PRIVATE SCHOOLING, LEARNING, AND CIVIC VALUES IN A LOW-INCOME COUNTRY*

Tahir Andrabi † Natalie Bau‡ Jishnu Das § Asim Khwaja ¶

Abstract

Private school enrollments have increased dramatically in low-income countries, where public school quality is low. Using rich panel data on educational outcomes, we estimate the effects of private schooling on learning in rural Pakistan. Across multiple identification strategies, we find that attending a private school increases test scores. Moreover, attending a private school appears to *increase* civic values, addressing the concern that private provision reduces positive externalities from education. However, average private school effects mask heterogeneous private school quality. To characterize the distribution of school quality, we estimate school value-addeds (SVAs) for all schools in the public and private sectors and provide evidence that SVAs are unbiased estimates of school quality. SVA is highly variant, with significant overlap in the public and private school distributions. Moving to a 1 sd better public school raises test scores by 0.321sd, while moving to a 1 sd better private school raises test scores by 0.223sd. Policies that move students between the public and private sector can deliver a wide range of effects depending upon implementation.

*We thank seminar participants at UCLA for valuable comments. Catherine Michaud Leclerc provided exceptional research assistance.

†Pomona. Contact: tandrabi@pomona.edu

‡UCLA, NBER, and CEPR. Contact: nbau@ucla.edu

§Georgetown. Contact: Jishnu.Das@georgetown.edu

¶Harvard. Contact: khwaja@hks.harvard.edu

1 Introduction

Private primary school enrollments have increased dramatically during the last decade in low-income countries. By 2010, there were more than 80 million children in South Asia in private schools (25 percent of all school-going children) relative to 6 million in the United States (11 percent).[1] In Pakistan, the focus of this study, the share of enrollment in private schools increased from less than 5 percent in 1990 to 35 percent in 2005. The number of private non-religious schools increased from 3,000 to 47,000 between 1982 and 2007 (in contrast to the share of private religious schools, which has remained unchanged at less than 3 percent since 1980).[2] This increase in private schooling is so stark that the *Economist* has written about it multiple times in recent years (2015, 2019). Cross-sectional variation also indicates that there could be large gains from private schooling.[3] Thus, the rise of private schooling raises urgent questions regarding the role of the state in the provision, financing, and regulation of such institutions, as well as widespread in interest in how private schooling can be used as a tool to improve learning outcomes.

In this paper, we provide evidence on the promise and limitations of private schooling. First, following Singh (2015) and Muralidharan and Sundararaman (2015) in India, we we estimate the causal effects of attending a private school on student learning in Pakistan. As is the case in India, we find that private schools are more productive than public schools. That is, they deliver greater learning at a lower per-student cost. However, we observe that this is true in part because of high teacher salaries in the public sector. Public schools, which are highly regulated in their teacher salaries (Bau and Das, forthcoming), face very different factor prices than private schools. As Figure 2 shows, were private schools forced to pay the average teacher salary in the public sector, it would more than close the per-student cost gap. Second, addressing the potential concern that private schooling may come at the cost of good citizenship, as the private sector may not internalize the positive externalities of good citizenship, we provide new evidence on the effect of private schooling on civic values. Finally, we show that a single private school effect estimate masks considerable heterogeneity. We estimate school value-addeds (SVAs) for *all* the schools in the market and show that these are forecast unbiased estimates of students' test score gains in schools. These estimates indicate that there is substantial overlap in public and private school quality, and policies that move students from the public to private sector can deliver a wide range

[1] In India, private school enrollments doubled between 1993 and 2005 to a country-wide average of above 20 percent with a market share above 50 percent in urban areas. Similar patterns are found in African countries; for instance, in The Gambia, private school enrollments increased from 2 (1998) to 20 percent (2008) in the space of 10 years (NCAER, 2005; Pratham, 2010; World Bank, 2019).

[2] Pakistan Integrated Household Survey (The 1991) and the National Education Census (2005) provide data on private schools; for religious school enrollments, see Andrabi et al. (2006).

[3] Figure 1 plots histograms of test scores in the public and private sectors in Pakistan, and shows that there is a large gap in learning outcomes.

of effects.

