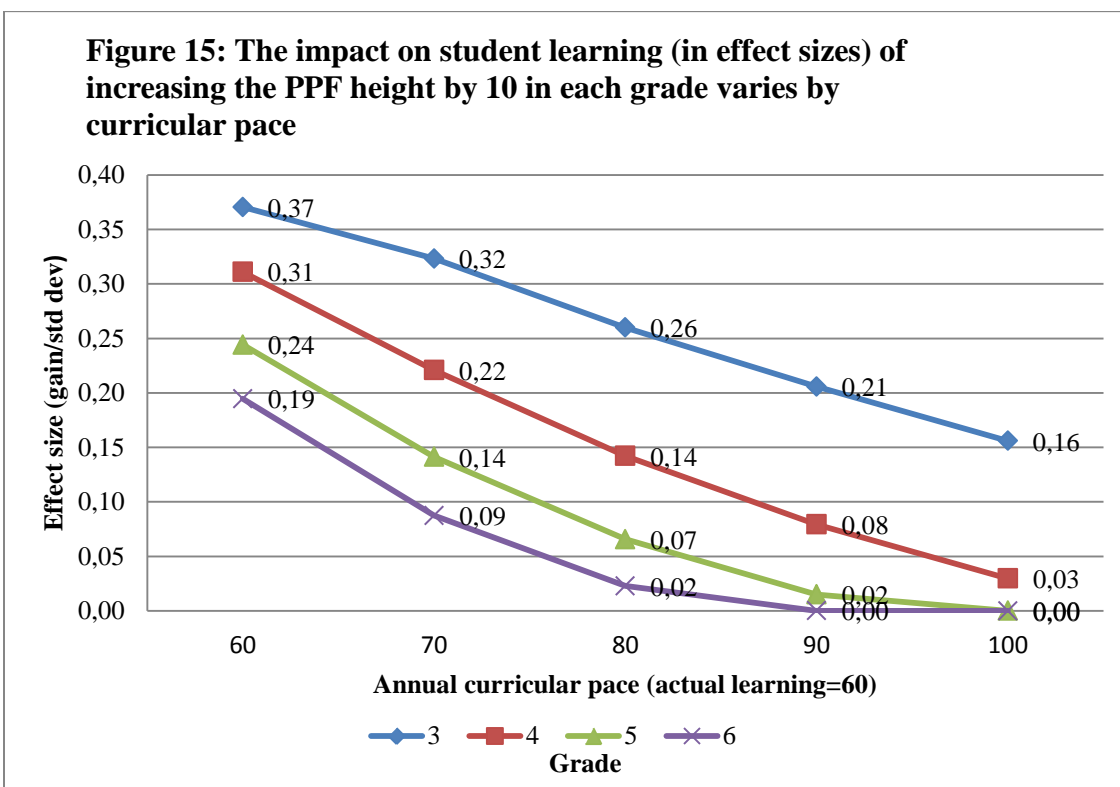

III.A.2) Impact of inputs across different contexts

The second implication of a curricular gap is that estimates of the impacts of changing the inputs—or even the PPF – will only be comparable across countries or educational settings with an equal (or similar) curricular gap. This is true irrespective of how rigorously estimated the learning impact is estimated. That is, suppose we did an unimpeachably rigorous experiment that raised inputs. Suppose we found that in context A the inputs raised scores by an effect size of 0.2 (20 points with a student standard deviation of 100). What we would know about the impact of raising inputs by exactly the same amount in context B? Even if we knew that inputs

would raise the PPF by a similar amount, we still would not know the impact in context B, unless we knew that the curricular gap in context A and context B were the same.

Figure 15 shows the results of doing *exactly the same thing*, that is, raising the PPF by 10 points (from 66 to 76) across grades and across curricular paces, which mimics any "input" increase that raised the PPF, as was done in table 5. The figure reveals two differences.

First, exactly the same experiment in grade 4 could produce an effect size of 0.31 (which is massive by the standards of the literature or an effect size of 0.14 (still pretty substantial, but only half as large) or 0.03 (which, depending on statistical power, might not even be statistically significant). All three of these estimates, although differing *by an order of magnitude* would each be the exactly correct estimate of the *causal* impact of raising the PPF by 10 points in grade 4 with the same underling PPF. But clearly extrapolating the results from a context where curricular pace matches actual learning to another where it is accelerated produces completely erroneous estimates. Estimates of the impact of interventions that raise *potential* learning have no external validity when there is variation across contexts in the curricular gap.

Figure 15: The impact on student learning (in effect sizes) of increasing the PPF height by 10 in each grade varies by curricular pace

Source: Authors' simulations

For instance, results on the impact of "class size" seem to differ substantially, even among "well-identified" estimates. For instance, there are experimental estimates of the impact of class size from the STAR experiment in Tennessee and estimates from Israel with class size effects identified off exogenous variation due to a class size rule both of which show substantial

impacts--at least on some student populations.[23] In contrast the class size effect from experiments in Kenya (Duflo, Dupas, & Kremer, 2012) and India (Banerjee, Duflo, Cole, & Linden, 2007) finds the impacts indistinguishable from zero.[24] Again, there are many potential explanations for why even the experimental estimates vary so widely since these experiments took place in a wide variety of contexts, but varying curricular gaps are one possible explanation as decreasing class size does nothing if students were not learning anything anyway. Only with independent measures of the curricular gap matched to the evidence about impacts can this be validated.

