WPS4067

Learning Levels and Gaps in Pakistan

Jishnu Das*
The World Bank

Priyanka Pandey
The World Bank

Tristan Zajonc
Harvard University

Abstract

This paper reports on a survey of primary public and private schools in rural Pakistan with a focus on student achievement as measured through test scores. Absolute learning is low compared to curricular standards and international norms. Tested at the end of the third grade, a bare majority have mastered the K-I mathematics curriculum and 31 percent can correctly form a sentence with the word "school" in the vernacular (Urdu). As in high-income countries, bivariate comparisons show that higher learning is associated with household wealth and parental literacy. In sharp *contrast* to high-income countries, these gaps decrease dramatically in a multivariate regression and once we look at differences between children in the *same* school. Consequently, the largest gaps are *between* schools. The gap in English test-scores between government and private schools, for instance, is *12 times* the gap between children from rich and poor families. To contextualize these results within a broader South-Asian context, we use data from public schools in the state of Uttar Pradesh in India. Levels of learning and the structure of the educational gaps are similar in the two samples. As in Pakistan, absolute learning is low and the largest gaps are between schools: the gap between good and bad government schools, for instance, is 5 times the gap between children with literate and illiterate mothers.

World Bank Policy Research Working Paper 4067, November 2006

---

* *Corresponding Author:* Jishnu Das (jdas1@worldbank.org). The discussion of learning in Pakistan draws from an ongoing report on Learning and Educational Achievement in Punjab Schools (LEAPS) authored by Tahir Andrabi (Pomona College), Asim Ijaz Khwaja (Harvard University), Tara Vishwanath (World Bank) and Tristan Zajonc (Harvard University). Funding for the survey was provided by the World Bank through grants from the Knowledge for Change and Poverty and Social Impact Analysis Trust Funds and the South-Asia Human Development Network. Pandey was not affiliated with the World Bank at the time data were collected from schools in India, for which necessary approval for the study was obtained from respective state governments in India. We thank Elizabeth M. King and Lant Pritchett for extensive discussions.

1. Introduction

Three million children were born in Pakistan in 1990, the majority in rural regions. By 2005, these children were at the doorstep of adulthood and 80 percent had completed all the formal schooling they would receive during their lifetime. The better educated arrived at this threshold with numerous advantages—they would earn higher wages in the labor market and enjoy better health outcomes, and the educational and health outcomes of their children would be substantially better than their own. More educated girls, although they may not participate actively in the labor market, would command greater autonomy and marry into higher socioeconomic groups. Even among children with the same level of formal schooling, those with more learning (as measured through test scores) would be at an advantage—a study of adult workers in Pakistan showed that literacy and numeracy had a higher return than primary school completion (Nasir and Nazli 2000).

Part of what determines a person's educational status is the time she spends in school, and this has been the focus of most educational research in Pakistan. Recent studies point to low educational participation, both in absolute terms and relative to the average income of the country (World Bank, 2003, World Bank 2005). With an adult literacy rate of 44 percent, Pakistan compares poorly to the South Asian average (54 percent), and in 2001-02, net-enrollment was 51 percent compared to 83 percent for India, 90 percent for Sri-Lanka and 70 percent for Nepal. For the country's level of income, the forecasted net enrollment rate (based on a regression of primary net enrollment on log per-capita income and the square of log per-capital income for 138 countries) is 77 percent: Pakistan's net enrollment is below what one expects for its level of income. Research also suggests that this low participation has a strong gender component (girls are less likely to go to school), is linked to the distance from schools (more so for girls than boys), and to the wealth and educational attainment of the child's parents (Holmes, 2003, Lloyd, Mete and Sathar, 2005, Alderman, Orazem and Paterno, 2001).

While we do know how many years children spend in school (and what determines the extent of their participation), we know little about how much they learn during this time. Household surveys reveal who joins schools, which children drop out and when, but say less about the basic competencies in language and numeracy that children acquire in school. Data requirements partly drive this emphasis—it is easier to ask about school enrollment than to test children. But it is also driven by educational policy. For example, the Millennium Development Goals call for universal primary education and equal enrollments among girls and boys, but say nothing about learning and competency. Similarly, a number of educational interventions focus on getting children into schools (midday meals, free textbooks, stipends for girls are all examples from very recent history) but do not monitor what these children learn and what effect these interventions have on their learning.

This may be a mistake. Although all children in Pakistan may be enrolled in primary schools by 2015, many may complete their education functionally illiterate and innumerate. International assessments of learning show that many low-income nations perform far below the mean for OECD countries (Mullis and others 2004). Furthermore, improving the quality of schooling may in itself be one of the best ways of getting children into school in the first place—after all, the returns on parental investment depend both on how much a child learns and the number of years she stays in school.

This paper reports on an independent survey of primary schools in rural Pakistan with a focus on the achievement of enrolled children, as measured through test scores. The first part of the paper looks at levels of learning. It gives a sense not only of "average learning" (as in reports from worldwide comparisons), but also points to the *specific* knowledge that children acquire in school. That is, we look at whether children can add 5+14 and whether they can make a sentence with the word "ball". Usually, tests administered in Pakistan take the curriculum as their starting point for test design. When children perform far below the curriculum level, the results are reported as "not achieving the desired competency", leaving the question of what these children actually know unanswered. In contrast, the tests presented here start from the basics (counting and addition in numeracy, alphabets in language) to provide some data on where children are on a knowledge-scale.[1]

The second part of the paper presents associations between children's learning levels and their attributes. We first construct an overall test score from the individual test questions. That is, we "norm" the test by weighting test questions according to difficulty using Item Response Theory, the established procedure in the international testing literature; this is discussed further below.

[1] There are several reasons for using knowledge-specific tests rather than other secondary data. In a number of tests available to us, the data for these three exercises are not available: we do not have access to the original questions used for a number of tests (and in some cases these are proprietary), the secondary data are presented as test-score "totals" rather than question-by-question responses, which does not allow us to construct the item response scores.

Using the weighted test score, we identify learning "gaps" across households (parental wealth and education), children (age and gender), schools (public/private), and geographical locations.

There are three take-home messages. The "bad" news is that children know little relative to what they need to know to function in society and relative to their curriculum. By the end of Grade III, a bare majority have mastered the mathematics curriculum for Grade I. They can add double digit numbers and subtract single-digit numbers but they cannot do much more. They cannot subtract double-digit numbers, they cannot tell the time, and multiplication and division are beyond reach for all except a tiny 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 in the vernacular. For the 40 percent of boys and 50 percent of girls who either never attend school or dropout by the end of Grade III, this is all the learning they will receive from their formal schooling for the remainder of their lives.

The first part of the "good" news is that differences across schools account for at least 50 percent of the overall variation in test scores. While there *are* differences across children from different parental backgrounds (children from wealthier backgrounds or with more educated parents know more), these differences are dwarfed by those across government and private schools and across good and bad government schools. In English, the difference between children in private and government schools is *twelve times as large* as the difference between children from poor and non-poor households after controlling for observed differences between children.

However, not all government schools are the same—the difference in learning between a high-performing and a low-performing government school is *twenty-four times* the difference between children from poor and non-poor backgrounds after controlling for observed child-level differences. Unlike in high-income countries where household factors are critical, the primacy of schools in Pakistan suggests that improvements in learning can be achieved through policies targeted at the school level. In the United States for instance, 20-25 percent of the achievement gap in black-white test-score gap arises from parenting practices at home (Jencks and Phillips 1998). Designing government policies that influence parents to improve their parenting practices are clearly harder than designing government policies that (say) influence teachers to come to school. Some governments can fire some teachers. No government can fire parents.

The second part of the "good" news is that this variation across schools is *not* due to differences across villages; indeed the largest differences are *between schools* in the *same* village, so that there are some good (and some bad) schools in every village. The policy implications are stark. If factors across villages drive variation in learning, district education boards, the educational atmosphere in the village and local returns to education, all of which are difficult to change, play an important role. The findings suggest that there are no "bad" villages and "good" villages; to improve learning, we need focus only on the characteristics of schools.

This paper does *not* address what school characteristics to focus on. A long history of research on learning achievement and school inputs points to technical and methodological issues that severely bias our understanding of the relationship between school inputs and learning achievement and we believe that a responsible answer is best left to further research. Towards the end of the paper, we speculate though, that the large differences across schools reflect differences in teacher commitment and motivation—whether children learn in government schools largely depends on whether their teachers are heroes or zeros.

Very low learning achievement is not a Pakistan-specific problem and requires a serious South-Asia wide policy response. The parallel analysis with data on government schools from the state of Uttar Pradesh in India helps contextualize the results from Pakistan within a broader South-Asian context. Contrary to the popular picture of Pakistan as a "failed-state" (Singer, 2001) and the popular "India-Shining" slogan, levels and patterns of learning are very similar. Children know little and most variation is across schools rather than across households: the gap in learning, for instance, between a child in a high-performing and a low-performing government school is five times the gap in learning between a child with a literate mother and a child with an illiterate mother.[2] Given the commonalities across the two countries, the analysis sheds some light on the necessary ingredients for a system of monitoring performance and learning; this is discussed in

---

[2] The results presented here are thus not meant as a comparison of performance in the two countries. The studies were independently designed with inherently different purposes. Consequently, the tests are different (and not comparable), the studies sample different portions of the relevant population, and the grades tested are different (third grade in Pakistan and fourth grade in India). Nevertheless, there are several reasons for using these data rather than other secondary data. In a number of tests available to us, the data for our analysis are not available: we do not have access to the original questions (and in some cases these are proprietary), the secondary data are presented as test-score "totals" rather than question-by-question responses, which does not allow us to construct the item response scores, and household attributes are not available in some cases. Furthermore, we cannot find *any* tests in India that allow us to replicate the analysis in Pakistan—there have been almost no large-scale tests of children in *rural* private schools (Pratham, 2006 is an exception); furthermore, samples have usually been drawn by picking a single school in every village so that comparisons across schools within the same village are not possible.

Section 6.

The remainder of the paper is structured as follows. Section 2 reviews what we know about learning from the existing literature. Section 3 describes the sample and data. Section 4 outlines what children know in language and numeracy, and section 5 presents a decomposition of the variation in learning and a discussion of the patterns that emerge. Section 6 concludes with three puzzles and questions for future research.

2. Literature Review

Since 1984 various agencies have conducted (at least) 19 assessments of primary education in Pakistan (Table 1). These efforts gained momentum after the World Declaration on Education for All (EFA) in 1990 and again with the joint UNESCO-UNICEF global initiative for Monitoring Learning Achievement in 1996. Two major assessments—the World Bank's Primary Education Project in 1984 and the BRIDGES project (Harvard Institute of International Development) in 1988–89 set the stage for a slew of surveys in the following decade. Both these projects assessed a large number of students (3,300 for the WB project and 11,000 under the BRIDGES) using the same test.

There were some noteworthy results (Shah 1984). In three tested provinces, girls scored higher in sciences while boys scored higher in mathematics. However, *all* groups performed poorly, leading the author to conclude that schools should focus on ensuring that students acquire basic competencies rather than increasing "peripheral luxuries" in the curriculum. In 1989, when the BRIDGES project retested children with the *same* test instrument, they found that performance had declined in the intervening five years. A clear message emerged: learning in Pakistani schools was a problem, and the experience of the last five years was an early warning that time alone would not yield better educated children.

Testing activity in Pakistan increased significantly during the 1990s and largely replicated the disappointing results of the 1980s in terms of children's learning. Some new results about the variations in school performance emerged. Different types of schools produced very different results, although there was little correlation between achievement scores and per-student costs (Mirza and Hameed 1994).