To estimate the effect of private schooling on test scores and civic values, we analyze detailed microdata on schools, students, and households collected over four years from 2003-2007 with several different identification strategies. First, following the value-added literature in the United States (Chetty et al., 2014) and Singh (2015) in India, we exploit the panel aspect of our data to generate value-added estimates of the effect of private schooling. This strategy uses past test scores to account for the selection of students into schools. Second, we also include child fixed effects in these regressions to account for selection on unobservables so that school quality is identified from students who switch schools. Focusing on students who switch schools also allows us to use an event study strategy to directly evaluate whether our value-added estimates of the private school effect are biased by the selection of students into schools. That is, we show that prior to a student switching into a private school, later private schooling does not predict her test gains but after the switch, it predicts them with the expected coefficient. Third, to exploit exogeneous variation in school switching, we instrument for school switches with private school closures. Fourth, we use distance to a private school relative to a public school, which has been shown to strongly affect private school enrollment in Pakistan (Carneiro et al., 2016; Bau, 2019), as an instrument for private school enrollment. While distance instruments may be biased if students' or schools' locations are endogenous (Altonji et al., 2005), an unusual feature of Pakistani villages allows us to control for much of the correlation between students' socioeconomic statuses and geographic locations. Across all these identification strategies, we find very similar results. An additional year of private schooling increases mean test scores by approximately 0.1-0.2 standard deviations. For civic values, where we only have cross-sectional data, the estimates with the two instruments also show that private schooling increases civic values.

However, the public and private sectors are unlikely to be homogeneous. If we naively interpret the cross-section of test scores reported in Figure 1 as capturing heterogeneity in school quality, we see substantial overlap in the public and private sector. Even causal measures of the effect of private schooling capture the effect of the private school the marginal student attended (under a given identification strategy) relative to the marginal public school. Consequently, there is no effect of private schooling that is invariant to the policy used to shift students into private schools. Characterizing the distribution of private and public school quality is important for understanding the effects of any policy that moves students between the two sectors.

To characterize the distribution of quality, we exploit the size and richness of our data from Pakistan, in which we observe 71,677 student-year observations. These data allow us to estimate SVA for each public and private school in the data. Having verified that (1) SVAs appear to be forecast unbiased in event studies with students who switch schools and (2) SVA can be used to replicate the private school effect size for the school switchers in the child fixed effect identification

strategy, we then treat SVAs as unbiased measures of school productivity. With these measures in hand, we can plot the distribution of private school and public school quality in rural Pakistan. Our results indicate that school quality is indeed highly variant in both sectors. Corrected for estimation error, a 1 sd better public school improves mean test scores by 0.32 student test score standard deviations, while a 1 sd better private school improves test scores by 0.22 sd.[4] While the mean private school has a higher SVA than the mean public school, there is substantial overlap in the distributions. The bottom 25 percent of private schools are worse than the top 48 percent of public schools, and a right tail of public schools are as good as the best private schools in most subjects. Furthermore, half of the variation in private school quality is across- rather than within-villages.

Using our SVA estimates, we consider two simple counterfactual voucher experiments. To capture the "lowerbound" effect of private schooling, we assign each child in a public school to the worst private school in her village. This results in a mean test score gain of 0.07 sd. To capture the "upperbound" effect, we assign each child in a public school to the best private school in her village, leading to an effect of 0.23 sd. Therefore, the partial equilibrium policy effects of a voucher program could range from 0.07-0.23 sd, and the implementation details of such programs will be important.

This paper relates to a small literature on test scores and private schooling from a number of low and middle income countries that complements the extensive literature on the United States.[5] The literature on the U.S. suggests that private schooling positively impacts graduation rates, with higher impacts for children in inner-city areas and Hispanic and black populations. The contributions from low-income countries establish a positive correlation between private schooling and test scores, which has often been used to argue for private schooling as a viable alternative to government schools in similar contexts. However, moving from correlations to causation has proved difficult with the exceptions of Singh (2015) and Muralidharan and Sundararaman (2015). We build on this literature in three ways.

First, we provide evidence on the causal effect of private schooling on students' learning and civic values in a new, previously-unstudied context. In this context, private schools do not merely lead to the same test score gains at lower costs. They generate greater test score gains. This is particularly important, because while private schools have lower per-student costs than public schools in many low-income contexts, this doesn't immediately imply that they use resources more efficiently. Rather, they face different labor prices. If a private school paid the average public teacher salary, its per-student cost would be approximately 8.4 USD (relative to 5 USD in the

[4]Our measures of SVA include teacher quality, since overall school value-added is the relevant school quality measure for policy.

[5]See Neal (1998) for a review of the US literature; see Wolf and Macedo (2004) or Wolf (2007) for a discussion of civic values and private schooling in the U.S. and Europe; see Jimenez et al. (1988); Tooley and Dixon (2003); Kingdon (1996a,b); Bashir (1994) for evidence from low-income countries.

public sector).

Second, importantly, we show that private schooling does not damage civic values. It is widely acknowledged that schools have multiple goals. Public schooling may have positive externalities, such as increasing citizenship and socialization that compensate for lower test scores (Meyer et al., 1979). If nation-building and the production of citizenship are not fully contractible, state provision may be more efficient (Pritchett and Viarengo, 2015). However, we find that attendance in a private school does not reduce civic values, both in terms of civic knowledge and civic disposition. Thus, government provision does not have positive externalities for citizenship, and may even negatively impact civic values.