III.B) Early remediation and tracking: narrowing the curricular gap

Accounting for the difference between the pace of learning and the curricular pace, or the curricular gap, also potentially explains why non-standard approaches to learning, such as remediation and tracking, can be successful. These approaches either shift the curricular pace down a notch to better coincide with students' learning potential (as is usually the case with tracking or training teachers in remedial pedagogy), or they accelerate the pace of student learning, usually through more targeted attention or tutoring, such that students can better keep up with curricular pace. Below we highlight several examples of narrowing the curricular gap in India and Kenya. Again, we are not presenting this as "evidence" for our simulation model, rather we are showing our simple simulation model can replicate analytically these experiences.

III.B.1) Pratham experiences in (early) remediation

Recognizing that government teachers face incentives to pass students to the next grade and follow a curriculum that may be too challenging for many students (Banerji, 2000), the Indian NGO Pratham has complemented standard lessons with a variety of remedial support and community engagement projects. Pratham is not only unique in its approach to education, but also in its willingness to experiment with project variations in cooperation with researchers in ways that produce rigorous evidence about impact. Here we describe results from several Pratham interventions, offering explanations of success or failure by looking at how they affect the curricular gap. [25]

One pioneering in-school remediation initiative by Pratham involved employing a community teacher, called a *balsakhi*, to provide lessons to a group of 15-20 students in grades 3 and 4,

---

[23] This is not meant as a review of the literature on class size as there are also estimates with equally plausible identification strategies that find little or no class size effect in the USA e.g. (Hoxby, 2000) and for other OECD countries, e.g. (Woessmann & West, 2006). We are just using the positive findings to illustrate variation across countries in the estimates that are widely accepted as "well-identified." That is, some difference in the estimates of "production" function impacts could be due to econometric issues that would lead estimates to not be consistent for causal impacts, but we show that even the "true" estimates that were perfectly estimated would differ systematically with the curricular gap.

[24] The regression estimate of receiving an extra teacher on scores of children 18 months later (a year after exposure) in Kenya was an effect size of 0.029 (table 5, column 1). In the *balsakhi* experiment the implicit class size impact,, or the effect of students who were not exposed to a balsakhi was also near zero (the 95% confidence interval ranged from -0.076 to 0.189).

[25] All of these interventions were part of Read India, Pratham's leading education program that supports rural primary school children in grades 1-5 in reading and math.

identified as low achieving, for about half the school day.[26] In nearly 200 schools, *balsakhis* covered fundamental concepts taught in grades 1 and 2, such as numeracy and literacy. The program successfully lowered the curricular gap buy making it easier for students with initially low ability to be able to master key components of the curriculum. Students saw a 0.14 standard deviation improvement in test scores (averaged across language and math) in the first year and a 0.28 standard deviation improvement in the second year, with greatest gains shown by children in the bottom third of the initial distribution (0.43 standard deviations).[27] Although they declined, test scores for the bottom third of students remained higher and significant one year after the program ended (0.10 standard deviations) (Banerjee, Duflo, Cole, & Linden, 2007).

Pratham also piloted an out-of-school remediation program aimed at reading and arithmetic fundamentals, and promoting community engagement in learning. Pratham staff recruited and trained village-level volunteers to teach low-performing students reading (at "reading camps"), combined with a) disseminating information about community rights and responsibilities regarding education; and b) community training on how to assess student reading and arithmetic levels, and produce a village report card. Volunteer tutoring was the only intervention that improved learning: of children who attended the reading camps, they were 22 percent more likely to be able to read at least letters, and 23 percent more likely to be able to read at least a word or paragraph.[28] This experience demonstrates that lowering the curricular gap need not even be very challenging. Village volunteer tutors were trained for a mere four days and as a result, could offer children fundamental skills required for them to succeed in school (Banerjee, Banerji , Duflo, Glennerster, & Khemani, 2010).

In addition to adding remedial personnel such as a *balsakhi*, Pratham experimented with encouraging government teachers to use Pratham's back-to-basics approach through training, distributing pedagogy materials, a remedial summer camp, and training local volunteers to tutor students in remedial material inside and outside of school, in the states of Bihar and Uttarakhand. The in-school interventions targeted at the average teacher were intended to slow the curricular pace – to ensure that children mastered key concepts in math and reading before moving onto more advanced lessons. These had the most modest impacts. Distributing materials alone had no effect. Training teachers was less effective than training teachers combined with providing volunteer tutors, which raised reading fluency scores by 0.13 standard deviations and math scores by 0.05 points after two years in Bihar.[29] The authors suggest that the greater success in

---

[26] Some classes received the balsakhi for two years and others for one, allowing authors to see if effects differed over time. They also tested students a year after the balsakhi program was finished.

[27] Results control for baseline test scores. Note that test gains for the bottom third are likely due to the fact that the bottom third of students were most likely to be assigned to a balsakhi. When looking just at children who received a balsakhi (rather than the average effect across students), authors estimate gains to be 0.6 standard deviations. There was no gain to the higher-achieving students not receiving a balsakhi, who might have benefitted due to smaller class sizes when peers were dismissed for tutoring.

[28] These results are conditional on participating, which was not random. At the village level, children who could not read at the outset were eight percent more likely to be able to read at least letters, but no more likely to read paragraphs or stories. Children who at the outset could read letters were four percent more likely to read at least paragraphs or words and three percent more likely to read stories.

[29] Increase in 'scores' for ASER data indicates a fraction of a reading or math level. In addition to reading fluency, ASER reading scores (separate test) increased by 0.12 SD. Math results above are ASER oral. Similar results were found for written Hindi (0.14 SD) and math (0.1 SD) tests. For children who could take written tests in grades 3-5, the training+materials+volunteers intervention showed an increase of 0.12 SD for Hindi and 0.14 for math. This indicates that higher-achieving children also benefited from the interventions.

