Chapter 1: Are students learning?

1.1 *This chapter of the LEAPS report examines how much primary school students are learning in rural Punjab Province and how much factors like school facilities, parental education, and wealth matter.* The sampling for the LEAPS report was done in two stages. In the first stage, three districts were chosen from the province of Punjab. The choice of districts follows the accepted division of the province into the "better" performing North and Central parts (Attock and Faisalabad) and a "poorer" performing South (Rahim Yar Khan). Within these districts, 112 villages were chosen randomly from a list of all villages with a private school in 2003. The total sample yields over 800 public and private schools. The LEAPS survey tested all 12,000 children in Class 3 enrolled in these schools on three subjects—English, Mathematics and Urdu. We describe the outcomes from these tests in two sections.

1.2 *Levels of learning.* Section 1 gives a sense of what children are learning in schools. It focuses both on "average learning" and the *specific* knowledge that children acquire in school. Can children add 5+14 and make a sentence with the word "ball"? Basic test data on counting and addition in Mathematics and alphabets in English and Urdu provide information on where children are compared to one another and where they stand relative to the curriculum.

1.3 *Learning levels and attributes.* Section 2 presents associations between children's learning levels and their attributes. It identifies learning "gaps" across geographical locations, schools (public/private), households (parental wealth and education) and children (age and gender), and looks at associations between learning and school characteristics, such as infrastructure and the student-teacher ratio.

1.4 *Summary of the findings.* First, learning levels are poor. Most children fail to meet the curricular goals set by Pakistan's Ministry of Education. Second, most variation in learning is *across* schools in the same village. That is, there are good and bad schools in all villages rather than good schools in certain types of villages and bad in others. Third, private schools significantly outperform government schools, and this difference dwarves all other associations. The worst government schools, where children learn virtually nothing, drive most of this gap. Fourth, there is no *single* characteristic of private schools that accounts for the vast difference with government schools—improving quality in government schools will require more than improving infrastructure or reducing the student-teacher ratio. These facts inform the structure and questions raised through the remainder of this report.

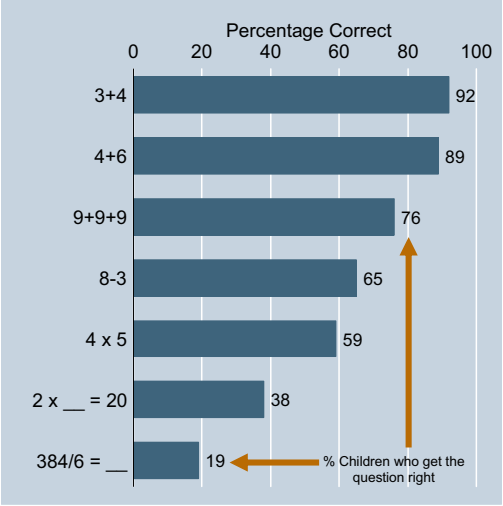
WHAT DO CHILDREN KNOW?

1.5 *Children are performing significantly below curricular standards.* The majority of children cannot answer simple questions in Mathematics, Urdu, and English. At the end of Class 3, many children have not mastered the Class 1 curriculum in Mathematics and a majority has not mastered the Class 2 curriculum. Only one-third can construct a sentence with the word "school" in Urdu.

1.6 *By the end of Class 3, just over 50 percent of the tested children have fully mastered the Mathematics curriculum for grade I.* They can add double-digit numbers and subtract single-digit numbers but not much more. They cannot subtract double-digit numbers, they cannot tell the time, and double-digit multiplication and simple long division are beyond reach for all except a small minority. In Urdu, they cannot form a sentence with the word "school" or the word "beautiful". Less than 20 percent can comprehend a simple paragraph. For children who never attend school beyond Class 3 (40 percent of boys and 50 percent of girls), this is all the formal learning they will receive for the remainder of their lives.

1.7 *The LEAPS test results in Mathematics suggest children are performing significantly below their current grade level.* Figure 1.1, presents a number of Mathematics questions from the test along with the percentage of children who answered the question correctly. 89 percent answer "4+6=" correctly. Subtraction is the next best skill but only 65 percent answer the question "8-3 =" correctly. More difficult subtraction problems such as "238-129=" are answered correctly by only 32 percent of children. Both multiplication and division skills have not yet solidified and more advanced topics such as fractions are beyond all but the best students. For example, only 59 percent can solve the basic multiplication problem "4 x 5=" and just 19 percent can solve the long division problem "384/6 =".

Figure 1.1: Children at the end of Class III cannot perform basic mathematical operations



Box 1.1: LEAPS Survey Tests

Several characteristics of the LEAPS test design and administration are unique and important in the Pakistani context.