Not surprisingly, different agents had very different perceptions of what was wrong with the system. Head-teachers criticized the standard of teaching, teacher absenteeism, and the lack of parental support at home. Teachers blamed low performance on the lack of adequate facilities and physical resources. Learning coordinators focused on the poor physical resources, teacher absenteeism, poor school administration and the lack of cooperation between teachers and the community (Arif and others 1999; Haque and others 2000, The Bureau of Curriculum Development and Extension Services, NWFP 1999; and the Bureau of Curriculum and Extension Centre, Balochistan 2000). One study at least, *did* find that teachers matter—the Multi-Donor Support Unit for the Social Action Program in 1995 collected information on 11,563 students in 527 schools throughout the country and reported that teacher performance, as measured through test scores was positively correlated with student scores. This effect was particularly pronounced in rural areas where trained teachers accounted for a 12-percent increase in student scores. Finally, private schools were doing better. Both in the Multi-Donor Support Unit test and a later test conducted by Action Aid, children in private schools performed significantly better than those in public schools, and in all subjects.

Despite the somewhat bewildering array of tests and outcomes, the lack of coordination between different testing bodies severely limits the ability to draw meaningful inferences. Consequently, over the past 15 years we do not know whether performance has improved, whether provincial differences in achievement have declined (Khan 1995), or whether increasing enrollments have had an impact on learning. As Table 1 shows, the available documentation is simply inadequate.[3] Even an educated guess at the answer to these pressing questions is impossible. Similarly, while some of the tests document gender differences (with differing results for different subjects), there has been no attempt to examine the variation in test scores along other dimensions—do children in rich families perform better than in poor families? What about in poor and rich villages? Or in families with literate/illiterate mothers? These gaps, which are the mainstay of educational policy in OECD countries, remain understudied.

The current state of knowledge regarding learning achievement in India is similarly hampered by a lack of coordination and systematization of testing tools, sampling, and reporting protocols. A

[3] In Assessing Learning Achievement, the UNESCO (2001) specifies that a "good survey" should include at the very minimal, careful documentation of the sampling methodology (with standard errors), as well as instrument construction and validation. Although we are currently in the process of trying to contact each individual author to update this review, in a number of cases, instruments have been discarded and published documents remain the only source of information. These documents provide little more than the information summarized in Table 1.

number of studies since the early 1990s report low achievement levels at the terminal grades of primary school. Examples include a large national study by the National Center for Educational Research and Training (NCERT) in 1994, which found that children scored an average of 47 percent in language and 41 percent in mathematics (Shukla and others 1994) and state-wise studies with smaller samples in Bihar, Tamil Nadu, Delhi and Madhya Pradesh (Bashir 1994; Hasan 1995; Govinda and Varghese 1993; Aggarwal 2000). A recent survey tested children from almost all districts from India and reports low levels of learning with substantial variation across states: 52 percent of children between the ages of 7 and 10 could read a small paragraph with short sentences at Grade I difficulty levels, 32 percent could read a story text and 54 percent were unable to divide or subtract (Pratham, 2005).

There is little consensus across these studies on the nature and extent of socio-demographic variation in achievement levels. In some studies, boys and students from upper castes perform better (Shukla and others 1994; Aggarwal 2000; Aikara 1997), while in others there are no gender differences (Hasan 1995). As in Pakistan, private schools outperformed government schools (Kingdon 1996; Tooley and Dixon 2006; ASER 2005). Most studies find that home background and school factors were both important in influencing students' learning (Kingdon 1996; Govinda and Varghese 1993).[4]

Thus, even though there has been a greater number of studies measuring learning in India, meaningful comparisons can be made neither across studies nor over time because of differences in sampling methodology, the tests used and the child characteristics collected as part of the survey.[5] As in Pakistan, there is no comprehensive picture of variation in achievement across relevant socioeconomic dimensions. For instance, the lack of consensus on gender and caste differences could reflect different test instruments and different sampling methodologies. In short, in both countries, it is hard to draw any straightforward inferences on either changes in learning over time or the gaps in learning across different groups.

---

[4] Although Kingdon (1996) looks at the impact of various student characteristics (wealth, gender, and parental education) and school characteristics on learning, the data is restricted to a sample of urban schools from one city in north India.

[5] Even though ASER (2005) allows one to compare learning achievements across districts/states in India at a point in time, there is no information collected on household characteristics. If repeated at regular intervals with a standardized instrument, this initiative would be a welcome addition to what we know about learning in Indian schools.

3. The Study Sample and Tests

3.1. Pakistan

The Learning and Educational Achievement in Punjab Schools (LEAPS) project in Pakistan carried out the tests analyzed here, with an eye toward piloting new educational interventions. Previous work by the project team documented a sea change in the delivery of education in the country (Andrabi and others 2004, 2005, forthcoming). Despite concerns about high and rising enrollments in religious schools, or madrassas, the team estimated that less than 1 percent of children are enrolled in madrassas and found no evidence of an increase in religious enrollments during the 1990s. Instead, there had been an explosion in *private schools* during the 1990s, which shows no signs of slowing down.[6] The LEAPS project was conceived as a multiyear study to pilot potential interventions and document the efficacy of various policies in an environment where villages had both private and public schools. The results presented here draw from the baseline survey of this project.

The team purposively chose three districts in Punjab on the basis of an accepted stratification of the province into North, Central, and South. Within these three districts—Attock in the North, Faisalabad in the Center and Rahim Yar Khan in the South—126 villages (42 in each) were chosen randomly from a list frame of all villages with at least one private school. In subsequent visits, 14 villages were eliminated and for the remaining 112 villages, the team mapped *all* existing schools, defining catchments as a 15-minute circle from the boundary of the village in Attock and Faisalabad (where the villages are concentrated) and a 30-minute circle in Rahim Yar Khan (where they are more spread out).[7] The mapping exercise yielded 828 schools, for an average of 8 public and private schools in every village.[8]

All children enrolled in third grade in these public and private schools were tested in Urdu,

---

[6] Of all new private schools in 2000 50 percent were set up in villages, were relatively cheap (with a fee to median GNI ratio of 2.5 percent compared to 9 percent for the United States), and their largest growth sector was the rural poor. By 2010, the team projected that in Punjab province (the largest in the country with 50 percent of the population) one-half of all villages would have at least one private school.

[7] In the villages that were dropped, either the population in these villages was too large (they were cities) or the private schools had shut down.

[8] This simple mapping exercise dramatically alters the vision of a village with a single school. The large number of schools is not because of the sampling strategy; a mapping exercise in the rural parts of the Sindh province (one of the poorest in the country) shows that there is an average of *five* government schools in every village. Anecdotal evidence from India suggests similar numbers, although there is currently no analysis that allows us to estimate the average number of schools (government and private) in a village.

Mathematics, and English.[9] The test was developed one year before it was administered and piloted extensively in rural and urban schools (Andrabi and others, 2002, for the test validation and documentation).[10] Two members of the LEAPS team monitored students during testing, and the children's teacher was not allowed in the same room. Finally, for a sub-sample of 10 randomly selected children in every school, we also completed a child questionnaire with basic information about the child's parents, ownership of certain assets, the distance to school and the number of siblings. We thus have test scores in three subjects for over 12,000 and combined test-scores and child characteristics for 6,241 children.

The sample of children is representative of children enrolled in third grade in a village with a private school in Punjab.[11] In 2000, 30 percent of all villages in Punjab province (and 50 percent of the population, since private schools disproportionately locate in larger villages) satisfied this restriction; it is likely that this has increased by another 10 percentage points by 2005. The particular strategy affords several advantages, in that we are able to compare performance across schools *in the same village*, and therefore understand the extent to which learning is driven by school-related factors that vary across villages and across schools in the same village.

3.2. India

The comparison sample from India is based on data from 31 districts in Uttar Pradesh and 9 districts in Madhya Pradesh (Table 3). Within each district, one "block" (several blocks comprise a district) was chosen randomly and within each block, five "Gram Panchayats" (henceforth GP) were chosen randomly.[12] The choice of districts was purposive, rather than random and sought to examine the effect of past institutions on the quality of public services delivered (Pandey 2005). A second set of districts were those that lay on either side of the border of MP and UP.

---

[9] Children from Grade 3 were chosen so that we could follow them for 2 years (till the end of Grade 5) before the transition to middle and secondary schooling. Since there are fewer government middle and secondary school, this often means that children study outside the village, which makes the construction of a panel of test-scores on the same children a daunting task.

[10] The test used the K-V curriculum as a starting point but did not use the K-V textbooks for the test-design. In the pilot, we noticed that children could verbatim repeat paragraphs from a textbook "as-if" they were reading them; in fact they were not—through the year, they had memorized the paragraph verbatim from the textbook. The test-design used here thus tests the curricular concepts developed by the Ministry of Education.

[11] Punjab is economically and educationally the most dynamic province in the country with systematically higher enrollments, lower dropouts and greater income than households in other provinces. Within Punjab, villages with private schools tend to be larger, more educated and wealthier than those without (Table 2).

[12] A gram panchayat (GP) is an administrative unit between a block and a revenue village. A GP comprises 2-3 revenue villages on average and the local village level government is formed at the GP level.

In each GP, one government school at the primary level (up to fifth grade) was followed from August 2002 until February 2003. In cases where there were two or more public schools providing primary schooling in the GP, one was randomly chosen. Each sample school was visited unannounced seven to eight times during this period by field investigators. In each of these visits, the survey team recorded the presence (or absence) of each teacher and his/her teaching activity, defined broadly to include teaching, writing on the board, supervising written work and keeping order in classroom.

Further, in every school a randomly chosen sample of ten children in fourth grade was followed during the survey period and, toward the end of the survey, the sampled children in the fourth grade were tested in literacy (Hindi) and numeracy; the test was based on the starting material of textbooks used in these schools.[13][14] The students were tested near the completion of their fourth grade. Since students are randomly chosen and the sample covers 40 districts, the sample is somewhat representative of public school children in rural north India.

4. What Do Children Know?

Table 4 documents the state of learning among children close to completing the third grade in Pakistan with a comparison of similar questions for the tested children in India. In Mathematics, children tested in both India and Pakistan can add single and double digit numbers. But results for other mathematical operations are poor: In Pakistan, a bare majority (65 percent) can subtract single-digit numbers (8 – 3), and less than one-third can subtract 3-digit numbers; in India 53 percent of all tested children could subtract 112 from 640 (640–112). Among the Pakistani children, 59 percent could multiply two single-digit numbers (5 x 4); in India 49 percent could multiply two-digit numbers (45 x 12) and 41 percent could perform a simple division exercise (936 ÷3). The last column in Table 4 shows the grade-level equivalent using the curricular standards developed by the Pakistani Ministry of Education (Government of Pakistan 2000). By the end of Grade III, a bare majority have achieved Grade K & I competencies and less than 20 percent are anywhere close to what is expected of them according to the curriculum.

[13] In public schools in UP and MP, Hindi is the only language taught between first and fifth grade.
[14] To ensure that enrollment does not decrease below an official norm, teachers in rural public schools sometimes maintain inflated school registers that include the names of children who have dropped out or have left to go to another school. Before randomly choosing the ten children per school, these children were excluded from the list-frame.

These results measure the extent of learning after spending four years in the education system in Pakistan and five years in India. Although the test went on to ask about division and simple fractions, these questions were simply out of the grasp of the vast majority of children. There were five students who scored more than 90 percent; but that only 19 percent knew how to divide 384 by 6 (Pakistan) suggests that these children were truly exceptional. In terms of the curriculum standards, most children were just at the benchmarks proposed by the curriculum wing for children in first grade in the case of Pakistan and second grade for India.

The results in English (Pakistan only) were equally poor (Table 5). Most children (86 percent) could recognize and write an alphabet when spoken and 70 percent could fill in a blank alphabet in a sequence (D _ F) but less than 50 percent could complete basic words like BALL or FLAG when given a picture and asked to fill in 2 blanks (BA_ _) or a single blank (FLA_). Indeed, for the latter only 29 percent were able to correctly fill in the blank. Again, the test went on to check children's ability to construct more complicated words and sentences as well as paragraph comprehension, but as in Mathematics, these concepts were too advanced for 90 percent of the children tested.