Third, while it is widely believed that schools differ in quality within sectors, and it has been demonstrated in the United States (Hoxby and Murarka, 2009), there is little evidence on school quality distributions from low-income countries. This is likely due to the lack of high quality, administrative data in these contexts, which prevents researchers from computing school value-added measures. We contribute by computing these measures and show that using only 4 years of longitudinal data, we are able to calculate forecast unbiased SVAs. If SVAs are credible proxies for school quality, this suggests that researchers with access to relatively short longitudinal data sets can use SVAs to evaluate policy effects in other contexts.

Lastly, we contribute by using our SVA estimates to characterize the distribution of public and private school quality and to characterize the range in the effects of reassigning students to private schools. Given the heterogeneity in school quality in both sectors, there is a wide range of interventions that policymakers could pursue to shift students between sectors. Each of these interventions would have very different results. Since different policies will shift children at different parts of the school quality distribution, understanding to what extent non-experimental (but still longitudinal) data can allow researchers and policymakers to characterize the full SVA distribution is critical.

The remainder of this paper is organized as follows. Section 2 describes the context and the data. Section 3 introduces our methods for estimating the effects of private schooling and reports their results. Section 4 estimates and validates the SVAs, characterizes their distribution, and reports on their correlates. Finally, Section 5 concludes.

2 Context & Data

2.1 Context

Several previous papers discuss the dramatic changes in Pakistan's educational landscape since 1990 (Andrabi et al., 2006, 2008, 2013). Contrary to popular belief and frequent media reporting, Andrabi et al. (2006) showed that enrollment in religious schools or Madrassas is low (roughly 1 percent) and has remained constant since the mid-80s. On the other hand, like in other South

Asian and African countries, there has been an explosion in the private sector share of primary education, both in terms of schooling availability and enrollment share. The last two decades have seen a more than ten-fold increase in the number of private primary schools (from 3,800 in 1983 to 47,000 by 2005), and currently, over a third of primary-level enrollment is in the private sector, with the fastest growth coming from rural areas (Andrabi et al., 2006). Andrabi et al. (2008) shows that private schools are for-profit enterprises in a largely unrestricted market, with no subsidies from the government and little (if any) de facto regulation and that the median annual fee in a rural private school in 2003 was Rs. 1000 ($18), so that a month's fee was roughly equal to the daily wage rate of an unskilled worker. Furthermore, these schools are small enterprises. The median school has 125 students and 7 teachers.

Computing the cost of schooling in public schools just on the basis of teachers' salaries and school-level expenditures suggests that educating children in public schools is fifty percent more expensive than in private schools. Estimates that also account for overhead administrative costs and capital expenditures suggest that costs could be as much as two times greater. The large difference in costs between public and private schools arises primarily from teachers' salaries and administrative overheads in public schools. Andrabi et al. (2008) shows that there are few fixed costs in running a private school (private schools are often setup initially in a room/part of the teacher/owner's house) and that teachers' wages in private schools were 20 percent of those in public schools. Although teachers in the private sector are less educated, the bulk of the difference in teachers' wages is not accounted for by differences in characteristics. This characteristic of education markets is not specific to Pakistan. Jimenez et al. (1991) show similar differences in Colombia, the Dominican Republic, the Philippines, Tanzania, and Thailand, and Kremer and Muralidharan (2008) and Muralidharan and Sundararaman (2013) show similar results for India.

The debate about private schools in low and middle-income countries, including Pakistan, centers around two points of contention. First, educationalists are troubled by the (relatively) poorly educated workforce in the low-cost private sector and have often argued that, with poor education, no teacher training (less than 26 percent have any training at all), and low wages, the private sector cannot possibly provide "quality" education (Barrett et al., 2007). Explanations for the increasing exodus from public schools are often based on the inability of parents to discern school quality (Habib, 1998).

The second point of contention is the relative role of public and private schools in the creation of 'good' citizens who can participate in the task of nation building. In most post-colonial countries, this was one of the key aims of the public schooling system and it is particularly relevant in the case of Pakistan (see Cohn and Scott (1996) on India and Bassey (1999) on sub-Saharan Africa). Dean (2005) provides a summary of the debates surrounding the broader holistic goals of Pakistan's education policy since the country's independence in 1947. The influential, first education conference

in 1947 (and subsequent statements regarding educational goals over the years) explicitly called for training in citizenship.[6]

This paper addresses these two key issues by bringing learning outcomes and civic values measures together with a causal empirical strategy. This allows us to evaluate school quality based on students' outcomes as opposed to schools' observed inputs.