Bihar may have been because in Bihar volunteers met with children after school whereas in Uttarakhand, tutors worked in classrooms supporting teachers, where they supplemented teachers' customary work (or even substituted for teachers), which involved adhering to the standard curriculum.[30]

While the in-school interventions targeted at the average teacher did not have much traction, a similar intervention with the same type of teachers did. Although implementers faced some challenges with targeting,[31] a month-long summer camp in which government teachers[32] taught low performing children in grades 3-5 for four hours a day improved reading levels, especially for those attending. A child attending on average gained over half a level in reading (out of five levels) in just a month (Banerjee, Duflo, & Walton, 2011).

How do the Pratham experiences help explain low learning levels? There are three distinct, but non-mutually exclusive, possibilities. One explanation is that low performing students are incapable of learning; which has been consistently refuted since the Pratham experiences as a whole demonstrate that learning levels of lagging students can be improved, often quite quickly.

A second possibility is that the PPF is just low inside government schools due to the quality of inputs and/or motivation of teachers. We know that teaching quality, at least with basic literacy and numeracy, is unlikely lagging since results suggest that teaching students fundamentals is not hard. Volunteers with motivation but very little formal training can be quite effective in remediating early learning gaps and improving reading and mathematics levels. The Bihar experience proves that that the existing government teachers *can* do what the volunteers with a few days training are capable of doing, if they were a) as motivated as locally hired volunteer teachers, and/or b) if they are not under intense pressure to follow the prescribed syllabus. There are clearly elements of both. But a key part of the question is whether government teachers *are* motivated and *do* attempt to teach the curriculum, irrespective of whether students are capable of absorbing that material, or whether they just are not motivated to do anything. The fact that government teachers were able to accomplish learning gains in Bihar in a summer camp (in which they were freed from the usual constraints of teaching the curriculum) but did not produce similar gains during the school year suggests that, as Banerji (2000) writes, "A teacher might be aware that some children in her class have serious difficulties with basic arithmetic concepts that the competency level of these children (sometimes 20 to 30 per cent of a class) is well below the prescribed curriculum for that grade level…Most primary school teachers take the least risky route of concentrating on those children who perform at grade level and can cope with the syllabus."

The third explanation is that the curriculum moves too fast relative to learning and quickly becomes nearly irrelevant to the skills of the average student. What remediation does is re-center instruction onto actual ability levels.

---

[30] The lower success in Uttarakhand could also have to do with higher baseline learning levels than in Bihar, and greater institutional support for the program in Bihar.
[31] 74 percent of the children who attended the summer camp were in grades 3-5, but only 39 percent were in the "low-performing" target group, defined as being in the ASER word level 3 or below.
[32] Participants were mostly teachers newly-appointed by the local government Note these local *panchayat* teachers are government teachers, different from "state" teachers, who are paid much more.

Our simple simulation model can easily reproduce the kinds of impact seen from remediation in the face of uncentered instruction. Figure 16 shows how large variation in PPFs could affect whether students of a particular initial skill level learn. A curriculum catering to the bottom 40 percent of students would be centered over an average skill level of 251 (red PPF), whereas a curriculum centered on the entire grade 4 distribution would be at 278, and an accelerated, overambitious curriculum (with a pace of 90) would be centered on an average student skill of 362 (blue PPF). In an overambitious environment, on the increment to student learning for the bottom 40 percent of students is only 21 points, which is why exposing them to lessons centered on their skill level for even half the time can yield remarkable learning gains.

Figure 16: A PPF centered on the bottom 40 percent of students can result in large learning gains, especially in an environment with an overambitious curricular pace
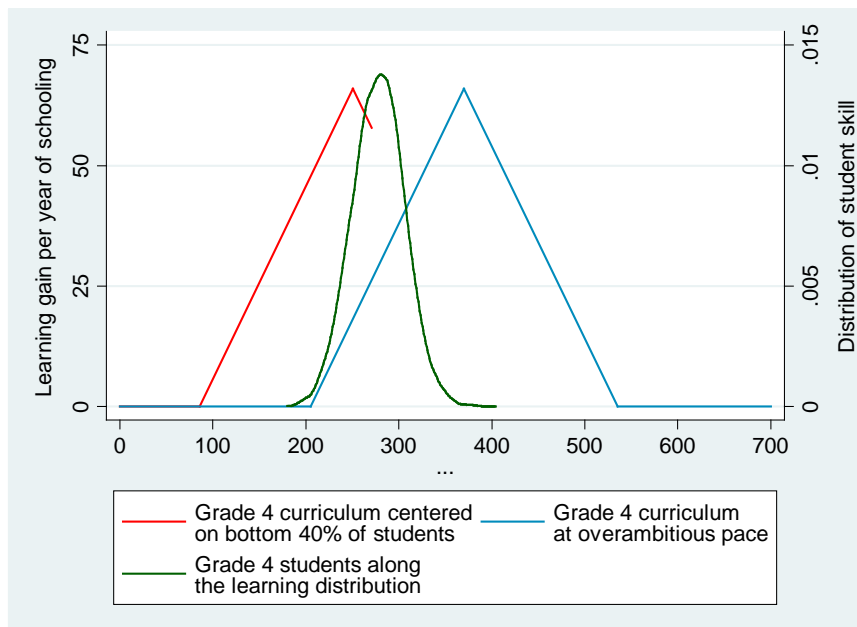
Similarly, Table 8 shows the results of using our basic simulation model to replicate the impact of remediation by assuming the bottom 40 percent of students are placed in a class with exactly the same PPF as a regular grade, except that it is centered on the actual average of those lower-performing students, rather than on the curriculum, for half the time. We calculate the "effect size" (learning as a fraction of student standard deviation) on students of that remediation in grades 3 and 4. With even modest curricular gap, the benefits of remediation are huge. At CP=80, the effect size on the bottom quintile is 0.5 effect sizes or higher and is 0.26 in grade 3 and 0.33 in grade 4—well into the range of gains of 0.43 for the bottom third. If there is no curricular gap (CP=60), then there is also no gain to remediation for quintile II students.