English is conceivably a language for the elite and to enjoy the benefits of reading and writing a basic knowledge of the vernacular may be sufficient. Unfortunately, the scores from the Urdu test (Pakistan) suggest that children are equally disadvantaged in the vernacular as they are in English. A minority of children can form basic sentences in Urdu given a word; only 31 percent can form a sentence using "school." Levels are similar in India—only 30 percent of the tested children could construct a sentence in Hindi using a commonly used word like "message" and a minority (46 percent) was able to use the word "house" in a sentence.

These numbers suggest that the average citizen in Pakistan (or India) will be in no position to compete in a rapidly globalizing world. Although international comparisons are inexact, in the case of Mathematics the recently completed Trends in International Mathematics and Science Study (TIMSS) gives a sense of where the children we tested fall in the global distribution. For instance, 72 percent of fourth-graders correctly answered the question 15 x 9 internationally compared to the 50 percent who could answer 32 x 4 in Pakistan (the former is a tougher question since it requires a "carry-over"); 73 percent could graphically represent simple information (bar charts) internationally compared to less than 10 percent in Pakistan. Both in India and Pakistan, average levels of learning are at the bottom of the international chart.

We might be tempted to counter the observation that only 50 percent of third-graders can compute 32 x 4 by pointing out that in Australia, New Zealand, and Norway, less than 45 percent of fourth-graders could compute 15 x 9. This is a near-sighted argument. The key difference between Pakistan and India and the low-performing OECD countries is the length of time that children stay in school. Household data suggest that 40 percent of all children in Pakistan and 42 percent in India will have completed their formal schooling history by third grade (Pakistan) and fourth grade (India). In contrast, nearly all children in the three OECD countries will complete the twelfth grade. And low achievement observed in fourth grade (largely due to a curriculum designed to enhance creative thinking at the lower grades) will be more than compensated by accelerated learning later on—by the time these children are in seventh grade, they score *above* the international average. By comparison, close to one-half of the 3 million born in Pakistan will leave school unable to add, subtract, multiply or divide; unable to read and write simple sentences in Urdu; and unable to read a short word like "BALL" in English.

5. The Educational Gaps

A second basic characterization of the data decomposes these averages into differences across population subgroups. Subgroup decompositions—say along income or ethnicity—have played a critical role in policy and educational legislation in the United States. For instance, the black-white test-score gap—17-year-old blacks score roughly one standard deviation lower on the NAEP math test than their white peers (Hanushek 2001)—has spurred a large body of research and widespread legislation on financial equalization across public schools. Equally, the gender gap in mathematics and science has been the focus of much recent discussion. There are two issues—one methodological and one technical—we address prior to a discussion of the test-score gaps.

On a methodological note, a key question is the choice of subgroups. Should we look at rich and poor families or gaps across girls and boys? Or should we concentrate on differences across the landed and landless? Following established norms and new literature on the equality of opportunity, we focus on two subgroups. Our first set of subgroup decompositions are for *immutable* characteristics such as gender, and arguably parental literacy and income. On moral grounds, we may believe that differences across children in these attributes should not affect their opportunities in life.

The second set of decompositions is along geographical lines. A guiding principle of most public education systems is that of standardization and equity. Indeed, in Pakistan, as in most other countries including India, *inputs* into the educational system are driven by this principle—salaries for teachers for example are dictated almost entirely by experience rather than ability or motivation. The principle of "equality" may also imply that *outcomes*, and in this case, a child's learning be the same, regardless of the village she is born in or the public school she attends. Thus, in addition to child characteristics, we also look at differences across villages and across schools within the same village. Given the focus of the LEAPS project in Pakistan, we also present the gap between private and government schools. In India, the data do not allow us to look either at the difference across private and government schools or across schools in the same village (there is currently *no* dataset that allows us to replicate the Pakistani analysis). A useful characterization though, is the difference across districts; despite the lack of comparability with the Pakistani geographical decomposition the results have independent implications for the design of future monitoring systems.

The technical issue is what test scores to use in the subgroup decompositions. One problem with the most natural candidate—the percentage of questions correctly answered in the test—is that the measured gap depends on the structure of the test. An example illustrates. Consider a test where there are ten questions on single-digit addition, five questions on double-digit addition and five questions on fractions. A child who knows single-digit addition will score 50 percent on the test, one who knows double-digit addition scores 75 percent and one who knows fractions 100 percent. Using the percentage of correct questions leads us to conclude, somewhat erroneously, that the gap between the first and the second child (25 percent) is identical to that between the second and the third child (also 25 percent), whereas its likely that the knowledge-gap between the second and third child (who can correctly answer questions on fractions) is a lot more than between the first and the second. Studies of learning achievement in low-income countries typically report results using a "standardized" version of the percentage correct (i.e., subtracting the mean and dividing by the standard deviation), but this approach suffers from the same problems.

The established norm in all test-score reporting in the United States or in international tests such as the TIMSS is to *weight* questions by their difficulty. Continuing with the example above, if the single-digit questions are weighted as "very easy", the double-digit as "easy" and the fractions as "very hard", the weighted averages scores the child who completes the fractions much higher compared to one who completes the double-digit addition. The process of assigning these weights and computing weighted test scores is known as *Item Response Theory*, and we use these methods to normalize our test results.

Knowledge gaps are then presented on a "knowledge scale", where the average child scores 500, the worst child scores 0 and the best 1,000. The distribution has a standard deviation of 150, so that a child who scores 350 is one standard-deviation below the average. In addition, the weighting ensures that the difference in "knowledge" between a child who scores 300 and one who scores 400 is *identical* to that between a child who scores 700 and one who scores 800.

Table 6 shows what these different knowledge scores mean in terms of actual learning. For instance, 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 200 (2 standard deviations below) 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.

5.1. Gaps due to Immutable Characteristics: Household Characteristics

Part of the gap in learning between children from poor and rich households could arise due to variation in parental income. However, parents of children in richer households are also likely to be more educated and the children are more likely to be enrolled in private schools. We thus present two types of learning gaps—unadjusted and adjusted—which differ in how they account for observable differences between children with different family backgrounds.

The *unadjusted* gap is the difference in the knowledge score between children from different family backgrounds. The *adjusted gap* is based on a regression of the knowledge score on observable attributes of the child and the household they come from (Table A1, Table A2). This regression includes the literacy of the parents, the school that the child is studying in (school fixed-effects), and the gender and age of the child. The *adjusted gap* is the coefficient of the relevant attribute in this regression. The interpretation suggests the extent to which two children with the same observable characteristics (controlled for in the regression) and studying in the same school still exhibit a difference in their learning due to (say) differences in parental wealth.

The *unadjusted gaps* are consistent with patterns we may expect *a priori*.

- Children from richer families report higher test scores and the differences are larger for language and lower for mathematics: richer children report test scores that are 70 points higher on the knowledge scale in English (just below 0.5 standard deviations) and 40 points (approximately 0.25 standard deviations) higher on the knowledge scale in Mathematics (Figure 1).[15]

- Children whose parents are literate report higher scores. Consistent with a large literature on the relative effects of father's and mother's education, the differences between children with literate and illiterate mothers are higher than those between children with literate and illiterate fathers. There are some differences in the patterns across subjects (Figure 1). The largest gaps across literate and illiterate parents are for English (48 points for fathers and 58 points for mothers) and the lowest for Mathematics (33 and 38 points, respectively).

- Females report lower scores than males for Mathematics (10 points or 0.06 standard deviations), but they are better in English and Urdu, and are worse in Hindi in India. Again, these gaps are small, accounting on average for a 10-point difference (English) and a maximum difference of 28 points in the case of Urdu (Figure 1).

- The results from India largely mirror those from Pakistan (Figure 2). Children from richer households report higher scores, but the difference is lower than in Pakistan (30 points or 0.2 standard deviations in Hindi and 28 points in mathematics. Children with more educated parents also report higher scores, although in India the gaps are *larger* for Mathematics than Hindi for differences by father's literacy and the same for differences by mother's literacy—these gaps are small in absolute magnitude compared to those in Pakistan (10 points for father's literacy in Mathematics and 0 for Hindi, and 20 points for both subjects for mother's literacy). There is some evidence of gender-gaps in India, although, unlike Pakistan, girls report worse scores for both Mathematics and English.

These gaps disappear or reduce sharply (with some exceptions) in a multivariate context, and once we control for the school that the child is attending. That is, the gaps between the rich/poor

---

[15] In Pakistan, wealth is measured through an asset index following Filmer and others (2004). In India, households are classified according to whether they have above or below median land holding (the latter includes landless households).

or children with literate/illiterate parents arise primarily because children are going to different schools—if we compare two children from different households studying in the *same* school, these gaps are small and insignificant (Figure 1, Figure 2).

- For parental wealth, the learning gaps are never more than 15 points or 0.10 standard deviations (Urdu, Pakistan) once look at the difference between two children in the same school with different parental wealth and control for other child and parental characteristics in a multivariate regression.

- Similar patterns hold for differences by father's literacy and mother's literacy for Pakistan; in the case of India, the effect of father's literacy drops to 0 for Mathematics and –5 points for Hindi, but mother's literacy remains as important.

- In contrast, the adjusted and unadjusted gender gaps are very similar with girls performing worse in both subjects for India, and better in all subjects for Pakistan, particularly for Urdu and English.

The data from Pakistan offer a unique opportunity to shed some light on the differences between public and private schools in rural areas—a topic that is often debated and remarked upon in the popular press and media reports. One popular perception is that of private schools as "fly-by-night" operators that survive because parents are illiterate and easily influenced. In contrast to this view, the data show a large and significant difference in both the unadjusted and adjusted gaps for all subjects between private and public schools (Figure 3). These gaps are largest for English, where private schools outperform public schools by 150 knowledge points, or almost one standard deviation, but they are also large for Urdu (100 points) and mathematics (75 points). Surprisingly, there is almost *no decrease* in the gap once we control for family wealth, parental literacy, the age and gender of the child and the village that the school is located in.

Three simple comparisons highlight the dramatic difference in scores between public and private schools:

- The adjusted gap between public and private schools in English is *12 times* the adjusted gap between rich and poor children.

- The adjusted gap between public and private schools in Mathematics is *8 times* the adjusted gap between children with literate and illiterate fathers.

- The adjusted gap between public and private schools in Urdu is *18 times* the adjusted gap between children with literate and illiterate mothers.

That across the two countries and in all subjects, there is a clear decline in the relevance of socio-economic characteristics once we control for the school that the child attends suggests that it is the school that matters. Clearly in the case of Pakistan, the learning-gap between private and public schools dwarves differences across children from different backgrounds. This is by no means a foregone conclusion: in most studies of high-income countries, the ability of differences across schools to explain differences in test-scores is *smaller* than that of differences across households (OECD 2005).

A priori¸ it also appears that there are 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. These observations suggest that gaps between the best and worst schools within each type may be just as important as the gaps between each type.

5.2. Gaps across Schools and Geographical Entities

This section confirms the substantial differences across schools, with tremendous variation in learning achievement across the worst and best government and private schools. Continuing with the attempt to contextualize the results from Pakistan within a broader South-Asian context, we present similar results on differences across schools for India. The patterns, though, are not strictly comparable due to different sampling strategies (in India there are multiple districts, but one sampled school in every village; in Pakistan there are many sampled schools in every village, but only three districts). For the Pakistani data, we show that there are bad and good schools in *every* village, so that most variation in learning is driven by differences *across schools* in the same village. In India, there is currently no data that allows us to analyze differences across schools within the same village. We confirm though, that there are large differences across schools and that differences across districts are important. Villages in the same district report similar scores, but learning achievements are very different across sampled districts.

Figure 4 highlights these differences. For Pakistan we ranked villages by average mathematics scores and took the 5th best and the 5th worst village (with a total of 112, this corresponds roughly to the 5th and 95th percentiles of the distribution). The 5th worst village had 10 schools; the 5th best had four. We then computed average mathematics scores for every school in each of these two villages and plotted these scores. Several patterns are remarkable.