2.2 Data

We use data collected as part of the Learning and Educational Achievement in Punjab Schools (LEAPS) project, an ongoing survey of primary schooling in Pakistan. The sample comprises 112 villages in 3 districts of Punjab, the largest province in Pakistan with a population of over 60 million. The 3 districts – Attock, Faisalabad, and Rahim Yar Khan – were chosen on the basis of an accepted stratification of the province into the better performing North and Central regions (Attock and Faisalabad respectively) and the poorly performing South (Rahim Yar Khan). Because the project was envisioned in part as a study of the rise of private schools, the 112 villages in these districts were chosen randomly from the list of all villages with an existing private school. Sample villages are generally larger, wealthier, and more educated than the average rural village. Nevertheless, because private schools are more likely to locate in larger villages, at the time of the survey more than 50 percent of the province's population resided in such villages (Andrabi et al., 2006).

Surveys were administered as part of a longitudinal study in the years 2003-2004, 2004-2005, 2005-2006, and 2006-2007. This panel allows us to study the effect of private schooling on students contemporaneously. There were three broad components of the survey: (1) information on all schools in the 112 villages; (2) low-stakes test scores of in-school children in every survey year, as well as scores on civic tests in the 2005-2006 (round 3) survey year, and (3) a detailed household survey of 16 randomly chosen households in every village.

School Surveys. The school survey covers all the schools within the sample village boundaries and within a short walk of any village household. Including schools that opened and closed over the first four rounds, 874 schools were surveyed at least once, while three refused to cooperate. Sample schools account for over 90 percent of the primary school enrollment in the sample villages. The school-level survey included detailed information on school infrastructure and expenditures, as well as schools' GPS coordinates.

In addition, we collected basic socio-demographic information and wages for all the teachers in

[6]This is very much in line with the global conception of civics and is quite similar to the goals of the NAEP civics framework in the United States. The civics framework states, "The possession of a vote by a person ignorant of the privileges and responsibilities of citizenship... is responsible for endless corruption and political instability. Our education must ... [teach] the fundamental maxim of democracy, that the price of liberty is eternal vigilance and it must aim at cultivating the civil virtues of discipline, integrity, and unselfish public service."

a school. The data on schools and teachers is described in Andrabi et al. (2007). One fact from this data is notable. Teacher salaries in the private sector are a fifth of those in the public sector, consistent with the fact that salaries explain a substantial fraction of the differences in per-student costs between sectors.

Tests. To assess learning outcomes and civic values, we tested children in each of the surveyed schools repeatedly over time. In round 1, a cohort of third graders was tested and these were then followed over time. In round 3, we began following a second cohort of 3rd graders, and 14,954 additional unique students were tested. Test scores are based on exams in English, Urdu (the vernacular), and mathematics administered mainly to students in grades 3-5 between 2004 and 2007.[7] There were 40 questions on average in every tested subject and the tests were designed to maximize precision over a range of abilities in each grade. Tests were scored and equated across years using Item Response Theory, as described in Das and Zajonc (2010). To avoid the possibility of cheating, the tests were administered directly by our project staff and not by classroom teachers. Additionally, we also administered a short one-page survey to randomly selected tested children within the schools to collection information on parental education and household assets.

In addition to the subject tests, we also administered a civics test to all children in round 3. The civics test follows a standard protocol similar to that of the civics portion of the National Assessment of Educational Progress exam in the United States. It was broadly divided into questions designed to elicit civic knowledge and civic dispositions. In the civic knowledge section, we ask about the political structure of the state and its history, basic geography of the country and region, political and historical personalities and familiarity with a popular song, a national slogan, and a historical poem. In the civic disposition section, we ask about trust in government institutions, preference for democratic methods of decision-making, gender bias through two questions on the relative ability of girls versus boys in learning and in positions of authority, and familiarity with the scientific method in terms of thinking about intellectual reasoning and skills.[8] To evaluate the effect of private schooling on civics, we form four indices: (1) a full index that includes all questions, (2) a knowledge index that takes the average score on the knowledge questions, (3) a pro-government disposition index, and (4) a gender bias index. Appendix Table A1 reports all the civics questions by the index they are assigned to and provides summary statistics for how students answered them.

Table 1 provides evidence on average learning levels and their dynamics over time for the first

[7] In round 4 of the data collection, sixth graders were also tested.