- LEAPS is the largest survey of learning outcomes in both public and private schools in rural Punjab. Children were tested in every school in the 112 villages surveyed, yielding information on over 12,000 children in 800 public and private schools in the districts of Attock, Faisalabad, and Rahim Yar Khan.
- The tests were developed with the help of experts familiar with the curriculum as well as international testing protocols. The tests were extensively piloted and revised to ensure that it would capture information not only where children were with respect to the curriculum, but also where they were on the "ladder of learning." To accurately capture what children knew, all tests started simply but ended with questions beyond the Class III material. What do we mean by very simple? In English, investigator made three alphabets sounds aloud (with a pause in between each) and asked the children to write each one down. Mathematics tests started with counting. In Urdu, the tests started with the written alphabet. These simple test questions minimized interactions between the test-takers and the students, which have been shown to affect results in other testing exercises.
- The testing was administered by the LEAPS team with no inputs allowed from the teachers of the tested children. Children were selected by constructing school-by-school rosters for the tested class. More than 90 percent of all children in the rosters were eventually tested, suggesting that the Class III population was effectively covered through the testing procedure.
- The test scores were then analyzed using the latest tools and methods from the educational testing literature to ensure strict comparability across schools and children. Both multiple choice and free response questions were included.

Students answered approximately 30 percent of the questions correctly in each subject—indicating that the majority of exam questions were significantly above the ability of most students. That said, a few children received almost perfect scores even though questions toward the end of the test were significantly beyond the grade-level in which they were tested. Details on the test instrument are presented in the Technical Annex to this document.

Comparing LEAPS results with Curriculum Standards

1.8 *The gravity of the problem in Mathematics becomes obvious when comparing how students are expected to perform under the new curriculum standards developed by the Ministry of Education* (http://www.moe.gov.pk-->curriculum) *with how they performed in the LEAPS test.*

Curriculum standard 1: In grades I-II children should be able to "count, read and write numbers up to 999."

LEAPS test results: Only 47 percent of all children in Class 3 can answer the simple counting question in Figure 1.2. Part of the reason why children are unable to answer this simple question but can do more complicated additions is because they are used to being asked the same questions over and over again; the moment the question does not directly match what they have seen before, they are stumped.

Figure 1.2: Counting and Subtraction Questions

Math Question: Counting	**Math Question: Subtraction**
Instructions: Tick the number which matches the number of stars. * * * * * 1 2 3 4 5	Solve: 238 -129

Curriculum standard 2: In Class 1 and 2, children should be able to "add and subtract up to 3-digit numbers."

LEAPS test results: Only 32 percent of children in Class 3 can correctly perform the subtraction shown above in Figure 1.2. There is no new "format" issue here that requires children to think differently. This is exactly how subtraction questions are given to children in basic tests. Children can barely perform single digit subtractions; anything more complicated is too difficult.

Figure 1.3: 3-Digit Division

Math Question: Division
Solve: $6\overline{)384}$

Curriculum standard 3: In Class 3-5, children should be able to "multiply and divide up to 6-digit numbers by 2 and 3-digit numbers."

LEAPS test results: Only 19 percent of all tested Class 3 children can divide a 3-digit by a single digit number as in the problem shown in Figure 1.3.

1.9 *The LEAPS test results in English also suggest children are performing significantly below their current grade level.* As Table 1.1 demonstrates, the average child can do little more than write the alphabet, and a fair number of children cannot even complete this task proficiently. In terms of alphabet familiarity, 86 percent can write the letter "D" when they hear it spoken and 70 percent can fill in the blank letter E in a sequence question like "D, __, F". Far fewer children can successfully spell a word they hear—only 46 percent can spell the word "bat" when it is spoken and only 20 percent can correctly spell "girl". Children are somewhat more successful in recognizing common words when presented with a picture of the object. Presented with a picture of a book and three possible words, 70 percent check "book" correctly. When shown a picture of a bird, a more difficult word, 57 percent of children can correctly select "bird" from the list "cow", "bird" and "boy". More

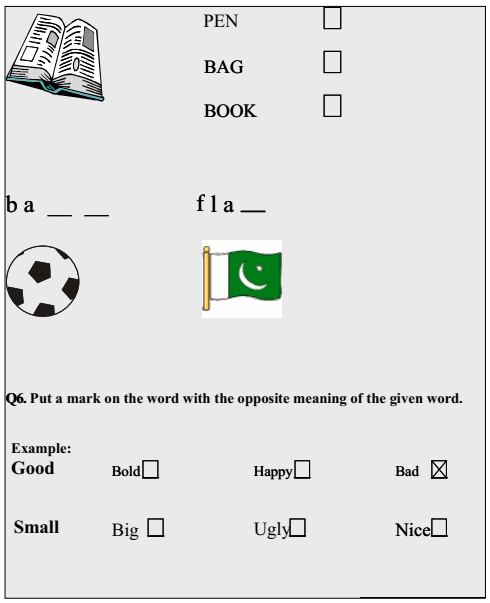
advanced language skills such as constructing a basic sentence given a particular word is beyond the vast majority of children. Only 11 percent can form a coherent and grammatically correct sentence using the word "school" (e.g., "I go to school").