First, although there is considerable difference in average mathematics scores in the two villages this 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 are in the region of "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.

Second, government schools clearly under-perform private schools (marked as triangles) 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, four government schools are *better* than the worst private school, and the top performing school is government rather than private. 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.

Third, the differences in math scores across schools in the same village are orders of magnitude larger than the differences in math scores across children with varying socioeconomic backgrounds. In the "bad" village, the difference between the worst and best-performing school is 500 knowledge points (more than 4 standard-deviations) and in the "good" village more than 400 knowledge points (3.3 standard deviations). The single largest unadjusted gap in mathematics score by socioeconomic status is across children from less and more wealthy backgrounds with the wealthy reporting mathematics scores that are 40 knowledge points higher on average. The unadjusted difference between the best and worst school in the *same* village is *12 times* this difference in the "bad" village and *10 times* this difference in the "good" village.

For India, we cannot repeat this exercise at the level of the village. However, a similar comparison at the level of the *district* (recall that in India one school was picked at random from each village, but five gram panchayats were surveyed in each of 30 districts) is instructive, both to understand whether the large differences across schools are particular to Pakistan and to understand the extent of district-level variation in the data (something we cannot say much about in Pakistan given that 3 districts were sampled). For a comparable exercise, we first ranked districts according to the average mathematics score of children in the district. Corresponding to the 5th and 95th percentile of the distribution, we then chose the 2nd best and 2nd worst district and computed village-level school scores (with one school sampled in every village, this is identical to school-level scores) for each of the five villages within the district. These are plotted in the first panel of Figure 4.

Similar to the data from Pakistan, differences across schools are large, both in absolute terms and when compared to differences by socio-economic characteristics. In the "bad" district, the difference between the best and the worst school is 200 knowledge points and in the "good" district 104 knowledge points. This is *8 times/4 times* the unadjusted wealth-gap and *10 times/ 5 times* the unadjusted gap by maternal literacy in the good/bad district respectively.

Unlike comparisons across villages in Pakistan, in India all the schools in the "bad" district are indeed worse than *all* the schools in the "good" district. There is no overlap, and the difference between the worst school in the "good" district and the best school in the "bad" district is 134 knowledge-points (almost one standard-deviation) or six times the unadjusted gap between children with literate and illiterate mothers. Neither is this a characteristic of the extremes of the distribution—indeed, out of 40 districts, we need to go to the 6th best and the 8th worst districts before we find a single overlapping school (roughly the 85th and 20th percentiles). Clearly, the language of "bad" and "good" *geographical* regions with respect to learning resonates in India, but not in the Punjab province of Pakistan, at least in terms of *villages*.

We can replicate this exercise using the data from the full sample rather than two villages in each country. One way the educational literature has approached this issue is through a "variance-decomposition" exercise, which asks how much of the total variation in test scores can be explained by district, village, and school characteristics, with the residual attributed to differences across children and noise in the data. A simple version of this ANOVA decomposition 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.

For both Pakistan and India, 50 percent of the total variation can be explained by variation across schools. District, village, and school effects are stronger for English (Pakistan) and Hindi (India) where they account for more than 60 percent of the total variation; they are weakest for Mathematics (India) and Urdu (Pakistan), though they still account for one-half of the total variation. In fact, 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. This 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).

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. In contrast, district effects in India are large. One-third of all variation in test scores is governed by the district that the child is studying in; the effect of the specific Gram Panchayat in a particular district is somewhat smaller (around 15 percent).

Figure 5 and Figure 6 present another way of looking at the same problem. We first regress test scores on village dummies and then plot the residuals—this is a measure of how much of the variation is accounted for by villages. We then add in family characteristics and plot the residuals again; finally we repeat the exercise with school dummies (in India, we proceed with district dummies, village dummies, and family dummies). To the extent that schools or villages explain a large portion of the variation in the test-score data, we expect that residual plot to be more "concentrated" once the appropriate dummies are accounted for. So, if districts matter a lot, we expect the residual plot from a regression of test scores on district dummies to be "tighter" than the distribution of all test scores.

A familiar story emerges. In Pakistan, the residual plot is slightly more concentrated once we include village dummies. Augmenting village dummies with child characteristics (wealth, parental literacy, age, and gender) smooth's out the distribution around the left-tail but a further concentration is hard to pick up through a visual inspection alone (although the variance does drop somewhat). Once school dummies are included, the picture changes dramatically—for all three subjects there is a substantial reduction in the variance of the residual plot and these effects are most pronounced for English and Mathematics, and less so for Urdu. In India, the tightening happens early on at the level of the district, and further controls for the village/school and family characteristics does decrease the variance, but there are no dramatic reductions (as an aside, note that the Hindi test was too hard—there are a large number of children at the extreme left of the distribution, and these seem to be concentrated in a handful of districts).

Based on these figures, Table 7 reports the 5th percentile, the median, and the 95th percentile of test scores, averaged at the school level for English, Mathematics, and Urdu, and for public and private schools in Pakistan.[16] While patterns are somewhat different across subjects, three messages emerge:

- In all subjects, the best schools are private. The "top" government schools come closest to the "top" private schools in Mathematics; they are further behind in Urdu and significantly behind in English, where the median private school is roughly equivalent to the top 5th percentile of government schools.

- The worst schools are always government schools, and by large margins in all three subjects. For mathematics and Urdu, the worst private schools outperform the worst government schools by 100 knowledge points (1.5 to 2.5 times the unadjusted wealth gap in Urdu and mathematics, respectively). For English, the equivalent difference is 300 knowledge points or more than 3 times the unadjusted wealth gap.

- The variation in private schools is always less than the variation in government schools. That is, if we look at the difference between the "top" performing and "worst" performing government schools, this difference is greater than the equivalent difference for private schools. For Mathematics and Urdu, the range of knowledge points between the bottom 5th percentile and top 95th percentile of private schools is 300 compared to 350 for government schools; for English its 200 in private schools compared to 450 in government schools.

[16] These decompositions should be viewed as a heuristic device only, since quantiles can only be recovered if the first stage regression encompasses the entire distribution rather than the expectation. Regressions are mean-unbiased, but do not necessarily allow us to recover unbiased quantiles presented in Table 7.

5.3. Discussion

Improving educational outcomes remains one of the key development challenges for Pakistan. While efforts to raise enrollment and retention rates must continue, the results reported here reiterate that enrollment and retention are only one measure of education. Enrollment and retention do not guarantee actual learning. Overall achievement is low and the gaps in achievement across schools are large. In contrast to popular writing on education in Pakistan and India, the levels of learning and the structure of the educational gaps are *very similar* in the two countries; if corroborated by representative and standardized tests, this would suggest that increasing learning is a difficult challenge across both countries.

In terms of total learning, children in the third grade in Pakistan can barely read and write or perform the simplest arithmetic operations such as counting and addition. More difficult but equally crucial skills such as forming a complete sentence given a word and solving more complicated multiplication and divisions problems are out of reach of all but a few children. Households characteristics are associated with achievement along the dimensions expected—children from richer and more educated families perform better—but these differences largely vanish after controlling for all the household characteristics and, in particular, the school in which the child is enrolled.

The most striking gaps arise not between children but between schools in the same village. Private schools score significantly higher than government schools, even after controlling for characteristics of the student population. But there are large variations within government schools. Even in villages with a great number of poorly performing schools there is usually one school that performs well. While the best schools are not always private schools, the worst schools are almost exclusively government schools. These dismally performing government schools often record achievement levels so low that the pupils tested must have virtually no cumulated knowledge or skills after four years of education.

At the very least, the findings point to the need for consistent evaluation of learning in Indian and Pakistani schools with some documentation of the learning-gaps across households and schools. The literature review highlights how different methodologies, the non-comparability of tests given at different points in time, and lack of uniformity in the data collected make it difficult to discern trends in learning over time and the changes in learning-gaps across households and tests. While this is increasingly recognized in both India and Pakistan (the ASER tests by Pratham in

India and discussions in Pakistan about the National Assessment and Educational System), a uniform and easily accessible methodology would allow both official and unofficial bodies to produce consistent data over time and regions.

A systematic assessment would enable comparisons across households, children, schools and geographical locations. The findings from Pakistan show large differences across government schools in the same village. Typically studies test children from only one school in every village. While this is valid when there are a small number of schools, administrative data on school-supply in Pakistan suggest that even in rural Sindh, a province with one of the worst educational outcomes, there are close to *six* government schools in every village. To enable meaningful comparisons over time, we suggest that periodic tests satisfy three requirements.

- First, design the test to be norm, rather than criterion-based (that is, the test is informative about where children are on an overall scale, rather than designed to check whether children have achieved one type of competency). One problem with tests in low-income countries is that they are designed around a curriculum suitable for the grade-level the child is in. Typically, children are three to four grade-levels below the curriculum. Consequently, tests are relatively uninformative about what children actually know. Norm based tests, with a list of questions available for test-designers, allow meaningful comparisons of what children know even when they perform below their grade-levels.
- Second, collect basic child characteristics along with the test. At the minimum, data on parental education, some indicator of wealth (children in third grade were able to accurately answer questions about household assets such as radios or televisions), child gender and age should be available to researchers so that gaps across these observable dimensions can be monitored carefully.
- Finally, allow for variations across districts *and* across schools in the same village in the sampling strategy. Our findings suggest it is necessary to include at least two to three schools in every village, two to three villages in every district and a large number of districts. The Pakistan study is the first that tests children in multiple public and private schools within the same village, but suffers from the small number of study-districts; conversely, tests in India assess learning across villages but seldom sample more than one school in every village.

On a substantive note, large differences across schools are consistent with two (broad) explanations. These differences could reflect student sorting rather than school quality. Although the adjusted gap across schools controls for observable socioeconomic characteristics, variables unobserved by researchers (such as child-ability/motivation or parental investment) play an equally important role. If children in private schools are more "able" than in government schools, this would explain a fraction of the observed gap in Pakistan. Similarly, differences in unobserved child characteristics are consistent with the variation *within* government schools in both India and Pakistan—these could reflect (unobserved) differences in the student-body rather than school quality.

A similar question of "selection on ability" confounds the interpretation of the small gender gaps in test scores. Given large gender gaps in enrollments (particularly in Pakistan) there are two potential interpretations for why similar gender-gaps do not arise in test scores. One possibility is that parents send only the most "intelligent" or "able" girls to schools, so that the average (unobserved) ability of an enrolled girl is higher than that of an enrolled boy. A second possibility is that the gender gaps in enrollment do not translate into differential household inputs into schooling across boys and girls once they are enrolled. That is, parents differentially enroll girls less than boys, but once they have decided to enroll them, spend as much time and money on both.

The observed patterns could also reflect variation in the quality of schooling rather than in the student body. These could either reflect school facilities and physical infrastructure, management styles and the influence of the head-teacher, or the teaching imparted by teachers themselves. Although the educational literature is generally inconclusive about the importance of school facilities and head-teachers, there is consistent evidence that teachers matter (unfortunately, we don't know *what* it is about teachers that matter).

In ongoing research, we document that variation in physical infrastructure across schools does not follow the same patterns as learning. In Pakistan, there is a marginal difference in infrastructure across private and public schools. In India, there is no difference in learning in schools with good and bad infrastructure; neither are there systematic differences in learning across schools with high and low student-teacher ratios (Figure 7). In stark contrast, the presence or absence of teachers in school and what they do when they are there is strongly associated with test scores—in India, schools where the average fraction of teachers actively engaged in teaching was high (more than 50 percent) report 75 more knowledge points in both Mathematics and Hindi tests (Figure 7). In Pakistan there is a clear difference in teacher effort measured through absences in public and private schools: the former are absent 3.2 days a month compared to 1.8 days a month for teachers in private schools.

Similar results regarding learning and teacher effort (usually measured through teacher absence) have been documented in recent research (Duflo and Hanna 2005; Pandey 2005). If these learning gaps truly reflect differences in teacher effort, they raise important puzzles and questions that may govern how (and whether) these two countries can appreciably improve the quality of human capital.