[8] We chose not to ask direct questions about radical Islamic ideology and extremism as in Shapiro and Fair (2010). Most of the children in our sample are between the ages of 8 and 12; it seemed overly intrusive to ask children at this age directly about sensitive questions given the overall situation in the country. Moreover, we felt that a broad based assessment of how children in different types of schools perform in terms of the overall program of nation-building and shared values as future participating citizens was perhaps more informative about the role of schools in the formation of civic ideology.

cohort of students for select questions from the exams we administered. Students' performance on the subject and civic values tests was poor relative to the curriculum. The average child in our sample can read simple words in the vernacular, Urdu, can recognize alphabets, can match simple words to pictures in English, and can add single digit numbers in Mathematics. He or she cannot, however, give antonyms in Urdu, construct an English sentence with words like "deep" or "play," or complete a division problem. Broadly, private school students start at a higher proficiency and are more likely be able to complete a question by the 4th year. For example, for a harder question such as dividing 384 by 6, 16 (22) percent of public (private) school students could answer it correctly in round 1, while 49 (62) percent could in round 4.

Although the results suggest poor academic performance, they are in line with other studies from low-income countries (and for civics, from the U.S. as well). For example, in India, only 52 percent of students between the ages of 7 and 10 could read a small paragraph with short sentences at the grade 1 level, and 52 percent were unable to divide or subtract (Pratham, 2005). On civics, the 2006 NAEP results from the U.S. show that 46 percent of fourth graders knew the cause of the civil war and only 24 percent knew why the United States expanded westward.

On the civics questionnaire, 46 percent knew that India neighbors Pakistan (the other choices were the U.S., Saudi Arabia, and Kuwait) with a significant fraction answering Saudi Arabia. While 97 percent correctly identified the founder of Pakistan, 40 percent got the current Prime Minister wrong, and 50 percent incorrectly identified India or China as the country that Pakistan got its independence from (it was Britain). Finally, only 30 percent could complete the country's national slogan, and the majority (63 percent) preferred to donate money in the case of disasters to private entities or nonprofits rather than government entities. Only 18 percent thought that "voting" was the best way to decide on what to eat for lunch relative to handing the decision over to a central authority (the teacher, the class monitor, or the smartest student).

Household Survey. Our final data source is the household survey, administered to 1,740 households in the 112 villages in each of the survey years. These households were picked randomly following a household census in 2003. Twelve households out of a total of 16 were chosen among those with a child attending grade 3 in 2003 and the remaining 4 were chosen among households with a child eligible to attend grade 3 but who was not currently attending. The stratification was designed to increase the number of children with matching information from the household-survey and test scores from the school-based testing exercise. The household survey collected detailed information on socio-demographic characteristics, consumption expenditures and assets, with additional specialized modules on education and time-allocation. Additionally, GPS coordinates were taken during the first round of household surveys, allowing us to calculate the distance between each household and every school in the village.

Combined Samples. For our main analyses, we create two datasets. For our SVA, school switcher, and school closures instrumental variables identification strategies, we do not need to observe the distance between a child and school. Therefore, for these analyses we create an unbalanced panel by combining the test score data with the in-school child surveys. This creates a sample of 31,382 unique children (71,766 child-year observations). Summary statistics for this data are given in Table 2. For our second instrumental variables strategy, where we instrument for private schooling using information on distance, we must rely on the household data. Thus, for this strategy, we combine the test data with the sample of children in surveyed households. This results in a smaller sample of 1,269 students (3,959 student-years). Summary statistics for these data are given in Table 3.

As Table 3 shows, the children we follow are 45 percent female and 11 years old on average. Their parents are relatively uneducated, with 29 percent of mothers having any primary schooling and 65 percent of fathers having any primary schooling. Private school students are systematically wealthier and more likely to have educated parents than public school students. 46 percent of private school mothers have some education (relative to 25 percent of public school mothers), and 72 percent of private school fathers have some education (relative to 63 percent of public school fathers).

Twenty-four percent of our sample was in private school when they were surveyed, and Tables 2 and 3 show the differences between children in public and private schools. These are large in terms of cognitive achievement, as well as household characteristics. Children attending private schools are slightly younger and come from wealthier and more educated households. Test scores are much higher for children in private relative to public schools. Private school children score 0.45 standard deviations higher on our Mathematics test and 0.55 standard deviations higher on our Urdu test. They score 0.78 standard deviations higher on our English exam. Performance on civics values questions is similar. Private school students score 0.23 standard deviations better on the civic knowledge index, which takes the average score across civics knowledge questions, and 0.10 standard deviations higher on the civic disposition index, which takes the average across questions on pro-government attitudes.

3 Measuring the Effects of Private Schooling

In this section, we outline the four types of empirical strategies we use to measure the effects of attending a private school and report their results.