Table 1.1: What do Children Know in English

Subject	The Question	Percentage who answered correctly	Corresponding Class for Curriculum
English	Write the letter "D" (spoken out loud)	86	1-2
English	Fill in the right letter D _ F	70	1-2
English	Tick the correct answer to match the picture (picture of book)	70	1-2
English	Fill in the missing letters (picture of a ball) BA _ _	45	1-2
English	Fill in the missing letters (picture of a flag) F L A _	29	1-2

1.10 *The gravity of the problem in English becomes obvious when comparing how students are expected to perform under the new curriculum standards developed by the Ministry of Education* (http://www.moe.gov.pk-->curriculum) *with how they performed in the LEAPS test.* According to the curriculum, the first benchmark for children in Classes 1-2 is that they "use reading readiness strategies; recognize words and sentences as meaningful units of expression and paragraphs as graphical units of expression." They are also (Benchmark 3) supposed to "Locate information from a visual cue or a graphic organizer and express the information verbally." Finally, under "Writing Skills" (C2 in the curriculum) children in Classes 1-2 are supposed to "Write words and sentences using writing techniques".

Figure 1.4: Basic English Questions



1.11 Figure 1.4 on the left shows three test questions for English. In the first, the instruction asks the child to tick the matching word, for the second the child has to fill in the blanks and for the third, tick the opposite word. All questions start off with a completed example, as in the third question, where "bad" is ticked as the opposite of "good" to show what is expected. Seventy percent of all children answer the first question correctly, 45 percent can correctly complete "BALL", 29 percent complete "FLAG" (putting in the single "G") and only 36 percent correctly tick that the opposite of "SMALL" is "BIG". The percentage correct on the "opposites" question is much worse than it seems— if children *randomly* ticked a box, 33 percent would get it correct so that performance is only *3 percent* above what randomness would imply. Given these results, it is not at all surprising that only 11 percent of all children can correctly form a sentence with the word "school".

Box 1.2: Rote learning is widely prevalent in both government and private schools.

We encountered rote learning first hand during the test development phase of the LEAPS assessment tool in a small private school in a village about 100 miles from Lahore. The children in the school struggled with a simple reading comprehension exercise conducted informally by the LEAPS team. We were puzzled because the same children had done quite well in a much more advanced English reading comprehension passage used in the school in their last internal examination. The puzzle was solved when we found out that the passage on the internal test was taken verbatim from the textbook used in the class. Each child had practiced and mostly memorized all the main passages in the prescribed anthology.

Testing children using template questions not only leads to official exams overstating children's subject mastery, it also results in them forgetting the important skill of decoding instructions. When administering the English exam in a second school in Rawalpindi, we found that students were completely stumped when the format of the question was changed slightly. The question was on understanding the difference in usage of a masculine vs. a feminine gender noun—a standard third grade question in Urdu. In Pakistani exams, the question is typically asked by having students convert a masculine noun into a feminine one and vice versa. Our question asked students whether a given noun was masculine or feminine. Most of the students could not answer that question even though the content was well below grade level.

Other examples abound. An essay on "your last actual holiday trip" led to a majority of students in a school in central Punjab answering in very similar tone about the beauties of Murree. In Math, a free response question showed the picture of a parallelogram and a rectangle drawn on graph paper and asked "How are these shapes different and how are they similar?" It drew a complete blank, even among fourth graders at an "elite" English medium school. Upon prompting, the students confided that they had never been exposed to that type of question. We eventually dropped that question in our actual test because of low discriminatory power vis-à-vis student ability. In plain English, nobody came even close to giving a satisfactory answer.

Source: Tahir Andrabi.

1.12 *The LEAPS test results in Urdu show children are performing significantly below their current grade level.* While English is a foreign language for most Pakistani children, we should at least expect that by Class 3 children have a basic grasp of reading and writing in Urdu. Yet only 31 percent (see Table 1.2) can write a grammatically correct and coherent sentence using the word "school." Likewise less then 30 percent of children can answer the most basic questions after reading a short paragraph. More difficult reading comprehension questions are answered incorrectly by nearly all children. Converting singular words to plural is particular troublesome- just 12 percent of children can correctly pluralize "habit" (*adat*). Children also struggle to separate words into their constituent letters –54 percent fail to separate the letters in the word "country" (*watan)*. And they have similar difficulties with combining letters into words–29 percent fail to combine the letters of the word "work" (*kaam*) and 73 percent were unable combine the letters of "morning" (*subah*). When presented with a picture of a book, only 73 percent of children chose the matching written word out of three options. In Urdu most children are performing just at the standards meant for Class 1.

Table 1.2: What do Children Know in Urdu

Subject	The Question	Percentage who answered correctly	Corresponding Class for Curriculum
Urdu	Tick the correct answer to match the picture (picture of house)	52	1-2
Urdu	Tick the correct answer to match the picture (picture of book)	73	1-2
Urdu	Write a sentence with the word "beautiful"	33	1-2
Urdu	Write a sentence with the word "school"	31	1-2

1.13 *The view that rising enrollment signals all is well is incorrect.* Learning has not improved dramatically since the time of the first LEAPS test in 2003. Although this report is primarily based on the first LEAPS cohort (class 3 children tested in 2003), we also tested a second cohort of third graders from the same schools in 2006. Table 1.3 displays the knowledge scores (explained below) and the percent correct score for both cohorts. Using knowledge scores that are comparable across years, the first cohort of third graders receives on average 500 across all three subjects (by construction). Three years later in 2006, the second cohort of same-aged children scored similarly in English but somewhat lower in Math (463) and Urdu (479). Learning lagged even as enrollment grew. This not surprising if one believes that the marginal enrolled child is at a greater disadvantage than the typical child or if increasing enrollment stresses the educational system.