6. Three Puzzles and Questions for Further Research

The first puzzle is the variation across government schools. Is this governed by teacher characteristics and effort? In private schools, Andrabi and others (2006) document that pay is linked to performance. Teachers who are more competent (as measured through their test scores and education) and teachers who are more absent are paid less. In contrast, government teachers in both countries are paid according to a fixed-salary scale that rewards experience and training, but little else. And they cannot be fired. Given that there are few differences in performance incentives for government teachers, why are test-scores so different across government schools?

There are two possible explanations. One is that teachers respond to the environment that they are in. A recent review of teacher absenteeism by Chaudhury and others (2006) finds that teacher absences respond (negatively) to better school infrastructure and local monitoring and (positively) to poverty in the local area. In Zambia (Das and others 2005) a large fraction of teacher absences are due to sickness arising from high HIV infection rates in the population. Given these complex and overlapping conditions, it is critical to understand to what extent better working conditions (better infrastructure, better healthcare and living conditions) and better monitoring or incentives can lead to higher teacher effort in government schools.

The second explanation is that unobserved *fixed* teacher characteristics such as motivation matter—some teachers are highly motivated and altruistic, others are not. In related work on Pakistan, we document that teacher qualifications and training are *lower* in private compared to public schools, but turnover is much higher among the former, suggesting that private schools may be better able to fire teachers with low levels of intrinsic motivation. Depending on the relative importance of these two explanations, the policy implications are very different. If teacher performance responds to the environment and to incentives, changing these to elicit higher effort will lead to better performance. However, if fixed characteristics account for some fraction of the variation in learning, the correct policy response is to allow more flexibility in hiring and firing teachers based on their initial performance.

The second puzzle is why children in Pakistani villages don't go to the better government schools in the village, given that there are good and bad government schools in *every village*. There are no residential restrictions on what schools children can attend, no admission restrictions either in public or private schools and out-of-pocket expenditures are identical across different government schools (there are no PTA fees, for instance). By sending their children to better schools in the same village, parents will reap the benefits of a better education later on. How to square the school choices of children in Pakistani villages with standard economic theory is not obvious: for instance, if it were the case that certain government schools are in the outskirts and attended only by the poor, while others are in the center and attended by the rich, we should see a large drop in the learning gap once we account for student wealth, but we do not. Understanding the limits to arbitrage in these rural contexts could shed light on what parents expect from schools, and what constrains them in their search for a better education.

A third puzzle centers on variation in teacher effort across districts in India. Earlier work by Kremer and others (2005) links higher teacher absenteeism with poorer districts, older teachers, permanent contracts, and local incentives such as paved roads and recent inspections. Recent research by Banerjee and Somanathan (2005) and Pandey (2005) points to a political economy–based explanation, whereby the potential for participation by the poor and the low caste, who are the main users of public schools, is lower in districts where there is greater concentration of power in the hands of the elite. Understanding how teacher accountability and effort can be improved in such a setting will be critical for India's future.

Bibliography

Action Aid Pakistan. (1999). "Comparative Analysis of Public, Private and NGO Schools." Islamabad

Aggarwal, Yash. (2000). "Primary Education in Delhi. How much do the Children learn?" NIEPA, New Delhi.

Aikara, Jacob. (1997). "Learner Achievement in Primary Schools, Unit for Research in the Sociology of Education." TISS, Mumbai.

Andrabi, Tahir, Jishnu Das, and Asim Khwaja. (2002) "Test Feasibility Survey PAKISTAN: Education Sector." Harvard University, Mimeo.

Andrabi, Tahir, Jishnu Das, and Asim Khwaja. (2005). "Private Schooling: Limits and Possibilities." Processed. The World Bank, Washington, D.C.

Andrabi, Tahir, Jishnu Das, Asim Khwaja, and Tristan Zajonc. (forthcoming) "Religious School Enrollment in Pakistan: A Look at the Data." *Comparative Education Review*.

Annual Status of Education Report (ASER). (2005). Pratham, Mumbai. Online at: http://www.pratham.org/aserrep.php

Arif, Charagh, Saima Chaudhary and Uzma Gilani. (1999). "Learning Achievement of Grades 3 and 5 Children in Rural Primary Schools." Punjab Curriculum Development and Research Centre. October, 1999.

Balochistan Bureau of Curriculum and Extension Centre. (1999). "Learning Achievement of Grade 3 and 5 Children in Rural Primary Schools of District Pishin, Balochistan."Quetta December, 1999.

Banerjee, Abhijit and Lakshmi Iyer. (2005). "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4): 1190-213.

Banerjee, A. and R. Somanathan. (2005) "The Political Economy of Public Goods: Some Evidence from India." MIT, Dept. of Economics, processed.

Bashir, Sajitha. (1994). "Achievement Performance at the Primary Level in Public and Private Schools of Tamil Nadu." Indian Educational Review 29(3-4, July-Oct): 1-26

Chaudhury, Nazmul, Jeff Hammer, Michael Kremer, Karthik Muraldhiran, and Halsey Rogers. (2006) "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*.

Das, Jishnu, Stefan Dercon, James P. Habyarimana, and Pramila Krishnan. 2005. "Teacher Shocks and Student Learning: Evidence from Zambia" (May 2005). World Bank Policy Research Working Paper No. 3602. Available at SSRN: http://ssrn.com/abstract=753497.

Drèze, Jean, and Geeta Gandhi Kingdon. (2001). "School Participation in Rural India". *Review of Development Economics* 5(1): 1-24.

Duflo, Ester, and Rema Hanna. (2005). "Monitoring Works: Getting Teachers to Come to School" NBER Working Paper 11880.National Bureau of Economic Research, Cambridge, Mass.

EDI. (1999). "Basic Skills Assessment – A Pakistan Case Study." Theme 3 at EDI Seminar held at Bhurban (February 7-10, 1999).

Filmer, Deon, and Lant Pritchett. (1998). *Educational Enrollment and Attainment in India: Household Wealth, Gender, Village and State Effects*. Washington, D.C.: World Bank.

Government of Pakistan. 2000. *Mathematics Curriculum For Classes K & I-V*." Curriculum Wing, Ministry of Education, Islamabad.

Govinda, R, and N.V. Varghese. (1993). "Quality of Primary Schooling in India. A case study of Madhya Pradesh." IIEP, Paris and NIEPA, New Delhi.

Hanushek, Eric A. (2001). "Black-White Achievement Differences and Governmental Interventions" *The American Economic Review* Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association. 91(2): pp. 24-28.

Haque, Manzoorul, Sharif Kalhoro, and Shoukat Saeed. (2000). "Baseline Survey of Learning.". May, 2000. Bureau of Curriculum Development and Extension Wing, Jamshoro Sindh

Hasan, Arif. (1995). "Baseline Survey of Learning Achievement in Primary Grades in Bihar." A.N. Sinha Institute of Social Studies, Patna.

Jencks, C., & Phillips, M. (1998). "*The black-white test score gap: An introduction.*" In C. Jencks and M. Phillips (Eds.), "The black-white test score gap": pp. 1-51. Washington, DC: Brookings Institution.

Khan, Habib, Seyed Shah, Mirza Ahmad, Nasir Amin, Humala Khalid and Ahmad Malik. (1999) "Measuring Learning Achievement at Primary Level in Pakistan" Pakistan Ministry of Education (Academy of Educational Planning and Management), Islamabad, 1999

Kingdon, Geeta Gandhi. (1998). "Education of Females in India: Determinants and Economic Consequences: A Case Study of Urban Uttar Pradesh." McNamara Fellowships Office, Economic Development Institute of the World Bank, Washington, D.C. Processed.

Kizilbash, Hamid. (1997). "Working with Rural Primary Teachers: Final Report on Improving the Quality of Rural Primary School Teachers in Pakistan." Ali Institute of Education, Lahore, Pakistan.

Kremer, Michael, Nazmul Chaudhury, Hasley Rogers, Karthik Muralidharan, and Jeffrey Hammer. (2005). "Teacher Absence in India: A Snapshot" *Journal of the European Economic Association* 3(2-3): 658-667.

Ministry of Human resource development, India. (2001). Selected Educational Statistics.

Mirza, Munawar, and Abdul Hameed. (1994). "Differential Achievement of Primary Age Students and the Cost Effectiveness by School Type." Primary and Non-Formal Education Wing, Ministry of Education.

Mullis, I.V.S., M.O. Martin, E. J. Gonzalez, and S. J. Chrostowski. (2004). *Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Multi-Donor Support Unit for the Social Action Programme. (1995). "Determinants of Primary Students Achievement – National Survey Results." Islamabad, July 1995.

N.W.F.P. Bureau of Curriculum Development and Extension Services. (1999). "Learning Achievement of Grade 3 and 5 Children in Rural Primary Schools." December, 1999.

Nasir, Z. M., and H. Nazli (2000). "Education and Earnings in Pakistan." http://www.eldis.org/fulltext/PIDE-edearnings.pdf.

OECD Publishing (2005). "Education at a Glance: OECD Indicators 2005." OECD Publishing.

PAKISTAN. (undated) "Strategic Framework for National Education Assessment 2001."

Pandey, Priyanka. (2005). "Service delivery and Capture in Public Schools: How does History Matter and Can Mandated Political Representation Reverse the Effects of History?" World Bank. Processed.

Pervez, Muhammad. (1995). "Basic Competencies of Children in Pakistan." UNICEF: Islamabad, March 1995.

Punjab Literacy Watch. (1999). "Levels of Pupil Achievement in Primary Schools of Punjab: A Sample Study." UNESCO. October, 1999.

Shah, Mubarik Hussain. (1984). "National Achievement Tests (Protocol Analysis)." Primary and Non-Formal Education Wing, Ministry of Education. Islamabad.

Shukla. S., Garg, V.P., Jain, V.K., Rajput, S. and Arora, O.P. (1994). Attainment of Primary School Children in India, National Council of Educational Research and Training, New Delhi.

Singer, Peter W. 2001. "Pakistan's Madrassahs: Ensuring a System of Education not Jihad." Brookings Institutions Analysis Papers 41.Washington DC.