3.1 Value-Added Approach

Empirical Strategy. For our first empirical strategy, we exploit the fact that we observe past test scores and use these test scores to control for selection of students into schools. This identifying

assumption is similar to the identifying assumptions used by Singh (2015) to study private school quality in India, Chetty et al. (2014) to study teacher quality in the United States, and Bau and Das (forthcoming) to study teacher quality in Pakistan. Our regression specification is

$$y_{igst} = \beta_0 + \lambda_g y_{igs,t-1} + \beta_1 private_{ist} + \alpha_g + \alpha_t + \mathbf{\Gamma X_{igst}} + \varepsilon_{igst}, \quad (1)$$

where i denotes a student, g denotes a grade, s denotes a school, and t denotes a year. Then, y_{igst} is a test score, $private_{ist}$ is an indicator variable equal to 1 if student i attends a private school in year t, $\mathbf{X_{igst}}$ is a vector of controls, α_g is a grade fixed effect, α_t is a year fixed effect, and λ_g, the coefficient on the lagged test score, is allowed to vary at the grade-level. In the basic specification, $\mathbf{X_{igst}}$ also includes controls for age, age squared, an indicator variable for female, and the interactions between female and year fixed effects and the age controls. In our more stringent specifications, $\mathbf{X_{igst}}$ also includes indicator variables for whether the father and mother have at least some education, and a household asset index,[9] all interacted with gender. Then, β_1, which estimates the effect of a year of private schooling, is the coefficient of interest. Standard errors are clustered at the school-level.

In addition to math, Urdu, and English, we also report estimates of the private schooling effect on civics. However, because civics were only tested in round 3 (2005-2006), we cannot include lagged civic scores. Thus, these results should be interpreted as associations rather than causally.

Results. Table 4 reports the estimates of β_1 from equation (1). Private schooling has positive and significant effects on all three subjects, though the largest effects are on English, consistent with the findings of Singh (2015) in India. An additional year of private schooling increases math test scores by 0.153 sd, English scores by 0.229 sd, and Urdu scores by 0.159 sd. The inclusion of additional socioeconomic controls in the even columns has little affect on the estimates. In terms of magnitude, we can compare the effect on mean test scores (0.13 sd) to average yearly learning in public schools (0.39 sd). Thus, the additional gain in learning from attending a private school each year is equivalent to one-third of an extra year of public school.

Table 5 reports the association between private schooling and the civics measures. For all of the civics indices, private schooling is positively and significantly related to civics. It is positively correlated with the Pakistan knowledge index and pro-government attitudes, and negatively correlated with gender bias. The inclusion of controls again has little effect on these relationships.

[9]The household asset index is formed by performing a principle component analysis of indicator variables for all the assets reported by the household. The index is the first factor predicted by this analysis.

3.2 School Switching: Child FE

Empirical Strategy. Our second empirical strategy exploits variation due to students switching schools to estimate the effect of private schooling. We include child fixed effects in equation (1) so that the new estimating equation is

$$y_{igst} = \lambda_g y_{igs,t-1} + \beta_1 private_{ist} + \mathbf{\Gamma X_{igst}} + \alpha_g + \alpha_t + \alpha_i + \varepsilon_{igst}, \tag{2}$$

where α_i is the child fixed effect. This fixed effect controls for any child-level time invariant unobservables, such as socioeconomic status or innate ability. Additionally, including this fixed effect means that β_1 is now identified by comparing the change in test scores for children who switch into or out of private schools to those who don't over time.

Validity of Strategy. One concern is that while child fixed effects control for time invariant unobservables, they may not account for time variant unobservables that are correlated with attending a private school. This is analogous to the violation of the parallel trends assumption in a difference-in-differences regression. To assess whether parallel trends hold prior to a child switching schools, we estimate

$$y_{igst} = \lambda_g y_{igs,t-1} + \sum_{k \in \{-2,..,2\}} \tau_k event\, k_{ist} + \mathbf{\Gamma X_{igst}} + \alpha_g + \alpha_t + \alpha_i + \varepsilon_{igst}, \tag{3}$$

where $k = 0$ is normalized to be the period when a student switched schools, and $event\, k_{ist}$ is an indicator variable equal 1 if it is k periods after an event. Thus, if trends are parallel, we expect that τ_k will be 0 before a child switches to a public/private school and take on the private school effect after the switch occurs.

Results. Table 6 reports the child fixed effect estimates with and without the socioeconomic controls from equation (2).[10] The effect on mean test scores is virtually identical to the effect estimated in Table 4. Figure 3 reports the coefficient estimates for equation (3) on mean test scores. While the number of pre-periods is limited, there is no evidence of pre-trends. Moreover, the increase in test scores happens exactly when children switch into a private school. Altogether, these results suggest that time invariant child unobservables are not biasing the value-added estimates.