Table 8: Impact of remediation: the "effect size" of learning gains from children exposure to centered instruction 50 percent of the total instructional time

Curricular Pace	Grade 3		Grade 4	
	Quintile I	Quintile II	Quintile I	Quintile II
60	0.19	-0.09	0.18	-0.12
70	0.35	0.10	0.38	0.12
80	0.50	0.26	0.55	0.33
90	0.65	0.42	0.68	0.54
100	0.81	0.59	0.70	0.70

Source: Authors' simulations. Relative to base case, 40 percent of students are exposed 50 percent of the "regular" PPF and 50 percent of a remediated PPF

III.B.2) Early grade reading assessment

One puzzle is why effect sizes are so small when it seems there ought to be huge, easy gains when starting from low levels of performance. Given this puzzle, one ought to find that if the problem of curricular gap is addressed squarely, there are potentially very large learning gains. In the Early Grade Reading Assessment Plus (EGRA) in Liberia, which was subject to randomized evaluation, authors found learning two to three times faster in treatment than control schools and large effect sizes in oral reading and reading comprehension of 0.8 (Korda & Piper , 2011).

There is an increasing body of evidence from interventions that follow the "five Ts" (Time on task, teaching teachers, texts, tongue of instruction, and testing) in early intervention can produce massive results with demonstrated effect sizes of 0.4 to 0.8. The experiences include *Breakthrough to Literacy* in Zambia (USAID, 2011), *Malandi District Experiment* in Kenya (Crouch & Korda, 2009), *Systematic Method for Reading Success (SRMS)* (South Africa) (Piper, 2009), and *Read-Learn-Lead (RLL)* (Mali) (Gove & Cvelich, 2011).[33]

III.B.3) Tracking by skills/learning in early grades

Tracking,[34] or organizing students in different groups or classes by ability, has proven effective at promoting learning in some contexts (eg, see Figlio and Page 2002, Zimmer 2003).[35] Proponents argue that tracking brings efficiencies in instruction – less diversity in student ability

---

[33] Many thanks to Luis Crouch for these references and his guidance on EGRA and remediation.

[34] There are many variations on ability grouping in education. For the purposes of this paper, we use the term tracking to mean sorting pupils into *different classrooms within schools* based on ability, with pupils generally taking all subjects together. A more refined approach would involve ability grouping by subject (sometimes called "setting"), but we have not seen this tested in developing countries. We not consider other types of tracking, or the sorting of students into ability groups *across schools* or grades, which make it difficult for students to move once they have been selected to enter a certain group (also called "selective" as opposed to "comprehensive" schooling in the UK).

[35] Much of the tracking literature is suffers from identification problems because the research largely relies on studies that compare a) test score gains of students in different tracks within the same school, or b) tracked and un-tracked students of similar ability in different schools, without accounting for the fact that the presence of tracking in a school could affect school choice, and unobservable characteristics about the school that could affect test scores. The Duflo et al 2011 Kenya study described below is the first we know of to randomly assign schools to tracked and non-tracked status.

in the classroom allows teachers to better tailor lessons to student needs. Critics in claim that tracking promotes inequality since lower-ability learners become entrenched in lower-ability tracks and miss out on benefits from learning with higher-ability peers (Slavin, 1990).[36] Why might grouping students by ability in an environment with an overambitious curriculum improve learning? Duflo, Dupas and Kremer (2011) argue: "A central challenge of educational systems in developing countries is that students are extremely diverse, and the curriculum is largely not adapted to new learners. These results show that grouping students by preparedness or prior achievement and *focusing the teaching material at a level pertinent for them* could potentially have large positive effects with little or no additional resource cost" (emphasis added).

The researchers compared schools in which classes were split based on ability verses randomly, allowing all students to benefit from smaller grade 1 and 2 classes. Because some students were placed in non-tracked, smaller classes, authors were able to attribute results to tracking versus class size reductions. Students across the achievement distribution benefited from being grouped by achievement, and effects persisted into grade 3 after children were re-joined into one class. Students in top half of distribution gained 0.19 standard deviations in test scores while students in the bottom half gained 0.16 standard deviations, controlling for baseline scores. [37] Not surprisingly, recalibrating lessons allowed students to excel in new material. Students in the lower half of the distribution demonstrated gains in easier questions or basic skills (perhaps because there was insufficient time to master fundamental and more advanced skills), while the upper half gained in more advanced skills or harder questions (perhaps because they already knew the basics).