UNESCO. (2001). "Learning Achievement in Primary Schools of Pakistan: A Quest for Quality Education." UNESCO, Islamabad, in collaboration with the Ministry of Education Curriculum Wing. March, 2001

Tables and Figures

TABLE 1. A BRIEF REVIEW OF TESTING IN PAKISTAN

Year of Test	Subjects Tested	Sample Size (schools)	Sampling Methodology	Test Outcomes Mean % (standard error)	Test Available?	Test Validation Documents Available?	Source
1984	Science Mathematics	3,300	Representative. Grades 4 and 5.		No.	No.	Shah (1984)
1988-9	Science Mathematics	11,000 (500)	Grades 4 and 5.		No.	No.	BRIDGES project at Harvard Institute for International Development (1989)
1994	Mathematics Science Social Studies Dinyat	15,991 (472)	Grades 3 and 5.		No.	No.	Mirza and Hameed (1994)
1995	Life Skills Knowledge Rote Reading Reading with Comprehension Writing from Dictation Writing Letter Numeracy and Arithmetic Mental Arithmetic Reading of Holy Qur'an	2582	Multi-stage, systematic-random sample design. All children ages 11-12.	26.1 63.7 26.8 61.7 18.1 69.6 67.7 44.2	No.	No.	Pervez (1995)
1995	Mathematics General Knowledge Comprehension	11,563 (527)	Grade 5.	45.6 74.4 69.1	No.	No.	MSU (1995)
1999	Mathematics Urdu General Knowledge	965 (50)	Grade 4.	60 71 75	No.	No.	Action Aid Pakistan (1999)
1999	Science Mathematics Urdu	2794 (145)	Grade 4. Sample not proportionate to universe.	72 58 72	No.	No.	Khan and others (1999)

1999	Mathematics Science Social Studies Urdu	200 (20)	Grade 3 and 5.	No.	No.	Arif and others (1999)
1999	Mathematics Science Social Studies Urdu	160 (20)	Sample 10 male, 10 female schools (80 students each). Grades 3-5.	No.	No.	Research Team of Bureau of Curriculum Development and Extension Services, NWFP (1999)
1999	Mathematics Urdu	1371 (31)	DEOs selected equal numbers of good, average and weak schools.	No.	No.	Punjab Literacy Watch (1999)
2000	Sindhi Mathematics Science Social Studies Islamiyat	300 (20)	Randomized sample 10 male, 10 female schools (150 students each). Grades 3-5.	No.	No.	Haque and others (2000)
2000	Mathematics Science Social Studies Urdu	801 (20)	10 male, 10 female schools. Grades 3-5.	No.	No.	Research Team of Bureau of Curriculum and Extension Centre, Balochistan (2000)

TABLE 2. SELECTED CHARACTERISTICS OF SAMPLED VILLAGES (PAKISTAN)

	(1) Rural Punjab	(2) ATK, FSD, RYK	(3) ATK FSD, RYK with Private School	(4) 112 Sample Villages
Village Population	2062	2665	4771	4125
	(16)	(49)	(129)	(228)
Percent Adult Literacy	37	34	45	42
	(0.0)	(0.0)	(0.0)	1.0
Household Size	6.88	7.15	7.19	7.15
	(0.01)	(0.02)	(0.04)	(0.10)
Village Area (Acres)	1639	1581	1864	1891
	(19)	(19)	(30)	(93)
Fraction Houses with Potable Water	0.019	0.021	0.025	0.021
	(0.000)	(0.001)	(0.001)	(0.003)
Fraction Houses with Electricity	0.097	0.086	0.110	(0.103
	(0.000)	(0.001)	(0.001)	(0.004)
Fraction Pacca (Cement) Houses	0.521	0.503	0.642	0.647)
	(0.002)	(0.006)	(0.010)	(0.022)
Number of Private Schools in Village	0.39	0.43	1.96	1.68
	(0.01)	(0.02)	(0.07)	(0.11)
Number of Government Schools in Village	2.33	2.71	3.43	2.95
	(0.01)	(0.05)	(0.12)	(0.20)

Notes: Cells contain means for the subsample of villages listed in the column heading. Parentheses contain the standard error of the mean. The statistics are based on the Census of Villages from the Population Census Organization (1998). ATK, FSD and RYK are Attock, Faisalabad, and Rahim Yar Khan districts, respectively. The 112 LEAPS sample villages were drawn randomly (stratified by district) from the villages in Column 3.

TABLE 3. SELECTED CHARACTERISTICS OF SAMPLED VILLAGES (INDIA): N=200

Variable	Mean	Std. Deviation
Distance from nearest pucca (mortar) road (kilometer)	2.12	3.00
Village area (hectares)	706	718
Population size	2035	1419
Literacy rate	0.35	0.12
Fraction of scheduled caste and tribe population	0.28	0.16
Number of government primary schools	2.01	1.16
Number of private primary schools	0.88	1.23

Note: The population, village area and literacy rate variables are from the Census of India, 2001. The remaining variables are from the data collected by the authors.

TABLE 4. COMPARING QUESTIONS AND PERFORMANCE ON SIMILAR EXAM QUESTIONS

Exam Question	India	Pakistan	India	Pakistan
Complete the addition problem:	55+12	36+61	87%	86%
Complete the subtraction problem:	640-112	238-129	53%	32%
	8803-4213		39%	
Complete the multiplication problem:	2*11	4*32	77%	50%
	12*45		50%	
Complete a sentence with the word:	Home	Beautiful	41%	33%
	Message	School	27%	31%

Notes: This table presents several questions from the Pakistan and India exams that are roughly comparable. The third and fourth columns state the actual question. The fourth and fifth column shows the percentage of students who answered the question correctly. This table is meant to show absolute not relative level of performance. The questions, while comparable, are not equivalent, and class three children were tested in Pakistan compared to class four children in India.

TABLE 5. MORE DETAILED RESULTS FOR PAKISTAN EXAM

Subject	The Question	Percentage who answered correctly			Corresponding Grade for Curriculum (Math only)
		All Schools	Government Schools	Private Schools	
Mathematics	4 + 6	89	87	92	K & I
Mathematics	36 + 61	86	84	90	K & I
Mathematics	8 – 3	65	63	71	K & I
Mathematics	5 x 4	59	55	67	II
Mathematics	238 - 129	32	29	39	II
Mathematics	Read and Write the time (Clock shows 3:40)	24	21	31	II
Mathematics	384 ÷ 6	19	17	24	III
Mathematics	4 x 32	50	46	59	III
Mathematics	Fractions: ½ + 3/2	19	17	22	III
Mathematics	Read a diagram of a scale to answer which part is heavier	12	11	14	III
Mathematics	Fractions: 7/5 – ¾	1	1	1	IV
English	Write the letter "D" (spoken out loud)	86	82	93	
English	Fill in the right letter D _ F	70	62	87	
English	Tick the correct answer to match the picture (picture of book)	70	61	89	
English	Fill in the missing letters (picture of a ball) BA _ _	45	35	69	
English	Fill in the missing letters (picture of a flag) F L A _	29	20	48	
Urdu	Tick the correct answer to match the picture (picture of house)	52	48	61	
Urdu	Tick the correct answer to match the picture (picture of book)	73	69	82	
Urdu	Write a sentence with the word "beautiful"	33	27	49	
Urdu	Write a sentence with the word "school"	31	24	46	

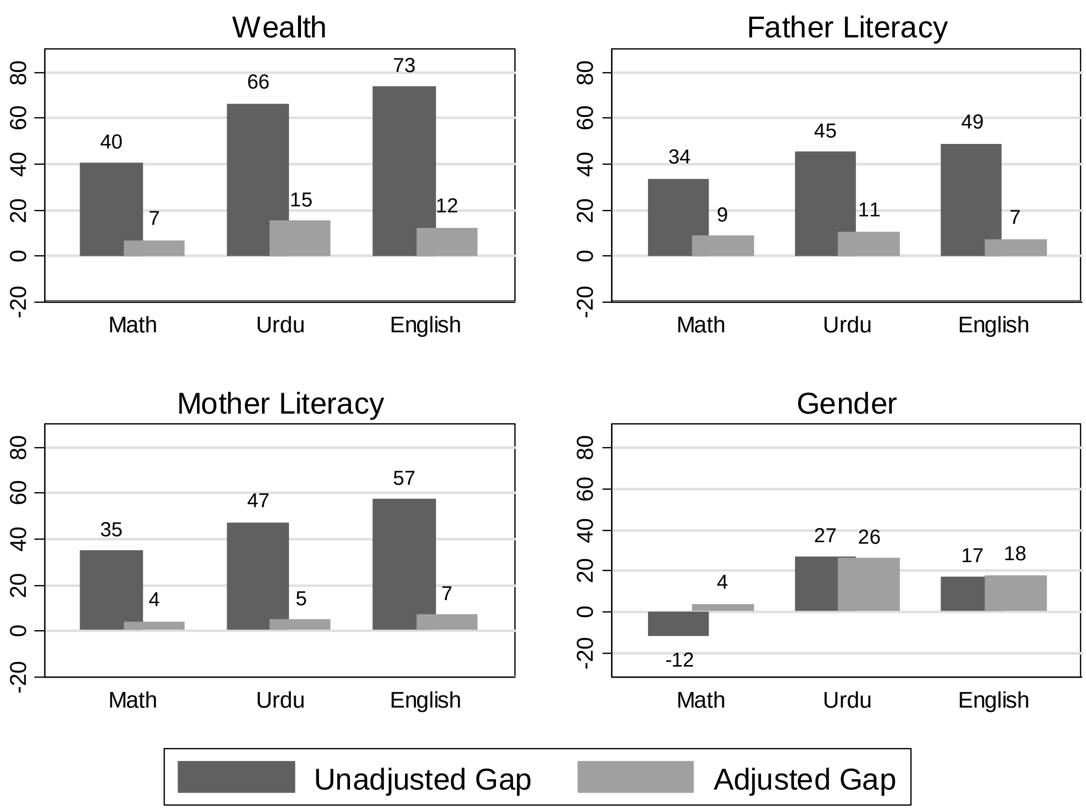
Notes: Questions are from the Pakistan LEAPS exam. Columns 3, 4, and 5 report the percentage of children who answered the question correctly in all schools, government schools, and private schools.

TABLE 6. KNOWLEDGE SCORES AND ABSOLUTE KNOWLEDGE

			India			Pakistan		
			Knowledge Points			Knowledge Points		
Exam Question	India	Pakistan	350	500	650	350	500	650
Complete the addition problem:	55+12	36+61	80	94	98	76	89	95
Complete the subtraction problem:	640-112	238-129	8	51	92	11	28	55
	8803-4213		4	29	81	-	-	-
Complete the multiplication problem:	2*11	4*32	42	87	98	22	49	77
	12*45		8	44	88	-	-	-
Complete a sentence with the word:	Home	Beautiful	2	27	89	2	15	78
	Message	School	4	36	90	1	10	76

Notes: Cells contain the likelihood a child with the score given in the column heading would answer the question given in the row correctly. The knowledge points 350, 500, and 650 represent the mean score (500) and ± one standard deviation (150).

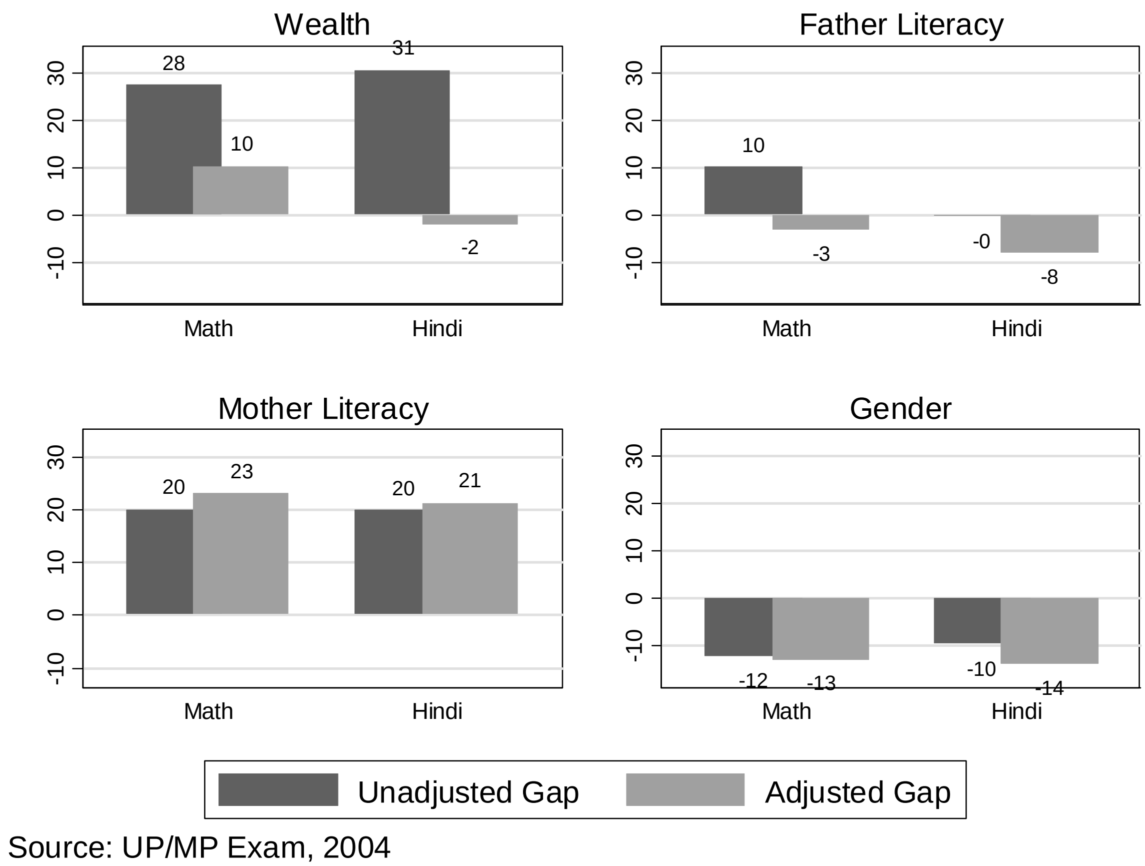
FIGURE 1. THE ADJUSTED AND UNADJUSTED KNOWLEDGE SCORE GAPS BY HOUSEHOLD CHARACTERISTICS (PAKISTAN)