3.3 School Switching: School Closure IV

Empirical Strategy. While estimating equation (2) is more conservative than equation (1), β_1 may still be biased since school switching is not exogenous. Children whose test scores are on an

[10] While mother and father education are time invariant, household asset index does vary over time and is not collinear with child fixed effects.

upward trajectory – e.g. because families have experienced positive wealth shocks – may be more likely to switch into private schools. Thus, our second approach exploits exogeneous switches due to private school closures. To do so, we restrict the sample to children who attended private schools when they were first observed. We then instrument for attending a private school with an indicator variable that is equal to 1 if the private school those students attended has been closed. The first stage is then

$$private_{ivst} = \beta_0 + \lambda_g y_{igs,t-1} + \mu_1 closure_{it} + \mathbf{\Gamma X_{igst}} + \alpha_g + \alpha_t + \alpha_v + \varepsilon_{igst}, \quad (4)$$

where v indexes a village, the instrument $closure_{it}$ is an indicator variable equal to 1 if a private school previously attended by student i has closed, and α_v is a village fixed effect. The second stage is then equation (1) except that it now controls for village fixed effects.

Since we only observe civics measures in round 3, we could not estimate the effect of private schooling on these outcomes using the child fixed effects strategy. However, the private school closure strategy does not require multiple rounds of outcome measures. Thus, we can use this strategy to estimate the causal effect of private schooling on civics outcomes with the caveat that we do not observe lagged civic values and therefore, cannot calculate yearly civic value gains. Estimates of the effect of private schooling on civic values should be interpreted not as the magnitude of the effect of one year of private schooling but as the net effect of attending private school across multiple years.

Before reporting the school closure instrumental variables estimates, in Appendix Table A2, we evaluate whether the instrument is correlated with children's characteristics by regressing student characteristics in the child survey on the instrument. The instrument is not significantly related to whether a child's mother has primary schooling, household wealth (as measured by the first principal component of the household's assets) or the child's gender, though it is marginally significantly (and negatively) related to father's education and the school's assessment of the child's ability. However, these coefficients are small. If we run a regression of school closure on all the outcomes in Appendix Table A2, and perform a joint test of their significance, the F-statistic is 1.49 (p=0.202).

Results. Table 7 reports the estimates that instrument for private schooling with school closures. Column 1 reports the first stage estimate and shows that a closure reduces the subsequent probability that a student attends a private school by a statistically significant 25%. We again find that the private schooling has positive and significant effects on test scores. The results are somewhat larger than before, with private schooling leading to test score gains of 0.38sd, though the confidence interval for mean test score gains includes our value-added and child fixed effects estimates.

Table 8 reports the instrumental variables estimates for the civics outcomes. Private schooling

appears to lead to higher scores on the full civics index (which takes the average over scores on all civic questions,[11] greater knowledge of Pakistani history and institutions, and more pro-government attitudes. The coefficient for gender bias also indicates that private schooling decreases gender bias, though this effect is not significant.

3.4 Distance IV

Empirical Strategy. Our second instrumental variables strategy exploits the fact that students are highly distance sensitive when they choose schools and uses distance to the closest private school relative to the closest public school as an instrument for private school enrollment. The validity of this instrument relies on two key characteristics of Pakistani villages.

First, historically, when villages were settled, richer households settled in the center of the village and poorer households settled on the periphery. In his discussion of settlement patterns in Punjab, Paustian et al. (1930) details the British administrative strategy of building water canals and leasing land in order to settle previously uninhabited regions. These were planned villages around the new canal projects, where settlers were chosen among those deemed "fit" by the British government and assigned land. Land grants to the initial settlers ranged from 22.5 to 27.5 acres – remarkably large farms in this context.[12] Later migration from other villages and (after 1947) from India led poorer migrants to settle on the village periphery. As a result, in 1930, Paustian notes, "The inner group of village houses is generally occupied by the peasants who till the land. The outer houses of the village are occupied by the village menials and artisans." Thus, both the wealth endowment to the initial settlers and the selection of "fit" individuals with exceptional farming skills ensured that the center of the village was occupied by wealthier individuals. These canal colonies, as they are known, are common to many parts of Punjab, including all the villages in the district of Faislabad and the majority in Rahim Yar Khan (the districts in our study that are in the center and the South).

The second key feature of Pakistani villages is that public schools are more likely to be located on villages' outskirts. Many of these schools were constructed under the Social Action Programs of the 1980s and 1990s. These were large programs of school construction supported by external donor funding. These programs required the provision of land (5 Kanals, or .625 acres) for the school by the village, and land often came from common village property known as the *shamlaat*. Land in the shamlaat could not be bought or sold without a bureaucratic "land conversion" process, much like the re-zoning of land in the United States. Consequently, the shadow price of shamlaat land was always lower than that of privately owned land. Unsurprisingly, at the time of school

[11] Gender bias questions are recoded so positive values no longer indicate greater gender bias.