Table 1.3: Rising Enrollment, Lagging Learning

Subject	First Cohort Class 3 in 2003			Second Cohort Class 3 in 2006	
	Knowledge score	Percent Score		Knowledge score	Percent Score
English	500	30	→	502	31
		29	↘	463	27
		38	↘	479	34

WHY DO SOME CHILDREN OUTPERFORM OTHERS?

1.14 *Knowledge scores are better at distinguishing between high and low performing children.* While the average child scores between 30 and 35 percent across the three subjects, some children score 0 percent and others score 100 percent. To understand what characteristics are associated with greater learning this study uses *knowledge scores* rather than the percentage correct on the test. There are a number of reasons for doing so. First, the percentage of questions correct can be a potentially misleading measure of "knowledge" if questions vary in difficulty. That is, the difference between 60 and 70 percent may be much larger in terms of the child's knowledge than the difference between 30 and 40 percent if questions in the test are progressively harder. Second, these knowledge scores, which are centered at 500 with a standard-deviation of 150 have a clear interpretation in terms of rankings: the difference between someone scoring 500 and someone scoring 650 is the same as moving from a rank of 50th out of 100 to a rank of 84th out of 100; the difference between 500 and 800 is the same as moving from a rank of 50th to a rank of 97th.

1.15 *Using these knowledge scores, the report describes three main findings in the data.* The first finding relates to the breakdown of overall variation in the data; the second and third then take a closer look at differences across children from different backgrounds and in different schools.

Table 1.4: Knowledge Scores

Exam Question	Question	Pakistan Knowledge Points		
		350	500	650
Complete the addition problem:	36+61	76	89	95
Complete the subtraction problem:	238-129	11	28	55
Complete the multiplication problem:	4*32	22	49	77
Complete a sentence with the word:	Beautiful	2	15	78
Complete a sentence with the word:	School	1	10	76

Box 1.3: Computing Knowledge Scores

Knowledge Scores are computed using a technique known as "Item Response Theory" which gives different weights to correct answers depending on the difficulty of the question. This is the same methodology used for international exams such as TIMMS and most national testing programs such the United States' SAT. The knowledge score can thus be interpreted as the student's knowledge or ability in a given subject area such as Mathematics, English, or Urdu. By construction the mean knowledge score is 500 and the scores range from 0 to 999, with a standard deviation of 150. Note though that a score of 500 does not mean that a student is meeting the curricular standards for Class III—it's simply the average score of children in this test. As a brief guide (see Table 1.4), a child at the mean of the knowledge-scale distribution (with a score of 500) can add two single-digit numbers in mathematics, complete alphabets in English and recognize simple words in Urdu. A child with a knowledge score of 300 (2 standard deviations below average) can barely count, may be able to recognize one or two English alphabets, and can write the Urdu alphabet. A child with a knowledge score of 800, can add, subtract, multiply, and divide large numbers and can understand, although not fully manipulate, fractions.

Enrollment and learning are two completely different processes

1.16 We have all heard about the high-performing and lagging regions in education; provincially, Punjab is performing well, Sindh and Balochistan poorly. Within Punjab, the North and Central are relatively better performing than the South. While correct for enrollment patterns, there is no evidence that learning levels follow similar patterns.

1.17 *The two main enrollment patterns show the dramatic gender-gap in enrollment and the positive effects of higher village literacy and wealth.* Table 1.5 (also in the introduction) shows enrollment patterns for children in rich/poor and high/low literacy villages. Here we add in test scores in Mathematics across the different village categories (Mathematics is chosen simply for readability; English and Urdu scores show very similar patterns). The two main enrollment patterns show the dramatic gender-gap in enrollment, with female enrollment always lower than male enrollment, and the positive effects of higher village literacy and wealth, whereby more children are enrolled in villages that are richer and/or more literate. These increases are fairly dramatic, and more so for girls compared to boys. For instance, moving from a low- to a high-literacy village increases enrollment by 27 percent for boys and 32 percent for girls when the village is poor, and by 11 (boys) and 29 percent (girls) when the village is rich. Although more literate villages are also wealthier, it is clear that even when villages are poor, literacy has a large effect on enrollment. Given the strength of the relationship, one would expect similar patterns to emerge for learning.

Table 1.5: While the percent enrolled increases with village wealth and literacy….

Village Wealth	Gender	% Enrolled by Village Literacy		
		Low	Medium	High
Poor	Male	56	75	83
	Female	38	65	75
Middle	Male	82	77	88
	Female	62	66	80
Rich	Male	73	82	84
	Female	53	73	82

Similar patterns do not emerge for levels of learning

Village Wealth	Gender	Mathematics Knowledge Scores by Village Literacy		
		Low	Medium	High
Poor	Male	512	526	565
	Female	463	495	503
Middle	Male	484	493	484
	Female	455	462	489
Rich	Male	497	515	525
	Female	507	522	522

1.18 *A similar pattern does not emerge for learning.* While the first row *does* suggest that moving from a low- to a high-literacy village increases Mathematics scores, a closer look reveals no systematic pattern. Villages with medium wealth appear to be doing worse for all literacy categories; among high-wealth villages, those with high literacy are doing worse than those with medium literacy. The best performing category are those in high-wealth villages with high literacy; the next best are those in low wealth villages with high literacy. Finally, there is no gender-gap in the knowledge scores—girls are performing as well as (or even better than) boys.