Source: LEAPS Exam, 2004

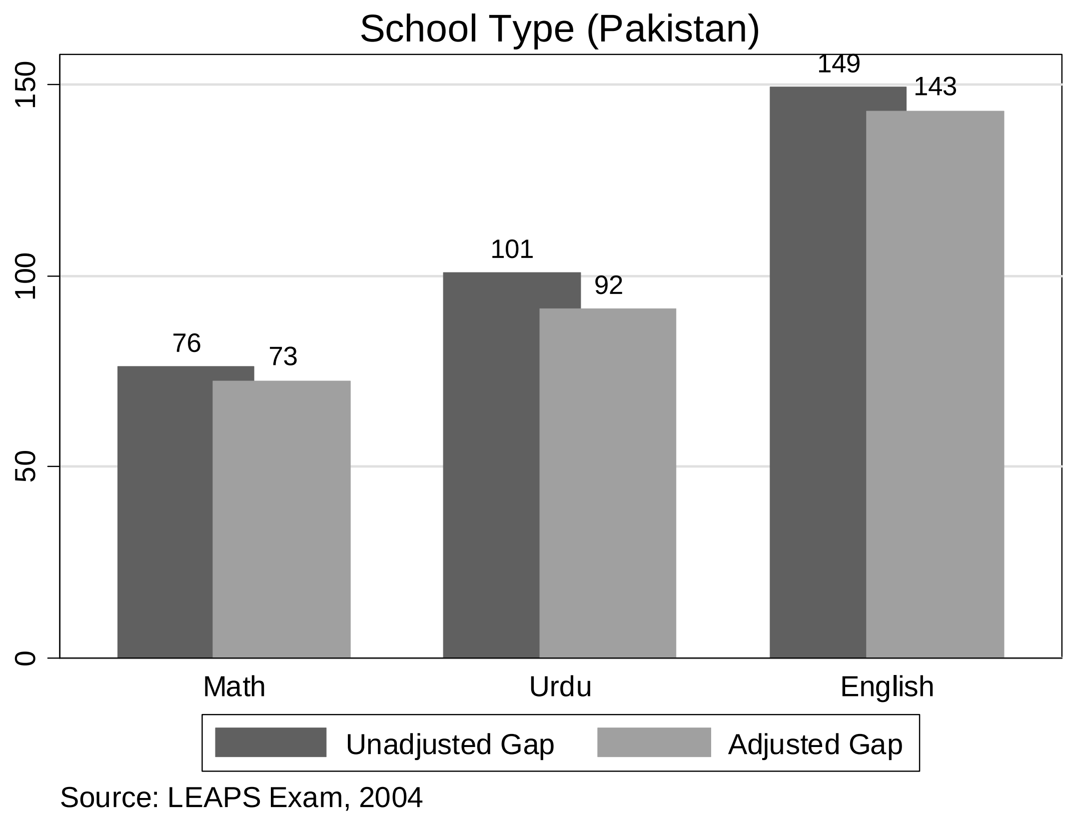
Notes: Gaps are mean differences: (rich – poor), (literate father – illiterate father), (literate mother – illiterate mother), and (female – male). Rich and poor are the top and bottom third of an asset index calculated using principal components analysis. The adjusted gaps are the coefficients from an OLS regression that includes wealth, father literacy, mother literacy, gender, age, age squared, and a school fixed effect (Tables A1 and A2).

FIGURE 2. THE ADJUSTED AND UNADJUSTED KNOWLEDGE SCORE GAPS BY HOUSEHOLD CHARACTERISTICS (INDIA)



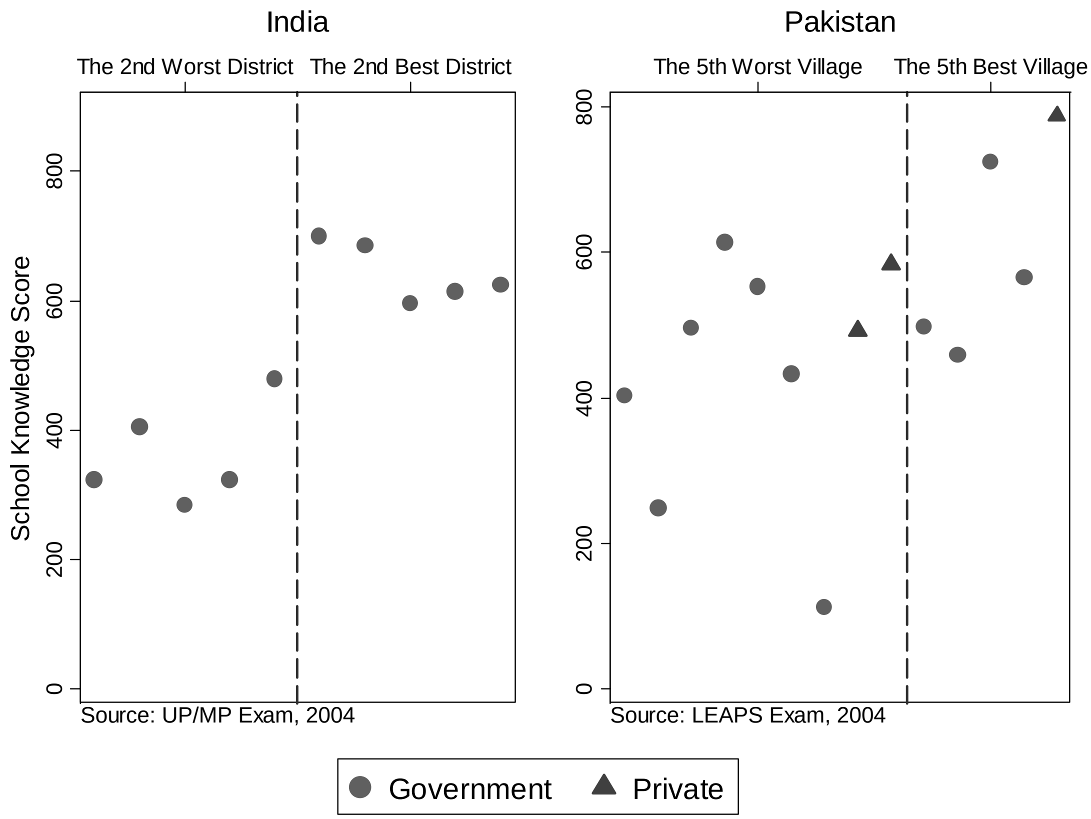
Notes: Gaps are mean differences in the knowledge score defined as: (rich – poor), (literate father – illiterate father), (literate mother – illiterate mother), and (female – male). Rich and poor are defined as families having above or below median land holdings. The adjusted gaps are the coefficients from an OLS regression that includes wealth (land holding), father literacy, mother literacy, gender, age, age squared, upper caste, backward caste, and a school fixed effect (Tables A1 and A2).

FIGURE 3. ADJUSTED AND UNADJUSTED KNOWLEDGE SCORE GAPS BY SCHOOL TYPE IN PAKISTAN



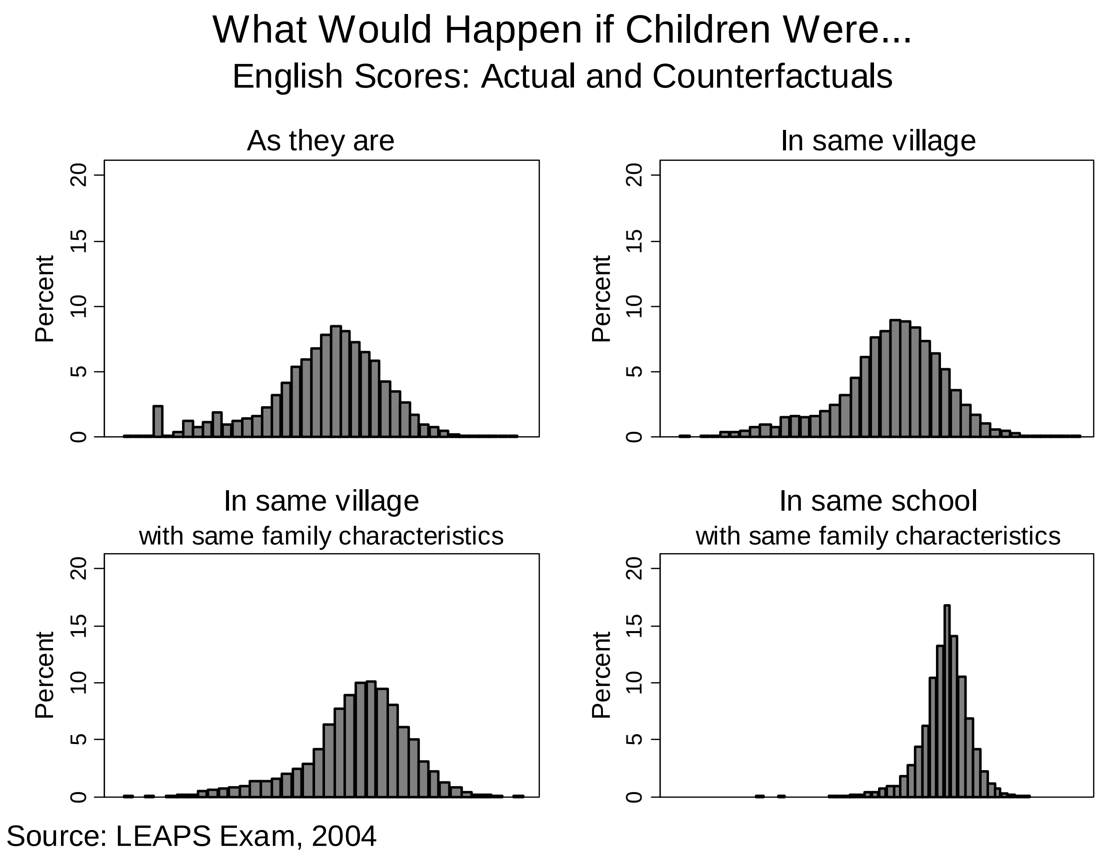
Note: School type gaps are mean differences between children in private schools and government schools. The adjusted gap for Pakistan is the coefficient on private school in a child level OLS regression that includes wealth, father literacy, mother literacy, gender, age, age squared, and a village fixed effect.