As with the other examples above, we cannot say conclusively that tracking in Kenya narrowed the curricular gap, but it seems plausible. The success of tracking depends on whether it shifts an overly-challenging syllabus to into the range of student learning potential – whether the PPF becomes more aligned with the distribution of student skill. Tracking would have little to no effect if the curricular and actual pace were already equal. However, tracking offers greatest benefits with an overambitious curriculum. Because tracking segregates higher and lower ability students in respective classes, teachers faced with a critical mass of students who cannot follow their lessons have little choice but slow down their curricular pace, and in turn the mass of students who end up learning nothing, as portrayed in figure 10, declines as more students can keep up with lessons in earlier years.[38]

III.C) Gains from changes in school management that allow "centering" the curriculum

In this section we look that the potential gains from private or community schools, or employing different types of teachers within public schools, with respect to how they structure curricula and syllabi. In no way are we advocating these approaches as universal superior mechanisms for

---

[36] In the USA, critics also claim that lower tracks receive fewer resources and are allocated lower quality teachers (Figlio & Page, 2002).

[37] Scores for students in tracking schools were 0.17 SD greater at the end of the program and persisted one later (controlling for baseline values and individual covariates). The fact that impact persisted suggests that in classes 1 and 2 students may have learned fundamental skills that they were able to apply in later grades.

[38] For higher-achieving students, the benefit from tracking may come from being in the classroom with similar achieving peers, even if teachers do not modify lessons, since teachers are by default already catering lessons to these students.

achieving learning; but instead explaining their performance at least in part due how these mechanisms narrow the curricular gap.

III.C.1) Private Schools

Why could private schools better align learning potential with curricular pace? Banerji (2000) explains that it is private schools' flexibility in adapting material to student ability. "Private schools do not follow the 'all pass' policy. Children who fail are kept back. Parents must seek remedial action outside the school. Through a heavy load of tests and reports, the school and parents ensure that a child does not fall behind."

The findings of a study in the Punjab province of Pakistan offer a similar explanation. This study demonstrates remarkable results from private schools, which make up 35 percent of primary school enrollment (Andrabi, Das, Khwaja, & Zajonc, 2011). Learning is deplorably low overall, but private schools are impressively less bad. In government schools, only 24 percent of third graders could form a sentence in Urdu using the word "school" while 46 percent of students in private schools could do so. There are massive differences in performance between private and public schools.   Simple cross-section estimates find learning of third graders in private school higher by effect sizes of 0.91, 0.57, and 0.45 in English, Math and Urdu respectively. The sophisticated estimates of the private school impact suggest learning per year in private schools is higher by an effect size of around 0.3 per year in each of these subjects. These are not "elite" private schools; these are mostly "mom and pop" small independent private schools that actually spend substantially *less* per student per year than government schools. In *cost effectiveness—* learning gain per year per rupee—private schools are *factor multiples* higher than government schools.[39]

The question is how the private school impact on learning can be so big when utilizing *less* resources and often using teachers at much lower wages, with less overall education and training as civil servant teachers. One obvious explanation is greater motivation and incentives of private schools, which leads to greater effort by teachers and school management. But, another possibility is that, while the government schools have a massive curricular gap, private schools can simply re-align the pace of instruction with actual student ability, and produce consistently much larger gains in learning and cumulative levels of skill. This realignment is often much easier for private schools and teachers to achieve, since they are frequently excused from the pressure of completing a single national curriculum, and may teach to a more homogeneous population in terms of ability.

Again, similar to the tracking explanation above, it is not that private schools demonstrate universal efficacy – this depends on whether they narrow the curricular gap relative to the gap in public schools. For example, the Pakistan findings are in striking contrast to the results from evidence and experiments in the USA where the heated debate is whether the pure causal learning impact of private schools is just small or zero. No one is proposing learning gains on

---

[39] These findings are consistent with other estimates from developing countries that private schools can provide more learning at substantially lower cost. Tooley and Dixon (2005) find students in private *unrecognized* schools in Hyderabad ahead of government schools by 0.84 and 1.51 standard deviations in math and English respectively. They also report private to government gaps between 0.3 and 1.15 standard deviations in Ga, Ghana and Lagos State, Nigeria. Interestingly, in Nairobi they find very small gaps.

the order of a full student standard deviation—and certainly not with costs per pupil a third to a half lower.

III.C.2) Community control of schools and/or teachers

One recurrent debate in education is between "official" schools and "community" schools or "formal" and "informal" schools. This contrasts "regular" government schools with civil service teachers controlled by a state bureaucracy and hired according to "meritocratic" criteria like degrees and teacher training, with "community" schools that often utilize local teachers, without civil service status, often working on contract, and rely much more on community inputs, and are often (implicitly if not explicitly) outside of the control of the government schooling bureaucracy.

Interestingly, in spite of having lower "quality" on most formal, input-based ratings, community controlled schools often perform as well, or better, on learning performance than regular schools. One possible reconciliation of this puzzle of "lower quality" leading to better outcomes is that when communities control schools, these schools teach to their students because accountability relationships are directly to the community and teachers do not have a commitment to the central curriculum.

A major challenge in evaluating community schools is that they are often located in areas where government schools haven't yet expanded due to remoteness or conflict or other reasons that might make it difficult for a centralized education bureaucracy to provide schooling. Chaudhury and Parajuli (2010) were able to get around this challenge by evaluating an initiative by the Government of Nepal to transfer school management responsibility from the Government to the community, allowing a school management committee, including parents and community members, to make staffing and financial decisions about the school, such as sending government-recruited teachers to return to district headquarters, to hire and fire community teachers, and to tie teacher salaries to performance.[40] They find that community schools did no worse than government schools on grade 3 curriculum-based achievement tests (math and language), and demonstrated improved access and equity, measured by the fraction of out-of-school children, especially those from disadvantaged caste or ethnicity groups. Community schools also demonstrated improved grade progression and repetition rates.