If you take away enrollment, a child is only as good as the school she attends

1.19 *Most villages have both good and bad schools.* Given that village wealth and literacy have little to do with children's performance in tests, one possible explanation is that these differences arise from good and bad schools in the *same* village rather than good schools in a rich/more literate village and bad schools in the others. To examine this hypothesis, we ranked villages by average mathematics scores and then selected the 5th best village and the 5th worst village for comparison (with a total of 112, this corresponds roughly to the 5th and 95th percentiles of the distribution). On average, schools in the 5th best village naturally do better than those in the 5th worst, but when schools are compared individually, it becomes clear that this village average masks a more complicated story

1.20 *That average mathematics scores in these two villages are different does not imply that all schools in the "bad" village are worse than all schools in the "good" village.* In fact, it is the opposite—of 10 schools in the "bad" village, 7 are better than the worst school in the "good" village. Of the 14 schools in these two villages, 10 "overlap": That

is, the majority of schools in the "bad" village perform similarly to the schools in the "good" village. Only the two worst schools in the "bad" village and two best schools in the "good" village do not have counterparts.

Box 1.4: Results from the Learning Decomposition

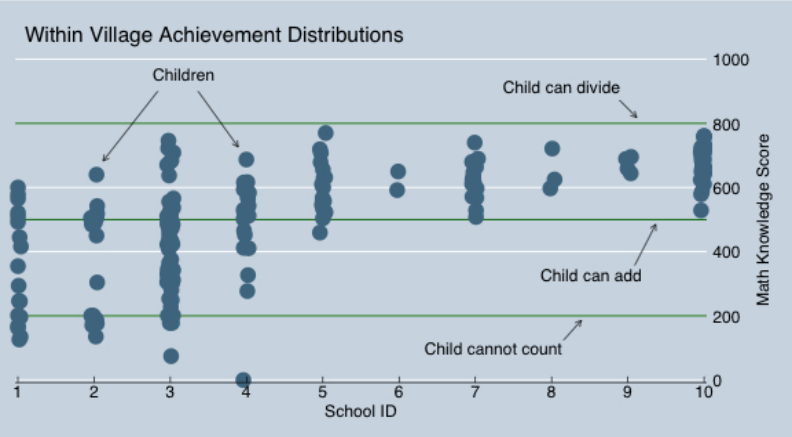
That fact that test scores are not associated with village wealth and literacy leads to the natural question "How much of the difference in test scores can all village attributes together explain?" A simple version of such a "variance decomposition" exercise sequentially regresses test scores on district, village, and school dummies; the residual variation is assumed to be driven by differences across children and unexplained variation in the data such as measurement error.

In our sample 50 percent of the total variation occurs across schools rather than across children in the same school. School effects explain the most variation for English (more than 60 percent of the total variation) and the least for Urdu (still 50 percent). Consistent with the finding that there are no "good" or "bad" villages, the portion of variation attributable to villages in Pakistan is small; for all subjects it accounts for less than 15 percent of the total variation.

Surprisingly child and household characteristics such as age, gender, parental literacy, and wealth add little explanatory power: With a full set of household and child covariates, and a school fixed effect, total explained variation never exceeds 68 percent. The percentage of variation explained by differences across schools is large compared to what has been found in other countries. The average variation across schools in 29 OECD countries, for instance, is 33 percent and only 6 countries (Austria, Germany, Hungary, Italy, Japan, Netherlands and Turkey) report between-school variation above 50 percent in Mathematics (OECD 2005).

1.21 *Given the large differences across schools in the same village, a natural follow-up question is to ask whether there are differences across children in the same school.* Do we see that most children perform at the same levels when they are studying together? One simple way to do this is to plot the scores, school by school in a given village, as in Figure 1.5. The figure plots each student's Mathematics score against a school identifier. This shows the full distribution, for every child, of schools in a typical village, grouped by the school they attend. The three horizontal lines indicate knowledge scores of 200, 500 and 800—children scoring 200 can barely count; those scoring 500 can add two-digit numbers and those scoring 800 can multiply and divide.

Figure 1.5: Different schools in the same village are very different….and so are different children in the same classroom



1.22 *More than one-half of the variation in learning is across schools*. There are 10 schools in this village and each one is identified by a unique number on bottom of the graph. As the visual inspection makes clear, some schools are fairly good and others are not doing well. This is simply a repetition of an earlier point—large variation across schools implies that in the same village, there are schools where every child can at least add, and schools where the vast majority of children cannot. Formally, following the terminology in Box 1.5, 50 percent of the variation in learning is across schools.