FIGURE 4. GEOGRAPHICAL VARIATION AND VARIATION ACROSS SCHOOLS



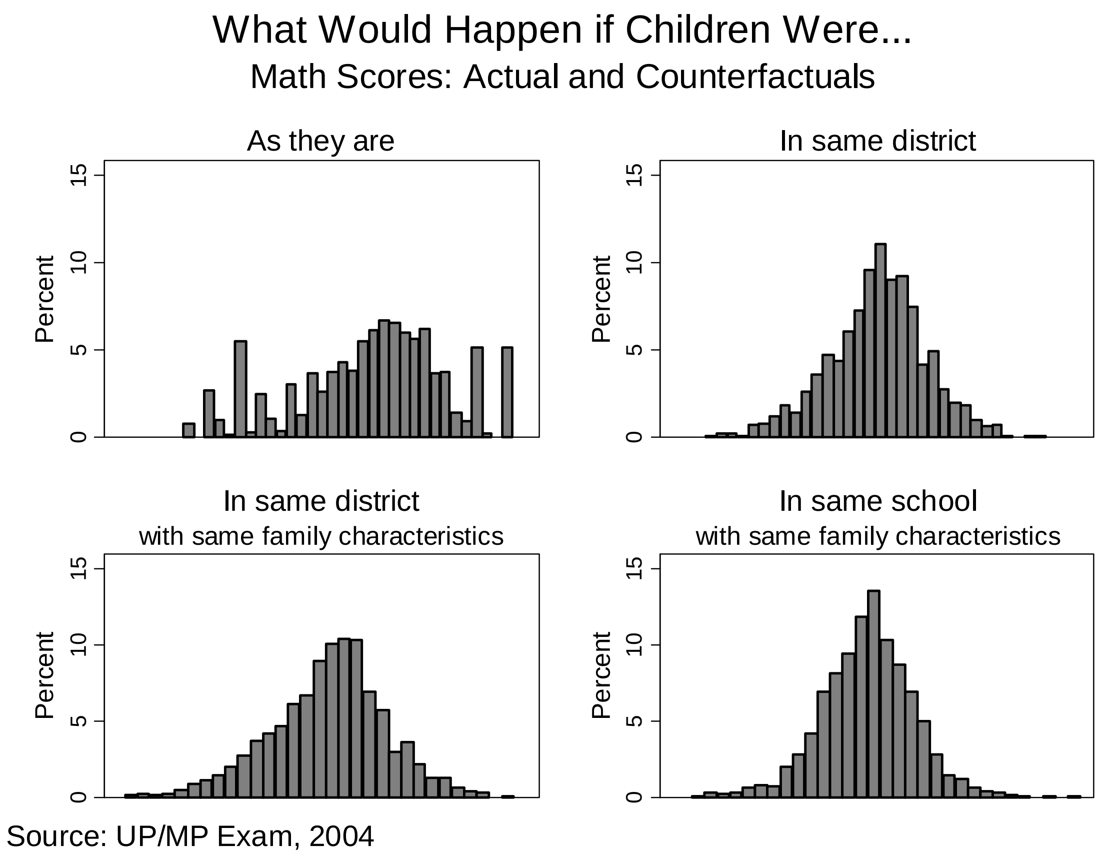
Notes: The first panel shows average Mathematics scores for schools in the 2nd best and 2nd worst districts in the UP/MP exam data. The second panel shows the average scores for schools in the 5th best and 5th worst villages in the LEAPS Mathematics exam data. The dotted line splits the two villages and districts, but the ordering of schools on the x-axis is otherwise irrelevant. Private schools were note tested in India.

FIGURE 5. COLLAPSING KNOWLEDGE SCORE VARIATION (PAKISTAN)



Notes: Panel 1 shows the distribution of English knowledge scores on the LEAPS exam. Panel 2 shows the residual distribution controlling for a village fixed effect. Panel 3 shows the residual distribution controlling for a village fixed effect and child age, age squared, gender, mother literacy, father literacy, and household wealth. Panel 4 shows the residual distribution including an additional school fixed effect.

FIGURE 6. COLLAPSING KNOWLEDGE SCORE VARIATION (INDIA)



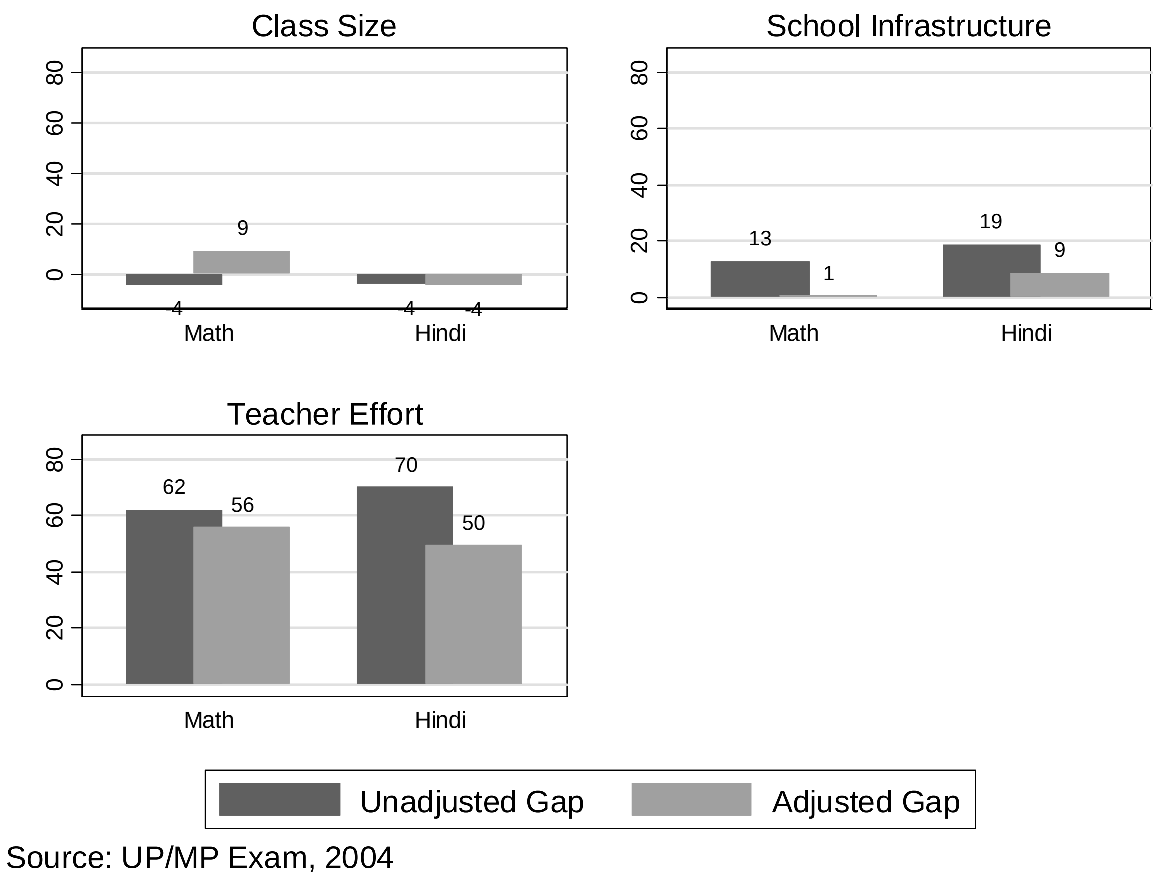
Notes: Panel 1 shows the distribution of math knowledge scores on the UP/MP exam. Panel 2 shows the residual distribution controlling for a district fixed effect. Panel 3 shows the residual distribution controlling for a district fixed effect and child age, age squared, gender, caste, mother literacy, father literacy, and household wealth. Panel 4 shows the residual distribution including an additional school fixed effect.

TABLE 7. ADJUSTED AND UNADJUSTED SCHOOL KNOWLEDGE SCORE DISTRIBUTIONS

		Private Schools		Government Schools	
		Unadjusted Score	Adjusted Score	Unadjusted Score	Adjusted Score
Math	5th Percentile	394	388	288	312
	Median	563	538	478	485
	95th Percentile	703	691	651	642
Urdu	5th Percentile	411	401	294	329
	Median	575	545	476	487
	95th Percentile	705	688	631	623
English	5th Percentile	488	460	226	275
	Median	596	558	463	469
	95th Percentile	725	714	620	610

Notes: Cells contain average school scores. Adjusted scores control for a village fixed effect, age, age squared, gender, wealth, mother literacy, and father literacy. This table is a heuristic rather than analytically accurate, since regressions identify only the conditional mean and not separate quantiles of the distribution.

FIGURE 7. ADJUSTED AND UNADJUSTED KNOWLEDGE SCORE GAPS BY SCHOOL CHARACTERISTICS IN INDIA



Notes: The infrastructure index is a sum of four indicator variables – whether the school has a pucca (cement and brick) building, drinking water, toilet facility and a playground. The gap is between schools above and below the median of this index. Small class size is 50 and large is above 50. The gap is small minus large. Results look similar if class size is split at 40. Schools with active teachers (teacher effort) are those where 50 percent or more of the teachers are engaged is teaching at the time of survey visits. The gap is high effort minus low effort. The adjusted gaps is the coefficient on variable of interest from an OLS regression that includes wealth (land holding), father literacy, mother literacy, gender, age, age squared, upper caste, backward caste, school infrastructure, class size, and teacher effort.

Appendix A

TABLE A1. OLS REGRESSION USED TO COMPUTE ADJUSTED KNOWLEDGE SCORES (PAKISTAN)

	School Fixed Effect			Village Fixed Effect + Private		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Math	Urdu	English	Math	Urdu	English
Female	4.0	26.1*	18.2*	-9.9*	28.6*	17.9*
	(4.6)	(4.6)	(3.8)	(3.6)	(3.5)	(3.2)
Age	0.6	-6.7**	-0.1	3.3	-3.9	0.8
	(2.7)	(2.7)	(2.3)	(3.0)	(3.0)	(2.8)
Age-squared	-0.0	0.1	-0.0	-0.1	0.1	-0.0
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Mother Literate	4.3	4.9	7.3**	3.0	4.4	7.6+
	(4.0)	(4.1)	(3.4)	(4.6)	(4.5)	(4.2)
Father Literate	8.7*	10.7*	7.3*	12.9*	15.8*	13.7*
	(3.3)	(3.3)	(2.8)	(3.8)	(3.7)	(3.5)
Middle Wealth	7.3+	12.8*	3.7	7.0	13.4*	5.2
	(3.7)	(3.8)	(3.1)	(4.4)	(4.3)	(4.0)
High Wealth	7.0+	15.2*	12.2*	18.1*	29.0*	26.7*
	(4.1)	(4.1)	(3.4)	(4.7)	(4.6)	(4.3)
Private School				72.7*	91.5*	143.3*
				(4.1)	(4.0)	(3.7)
Constant	487.8*	530.5*	491.8*	439.4*	465.0*	422.8*
	(18.4)	(18.8)	(15.5)	(20.0)	(19.5)	(18.1)
Fixed Effect	School	School	School	Village	Village	Village
Observations	6347	6347	6347	6241	6241	6241
R-squared	0.54	0.53	0.68	0.18	0.23	0.35

Standard errors in parentheses

+ significant at 10%; ** significant at 5%; * significant at 1%

Notes: Dependent variable is child knowledge score for math, Urdu, or English. Data from LEAPS Exam, 2004. Sample consists of (up to) 10 randomly selected children per school; the OLS regression is not reweighted by the total school population. Wealth is calculated using principal components analysis from an asset survey; middle wealth is the middle third; high wealth is the top third; the bottom third is omitted.

TABLE A2. OLS REGRESSION USED TO COMPUTE ADJUSTED KNOWLEDGE SCORES (INDIA)

	(1) Math	(2) Hindi
Female	-12.8**	-13.8**
	(6.0)	(5.4)
Age	-74.8**	6.0
	(30.4)	(27.6)
Age-squared	3.4**	-0.3
	(1.5)	(1.3)
Mother Literate	23.2*	21.3*
	(8.3)	(7.5)
Father Literate	-3.0	-7.8
	(7.0)	(6.3)
Rich	10.4	-1.8
	(7.4)	(6.7)
Middle Caste	-8.1	-9.0
	(6.9)	(6.3)
High Caste	-5.6	0.2
	(10.3)	(9.3)
Constant	906.3*	481.0*
	(155.4)	(140.8)
Fixed Effect	School	School
Observations	1479	1479
R-squared	0.60	0.66

Standard errors in parentheses

+ significant at 10%; ** significant at 5%; * significant at 1%

Notes: Dependent variable is child knowledge score for math or Hindi. Data from UP/MP Exam, 2004. Rich is defined as having above median land holdings. High castes are upper castes, middle castes are backward castes and the omitted category is low castes who are scheduled castes and tribes.