While utilizing locally-hired non-tenured teachers isn't offering local control of the school, it is similar to community schools in the sense that such teachers are usually recruited and supervised by the community, often by through school-level committees. These teachers, who, for example, receive contracts that are renewed annually by the school committee, could improve learning in several ways. First, they are likely under less pressure than the typical civil-servant teacher to follow the curriculum and instead focus instruction that meets local, community demands. Thus they have more discretion over what they teach and are better positioned to teach to students' ability, lowering the curricular gap. Second, they often come from the community and thus know the local context and student ability better, and are more attuned to student and community needs, thus further reinforcing community accountability. Third, locally-hired teachers may just

---

[40] Conversion from government to community management was voluntary, so the researchers employed a strategy of encouraging communities through an advocacy campaign in randomly-chosen areas to convert to community management.

work harder and show up to class more than civil servant teachers, as they could be dismissed due to absence or poor performance. This diligence in the classroom could lead to improved learning.

Muralidharan and Sundararaman (2010) examine the impact of randomly assigning contract teachers to 100 public schools in the Indian state of Andhra Pradesh. Just as community managed schools did no worse than government schools in changing academic performance, the authors found that contract teachers did no worse in improving student learning than civil-service teachers (who are five times as expensive). And students in schools with an extra contract teacher performed better in mathematics and language tests by 0.15 and 0.13 standard deviations, respectively, with the greatest gains (0.23 and 0.25 standard deviations) for schools in remote areas and grade 1 students.[41]

What motivated the learning improvements in Andhra Pradesh? It was indeed the case that contract teachers were "only" absent 16.3 percent of the time compared to 26.8 percent for civil service teachers, averaged over two years; and contract teachers showed higher rates of teaching activity. But it seems implausible that an approximately 10 percentage point difference in attendance could be driving the lack of *any* difference in student learning between highly-trained and well-paid teachers and those that have little to no training, less schooling and much lower pay. The community accountability explanation, which the authors don't test, remains possible – even if contract teachers are still absent some of the time, they may be more attentive to student needs as the community is monitoring and evaluating their performance. Or narrowing the curricular gap could offer a plausible explanation.

Figure 17 demonstrates how a teacher with much lower skills could promote just as much learning as a better trained or educated colleague – it depends on where the respective teachers are targeting their lessons. The smaller blue PPF is centered on the average grade 5 student (340) with a curricular pace of 60, whereas the red PPF is centered on an average student who can follow lessons with a pace of 90 (460). Although the height (maximum learning gain per year) and width (range of students for whom lesson is applicable) of the blue PPF are half that of the more accelerated red PPF, students learn more, or just as much, with the more modest curriculum since it is better tailored to student ability.

---

[41] It is possible that the contract teachers were more likely to be assigned to grade 1, since grade assignment was decided by the school. At least in the first year, contract teachers were more likely to be assigned to grades 1-3 as opposed to grades 4-5.

Figure 17: If community or contract teachers better center curricular pace on actual student skill, they can achieve the same learning gains as "better" teachers (e.g. higher PPF) but who are centered in the curriculum, not the student
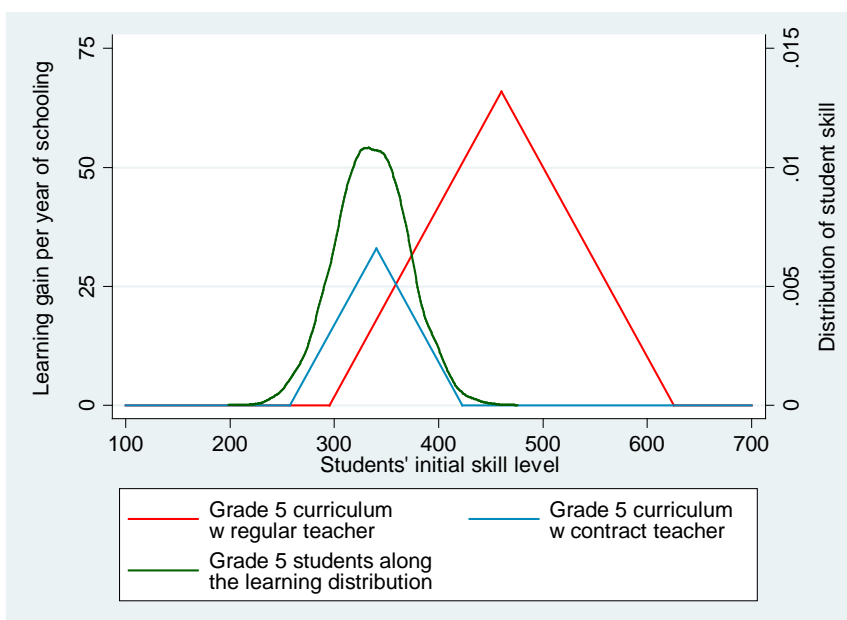
New interventions are underway that mix together community control, contract teachers, and a focus on early intervention and remediation. The Teacher's Community Assistant Initiative (TCAI) in Ghana for instance combines all these elements into a single intervention. It is now being rigorously evaluated by the NGO Innovations for Poverty Action and will have results soon.[42]

IV.) Conclusion

This paper takes the results from a number of recent experiments and studies, and encompasses their findings in a simple parsimonious formal simulation model. To the standard discussions of education quality, we add the key question of the alignment of teaching and instructional materials and curricular objectives, with the ability of students. This based on the common sense notion that material can be either too easy or too hard, and students learn the most when instruction is tailored to students' skill level.