1.23 *Of the total variation in learning, 50 percent is within the same class*. Even within schools, children are at very different levels of learning. Schools 6 through 10 all have children who are performing at a roughly similar level and reasonably well, but in schools 1 through 5 children *in the same class and taught by the same teacher* are miles apart in what they know. In school 2, two groups of children are clustered over 200 knowledge points apart. School 1 baffles—even though it performs the worst on average, it displays the full spectrum of learning in the same class, ranging from children who cannot count to children who are close to performing division problems. While in schools 6 and 8 with 2 and 3 children in the class it is possible to see how teachers may actually accommodate children at different levels of understanding, school 3 has 61 children taught by the same teacher. It is unclear how this teacher can cater to the needs of children who can divide and children who cannot count at the same time. Again, following Box 1.5, of the total variation in learning, 50 percent is within the same class.

1.24 Two messages thus emerge. First, higher enrollments do not necessarily mean better learning—enrollment and learning are different processes. Second, the largest differences in learning are not across children from different types of households or children living in different types of villages, but children enrolled in different schools. Given that every village in our sample has at least one private school, it is then natural to ask what fraction of the variation in schools within the same village arises from the difference between government and private schools.

Private schools outperform than government schools. The difference is much larger than those across socioeconomic groups, such as children from rich and poor villages.

1.25 *Government schools underperform private schools on average in both the best and worst villages.* Figure 1.6 looks at schools in the 5th worst and the 5th best village—with one added fact. Here, we separate the private and government schools, using triangles to show private and circles to show the government schools. Government schools clearly under-perform private schools on average in both villages. However, the mean comparison does not imply that all government schools are worse than all private schools—in the "bad" village, three government schools are better than the worst private school, and the top performing school is a government school. At the same time, the worst schools are also run by the government so that the four worst schools in the "bad" village are all government-run, and the bottom-of-the-pile schools are truly dismal—with an average score below 100, children in this school can barely count after four years of education.

Figure 1.6: Different schools in the same village are very different

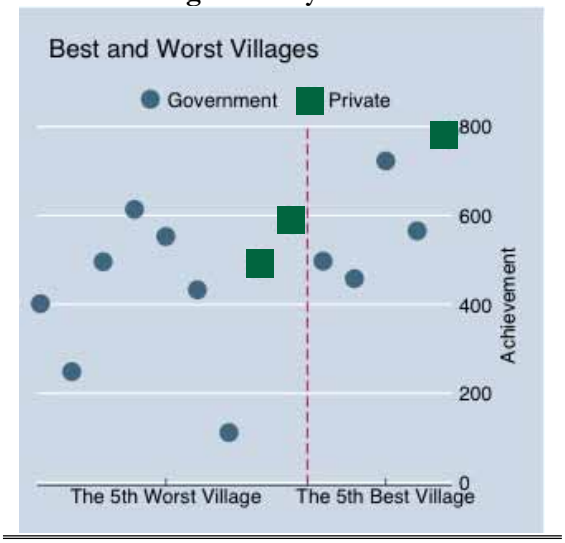
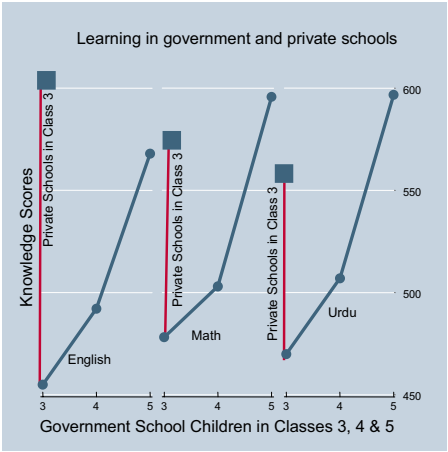


Figure 1.7: Private school students are years ahead



1.26 *The differences between government and private schools are systematic.* Children in private schools score significantly higher than those in government schools, even when they are from the same village. To understand how large these public-private differences are, Figure 1.7 shows the difference in knowledge scores between children in public and private schools.[5] As a rough guide, a knowledge score difference of 150 points translates into an increase in the ranking of the child from 50th to 85th out of 100 children; a knowledge score difference of 300 increases rankings from 50th to 97th out of 100. The knowledge scores of children in private schools are between 76 (Urdu) to 149 (English) units higher than those in government schools. Children in government

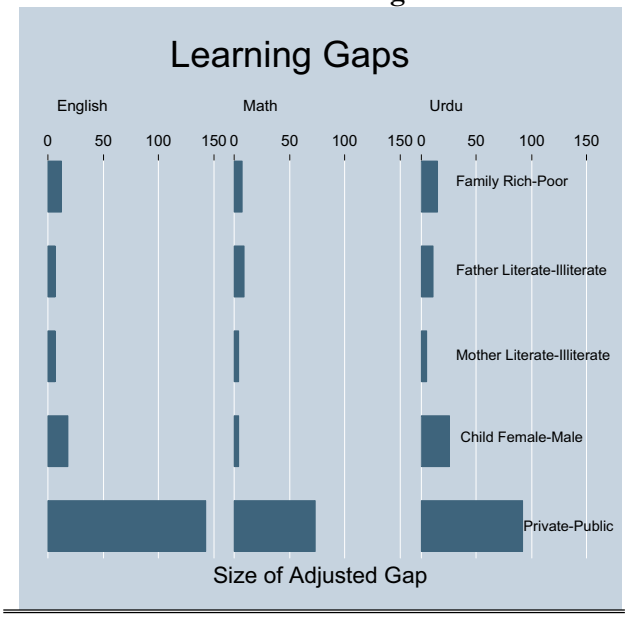
[5] The adjusted gap has a simple interpretation -- it shows the extent to which two children with the *same* observable characteristics and studying in the same school would score differently due to (say) differences in parental wealth. Similarly, in comparing public and private schools, we ask whether two schools—one public and one private—with similar students still show differences in learning. A natural question is whether the differences between public and private schools are driven by differences in school characteristics—we address this issue below.

schools will be among the worst performing 20 percent in private schools in English, and the worst performing 30 percent in Urdu.

1.27 *Children in public schools will take 1.5-2.5 years to catch up to private school children in grade 3.* To understand the size of the public-private gap in test-scores, Figure 1.8 also shows how long it takes for the same public school children who were tested in Class III and followed through to grade V to "catch-up" For all three subjects, children in public schools will report the same test scores as children in private schools after 1.5-2.5 additional years of learning. In English, government school children in grade V have still not caught up with private school children in Class III. Even in Urdu, an additional 1.5 years of schooling is required before government school children catch up with their counterparts in private schools.

1.28 *The public-private learning gap is much larger than that across children from different socioeconomic backgrounds.* Another way to benchmark the private-public gap in learning is to compare it to differences across widely emphasized parental dimensions, such as parental literacy and wealth. The gap between public and private schools in English is 12 times that between rich and poor children. The gap between public and private schools in Mathematics is 8 times that between children with literate and illiterate fathers. The gap between public and private schools in Urdu is 18 times the gap between children with literate and illiterate mothers (Figure 1.8).

Figure 1.8: The gap between private and public schools is 8 to 18 times the gap between socioeconomic backgrounds



1.29 *A priori, one might expect to find that there are also differences among government schools and among private schools.* Casual observation suggests that private schools range from elite institutions to temporary schools run by a local high school graduate seeking supplemental income. Likewise, government schools sometimes appear well-managed and staffed but just as often appear dilapidated and mismanaged. Indeed, the data confirm that the stark difference between government and private school performance is not because all government schools are performing poorly and all private schools well. To the contrary, the best government schools are competing head-to-head with even some of the top private schools. Of the top 10 percent of schools in Mathematics, 39 percent are government schools. In fact, out of the 804 schools tested, the top scoring school in mathematics is a government school. Even on the English exam, 17 percent of the top 10 percent of schools are government schools.

1.30 *The gap between top and bottom ranked government school is truly massive compared to any other gaps found.* The real problem with poor government school performance is the much higher variation in test scores across government schools and the performance of the worst schools. In English the worst government school scored 84 and the best scored 845—a ten-fold difference in performance. By comparison, the worst private school scored 351 and the best scored 850—only a 2.5 fold difference. The difference between good and bad government schools is almost 10 times that between children from rich and poor families. The difference is almost 15 times that between children with literate and illiterate parents. This extreme range of performance in government schools is not the result of two outlying schools. The vast majority of poor performing schools are government schools. Government schools account for 91 percent of schools performing in the bottom 10 percent in Mathematics, 95 percent of the bottom 10 percent in Urdu and the entire bottom 10 percent in English. When government schools fail, they fail completely, and the difference between schools dwarfs the difference between children from different households.

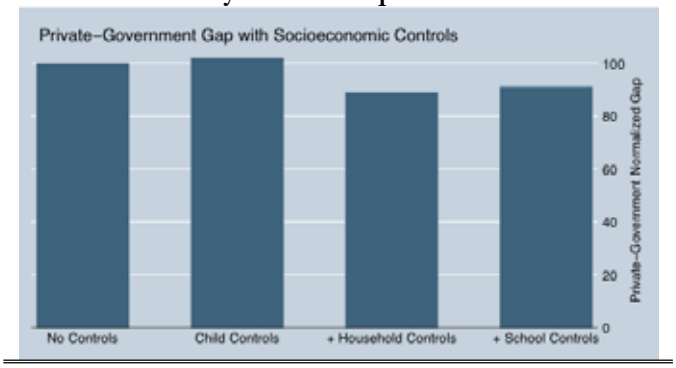
1.31 *The real problem in the government sector is a "long-tail" of dismal schools with no counterparts in the private system.* A clear picture emerges of learning achievement in the sample villages. All villages have good and bad schools. Although there are some differences across children from different backgrounds, these are small compared to that between government and private schools. An average private school has children with higher test scores than an average government school. Third, average performance masks the fact that the top-performing government schools are almost as good as the top-performing private schools; the real problem in the government sector is a "long-tail" of dismal schools that have no counterparts in the private system.

1.32 *An immediate reaction to these results is to point to the differences between private and public schools in their student-body and in their inputs.* One reaction may even be to say that in many ways this is a remarkable accomplishment for these government schools. Attaining top scores with open admittance and large class sizes is considerably more difficult than with a closed admissions process and small classes, regardless of the quality of instruction. Some might say government schools perform poorly because they have children from poor and illiterate backgrounds, insufficient resources, poor infrastructure and very high student-teacher ratios. But does this argument hold water?