We show that if the curricular pace—the level and material teachers are expected to teach—moves faster than actual student learning, this *alone* can generate enormous differences in cumulative learning. Calibrating a baseline model to reproduce typical OECD grade 8 results in internationally comparable assessments (e.g. PISA or TIMSS), we show that *all* of the observed learning differences between poor performing and OECD countries *could* be accounted for only by an overly accelerated curriculum in poor countries – even if the countries have exactly the same potential learning.

---

[42] See http://www.poverty-action.org/remedialeducation/scalingup

That is, the observed learning profiles (rates of learning per year of schooling completed) can be flat just because the material being taught to too hard for students as the curriculum has moved ahead, leaving students behind.

There is widespread evidence of a mismatch between curricula and student achievement—essentially every curricular referenced assessment given shows that students actual performance is many grade levels behind what the curriculum says they should have mastered.

If an educational system recognizes an existing large gap between curricular pace and actual learning, then the change from "business as usual" expansion of inputs is almost complete.

- More focus on "inputs" is unlikely to have much impact—and may detract attention from the real problem for a very long time.
- Efforts that "re-center" teaching on student actual skill/ability can have enormous pay-offs, whether through remediation (such as Pratham's programs in India or Early Grade Reading Assessment Plus-like programs), private schools, community-controlled schools, tracking, multi-grade teaching, or the adoption of other pedagogical reforms that allow teachers to focus on student mastery of basic skills. Our point is that these may or may not work in different contexts, depending on the extent of the curricular gap, and how much flexibility systems allow teachers and lessons to respond to it.
- Re-centering curricula and in turn teaching could offer a low-cost solution to improving learning, if systems are willing to re-focus learning goals on the average learner. (We know they are able of remedial teaching from the Banerjee, Duflo and Walton 2011 India results.)
- A focus on meeting early and achievable goals, rather than end of school high stakes for student examinations like university entrance exams, could re-orient teaching and learning away from what happens for a small elite towards the typical student, who is also then more equipped to move ahead.

We are proposing that there are large gains in learning from actions that are near costless—or even cost saving. Economists, ourselves included, are always (and often rightly) suspicious of "hundred dollar bills on the sidewalk"—if what we are saying is true, why haven't others already noticed this and made the changes? This is an important argument and one which we address at length elsewhere. We believe that the answer is that existing school systems have gained legitimacy and internal and external support through "isomorphic mimicry" (Pritchett, Woolcock, & Andrews, 2010) in which schooling systems are rewarded for "looking like" good schools rather than actually promoting learning. In this context, systems are geared to process compliance and inputs and hence just are not equipped to officially "see" (in the sense of Scott 1998) the realities of learning. In the context in which looking like good schools is rewarded, then each curriculum reform adds new skill sets and topics to be covered and helps with the camouflage of looking like functional or even "best practice" systems. Simplification or slowing down would look like failure—which is politically more salient than actually being a success- and hence systems actively resist "disruptive" (Christensen, 1997) innovations (like remedial teaching or contract teachers or radical revisions of the structure of classrooms around technology). The practical question is therefore how to sabotage the camouflage so that systems are motivated to adopt innovations, even ones that don't necessarily look good.

Works Cited

Andrabi, T., Das, J., Khwaja, A. I., & Zajonc, T. (2011, July). Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics, 3*(3), 29-54. Retrieved from http://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/cid-working-paper-no.-158

ASER. (2008). *Annual Status of Education Report (Rural) 2007.* New Delhi: Pratham Resource Company.

ASER. (2009). *Annual Status of Education Report (Rural) 2008.* New Delhi: Pratham Resource Company.

ASER. (2011). *Annual Status of Education Report (Rural) 2010.* New Delhi: Pratham Resource Company.

ASER Pakistan. (2011). *Annual Status of Education Report (Rural) 2010 Pakistan.* Retrieved from http://www.aserpakistan.safedafed.org/document/aser/2011/Total%20PDF%20Aser%20Report.pdf

Banerjee, A. V., & Duflo, E. (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty.* New York: Public Affairs.

Banerjee, A. V., Banerji , R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy, 2*(1), 1-30.

Banerjee, A. V., Duflo, E., & Walton, M. (2011). *Preliminary Report on the Impact Evaluation of Read India.*

Banerjee, A. V., Duflo, E., Cole, S., & Linden, L. (2007). Remedying Education Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics, 3*(122).

Banerji, R. (2000, March). Poverty and Primary Schooling: Field Studies from Mumbai and Delhi. *Economic and Political Weekly, 35*(10), 795-802.

Barro, R. J., & Lee, J.-W. (2011, November). A New Set of Educational Attainment in the World, 1950-2010. *NBER Working Paper No. 15902*. Retrieved from http://www.barrolee.com/

Bhattacharjea, S., Wadhwa, W., & Banerji, R. (2011). *Inside Primary Schools: A study of teaching and learning in rural India.* ASER and Pratham Mumbai Education Initiative. Retrieved from http://images2.asercentre.org/homepage/tl_study_print_ready_version_oct_7_2011.pdf

Chaudhury, N., & Parajuli, D. (2010). Giving it Back: Evaluating the Impact of Devolution of School Management to Communities In Nepal.

Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Harvard Business School Press.

Coleman, J. S. (1966). *Equality of Educational Opportunity Study.* United States Department of Health, Education, and Welfare.