The nature of the student-body, the quality of school infrastructure and Student-Teacher Ratios (STRs) have little to do with learning outcomes

1.33 *Even among schools with similar students, similar infrastructure and STR, private schools substantially outperform government schools*. Figure 1.9 starts with the raw gap "normalized" to 100 points, and limits the analysis to Mathematics. A multiple regression framework is used to see how much the gap reduces when including additional controls for child, family, and school characteristics. The first bar introduces child characteristics (gender, age, height-for-age, parental perception of intelligence and self-reported health-status), the second bar introduces family characteristics (education, wealth and home environment measured by the availability of books/media at home), and the third introduces school characteristics (infrastructure index and STR). While adding controls for child and household characteristic reduces the gap by 25 percent, the difference between public and private school test-scores remain large and observed school characteristics fail to explain any of the difference.

Figure 1.9: Controlling for school, household and child characteristics reduces the public-private gap by at most 20 percent



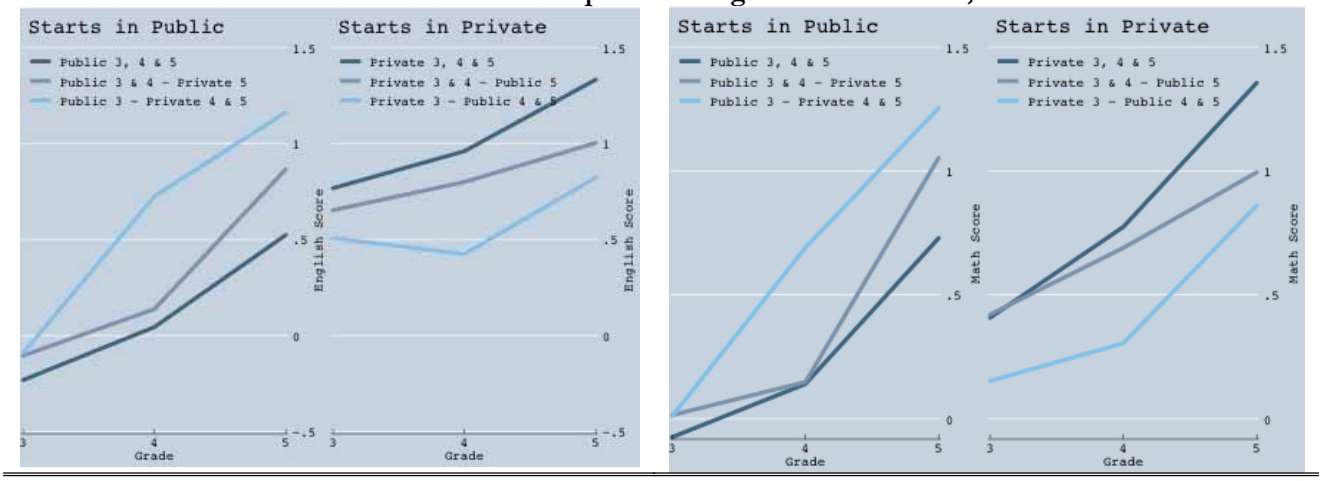
Box 1.5: Field Notes - How Teachers Adjust To High Student-Teacher Ratios

Why is it that greater STR's have no effect on test scores? One reason could be that teachers are imaginative in finding ways to teach even when the STR is high. For example, in a Government Girl's Elementary School in Attock's Mauza number 35 grade 3 has 94 students and a single teacher. It was a big problem to seat them all in one room when it came time for testing, so we used two rooms. When I asked the class teacher how she managed to teach in the class room with such a large number of girls, she said "*it is difficult to teach and manage them in a room. The way I manage them is that there is a monitor for each column and she is sitting at position 1 in each column, she helps a lot while listening to the lessons and keeping the students quiet*". After the testing I found that the school is performing well compared to the other government schools in this Mauza. It made me wonder how well this teacher could do if she had fewer children!
Source: Sarfraz Bhatti.

1.34 *Controlling for a wide group of child, household, and school characteristics makes relatively little difference to the magnitude of the public-private learning gap*. Educationalists and economists will also be concerned with other variables that we cannot control for in the multivariate regression framework; for instance, the ability or intrinsic motivation of the child. If parents send more motivated children to private schools, maybe the test-score difference is due to differences in motivation rather than the type of school. Using the three years of

data collected in the LEAPS study, it's possible to show that it is unlikely that the differences are driven by this sort of "selection" on unobserved characteristics.

Figure 1.10: When the same child switches from a government to a private school, she learns more…when she switches from a private to a government school, she learns less



1.35 Figure 1.10 shows one of the formal tests in Andrabi, Das, Khwaja and Zajonc (2007), which plots the test-scores of children who stayed in the same school (type) and those who switched schools over the three years. Children who *switched* into a government school learned less in the year of the switch compared to their counterparts who remained in private schools. This pattern is identical for all three subjects. In contrast, children who *switched* to a private school learned more in the year of the switch compared to those who remained in a government school. Since these same children were followed over time, factors like motivation are unlikely to explain the strong positive correlation between learning switches from public to private schools. While ultimately the public-private gap can best be answered in an experimental framework, all the multi-year observational data suggests that the raw gap in the data is very close to the impact of private schools on learning.