Crouch, L., & Korda, M. (2009). *Improvements in Reading Skills in Kenya: An Experiment in the Malindi District.* Research Triangle Park, NC: RTI International. Retrieved from http://ddp-ext.worldbank.org/EdStats/KENdprep09.pdf

Das, J., Pandey, P., & Zajonc, T. (2006). Learning Levels and Gaps in Pakistan. *World Bank Policy Research Working Paper No. 4067.*

Das, J., & Zajonc, T. (2010, July). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics, 2*(92).

Dubeck, M. M., Jukes, M. C., & Okello, G. (2012, February). Early Primary Literacy Instruction in Kenya. *Comparative Education Review, 56*(1), 48-68.

Duflo, A. (2011). *Teacher Community Assistant Initiative: Baseline Survey Report.* Innovations for Poverty Action.

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*, 1739-74.

Duflo, E., Dupas, P., & Kremer, M. (2012). School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. Retrieved from http://www.nber.org/papers/w17939

Educational Initiatives. (2010). *Student Learning Study:Status of Student Learning across 18 States of India in Urban and Rural Schools.*

Figlio, D. N., & Page, M. E. (2002, May). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics, 51*(3), 497-514.

Filmer, D., Hasan, A., & Pritchett, L. (2007, August). A Millennium Learning Goal: Measuring Real Progress in Education. *Center for Global Development Working Paper No. 97.*

Glewwe, P. W., Hanushek, E., Humpage, S. D., & Ravina, R. (2011, October). School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. *NBER Working Paper No. 17554*. Retrieved from http://www.nber.org/papers/w17554

Glewwe, P., Kremer, M., & Moulin, S. (2009, January). Many Children Left Behind? Textbooks and Test Scores in Kenya. *American Economic Journal: Applied Economics, 1*(1).

Gove, A., & Cvelich, P. (2011). *Early Reading: Igniting Education for All. A report by the Early Grade Learning Community of Practice. Revised Edition.* Research Triangle Park, NC: Research Triangle Institute.

Gundlach, E., Woessmann, L., & Gmelin, J. (2001, April). The Decline of Schooling Productivity in OECD Countries. *The Economic Journal, 111*(471), 135-147.

Hanushek, E. A. (1986, September). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature, 24*(3), 1141-1177.

Hanushek, E. A. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer, 10*(2), 227-246.

Hanushek, E. A., & Woessmann, L. (2009, January). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *NBER Working Paper No. 14633*. Retrieved from http://www.nber.org/papers/w14633

Hernandez, D. J. (2011). *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation.* Annie E. Casey Foundation. Retrieved from http://www.gradelevelreading.net/wordpress/wp-content/uploads/2010/10/DoubleJeopardyReport040511FINAL.pdf

Hoxby, C. M. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of Economics*, 1239–1285.

Korda, M., & Piper , B. (2011). *EGRA Plus: Liberia--Final Program Evaluation Report.* Research Triangle Park, NC: RTI International. Retrieved from https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=283

Kremer, M. (2003, May). Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. *The American Economic Review, 93*(2), 102-106.

Kremer, M. R. (1995, August). Research on Schooling: What We Know and What We Don't. A Comment on Hanushek. *World Bank Research Observer, 10*(2), 247-254.

Muralidharan, K., & Sundararaman, V. (2010, May). Contract Teachers: Experimental Evidence from India. Retrieved from http://www.econ.ucsd.edu/~kamurali/papers/Working%20Papers/Contract%20Teachers%20(24%20May,%202010).pdf

Muralidharan, K., & Sundararaman, V. (2011, February). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy, 119*(1), 39-77.

Piper, B. (2009). *Integrated Education Program: Impact Study of SMRS Using Early Grade Reading Assessment in Three Provinces in South Africa.* RTI International, Research Triangle Park, NC. Retrieved from https://www.eddataglobal.org/courses/survey/resources/Handout_3_FINAL_RSA_SMRS_EGRA_Impact_Study.pdf

Pritchett, L., Woolcock, M., & Andrews, M. (2010). Capability Traps? The Mechanisms of Persistent Implementation Failure. Retrieved from http://www.hks.harvard.edu/fs/lpritch/Governance/capability_traps(june2010).pdf

Scott, J. C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed.* Yale University Press.

Slavin, R. E. (1990). Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis. *Review of Educational Research, 60*(3), 471-499.

Tooley, J., & Dixon, P. (2005). *Private Education isPrivate Education is Good for the Poor: A Study of Private Schools Serving the Poor in Low-Income Countries.* The Cato Institute. Retrieved from http://www.cato.org/pub_display.php?pub_id=5224

USAID. (2011). *Improved Learning Outcomes in Donor-Financed Education Projects: RTI's Experience.*

Uwezo. (2010). *Uwezo Uganda: Are Our Children Learning?* Retrieved from http://www.uwezo.net/uploads/files/Uwezo2010%20Assessment.pdf

Uwezo. (2011). *Are our children learning? Annual Learning Assessment Report Tanzania.* Retrieved from http://dl.dropbox.com/u/34915105/ALA%20UWEZO.pdf

Uwezo. (2011). *National Report Kenya.* Retrieved from http://dl.dropbox.com/u/9310905/Kenya/Assessment%202011/Results/National%20Report.pdf

Woessmann, L., & West, M. W. (2006). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. *European Economic Review, 50*(3), 695-736.

World Bank. (2005). *Expanding Opportunities and Building Competencies for Young People. A New Agenda for Secondary Education.* Washington, DC: World Bank.

World Bank. (2011). *Improving the Odds of Achieving the MDGs: Global Monitoring Report 2011.* Washington, DC: World Bank.

Zimmer, R. (2003). A New Twist in the Educational Tracking Debate. *Economics of Education Review, 22*, 307-315.