[12] According to the census report of 1868, for instance, the cultivated area in Punjab amounted to 1.25 acres per capita, of which irrigated land was only 0.06 acres per capita. Grants of 22.5 to 27.5 acres of irrigated land represented a sizeable gain in agricultural capacity for the original settlers.

construction, many villages donated land in the shamlaat rather than attempt to collectively purchase expensive private land in the center of the village. As a result, a significant fraction of public schools were located on the outskirts of villages. In contrast, private schools typically locate near a village's center to be closer to richer families and to reduce their distance to the largest number of households.

The instrument exploits variation in the distance between a household and the closest private school relative to the closest public school. On its own, relative distance will clearly be correlated with the distance to a village's center as private schools are likely to place close to the center and public schools place on the outskirts. Thus, we calculate (see a discussion of this calculation in Appendix A) and directly control for distance to the center. Consequently, the variation we exploit comes from (for example) two households that are on the periphery of the village but one happens to be on the same side of the village as the public school while the other is on the opposite side. Appendix Figure A1, which plots school and household locations for all villages in the data, where the center of each village has been normalized to be at (0,0), illustrates the variation we use to identify the private school effect. Appendix Table A3 verifies that conditional on controlling for distance to the center, relative distance to private schools is indeed unrelated to measures of parental education, assets, consumption, or family size. In a regression of the instrument on the individual characteristics in a the table, the F-statistic from a joint test of those characteristics is 0.67 (p=0.752).

To use distance to identify the effect of private schooling on learning and civic values, we need a sample with both distance information (the household sample) and test scores information (the tested sample). However, restricting our sample to the individuals who appear in both data sets greatly reduces our sample size and throws out information from the household survey on the relationship between private schooling and distance. Therefore, for our main results using this identification strategy, we use the two sample 2SLS methodology of Inoue and Solon (2010). This methodology allows us to estimate the instrument and the endogenous variable using the full sample of children for whom we observe distance information and enrollment information in the household sample, even if those children were not tested. We can combine this with estimates of the effect of the instrument on test scores (using the full sample of children for whom test scores and the instrument are available) to back-out the IV estimate.

The first stage is

$$\begin{aligned} years\,private_{igst} =& \mu_1(Distpri_i - Distgov_i) + \mu_2(DistCenter_i) + \mu_3(DistCenter_i) \times female_i \\ &+ \mathbf{\Gamma X_{igst}} + \alpha_v + \alpha_g + \alpha_t + \varepsilon_{igst}, \end{aligned}$$

where $years\,private_{ist}$ is the number of years a child has attended private school by time t, the

instrument $(Distpri_i - Distgov_i)$ is the difference between the distance to the closest private and public schools, $DistCenter_i$ is the distance to the village center, and $female_i$ is an indicator variable equal to 1 if the child is female. The controls $\mathbf{X_{igst}}$ include age controls, their interaction with gender, and the interactions between gender and the grade and year fixed effects. We focus on $years\ private_{ist}$ instead of private schooling as our variable of interest since the instrument varies little over time and we cannot control for lagged test scores in the two-sample IV strategy since we do not observe them for non-tested children. Using $years\ private_{ist}$ as our endogenous variable ensures that the coefficient we estimate can still be interpreted as the effect of one additional year of private schooling. The second stage is then the same specification except y_{igst} is the outcome variable and the instrument $(Distpri_i - Distgov_i)$ is replaced with the endogenous variable $years\ private_{igst}$.

Results. Table 9 reports the results from the two sample IV for test scores. The first column indicates that a 1 km increase in the relative distance to a private school decreases a child's years spent in private school by about one-third of a year. The effect of a year of private schooling on mean test scores (0.155 sd) is very close to the value-added estimate. Appendix Table A4 reports the results when we use a standard IV using the subsample of participants for whom distance, test score, and years of private schooling information are available. Doing so greatly reduces the sample size, and the resultant point estimates are no longer significant, though still positive.

Table 10 reports the two sample IV estimates for civic values. A year of private schooling marginally significantly improves the full civic value index by 0.03 sd, improves knowledge of Pakistan by 0.03 sd, and reduces male biased responses by 0.159 sd. Appendix Table A5 reports the instrumental variables estimates for civic values that do not use the two-sample strategy. The sample size is again much smaller, and the point estimates are again imprecise but all are consistent with private schooling improving the civics measures.

4 Heterogeneity in School Quality

The previous section provides evidence that on average private schools out-perform public schools, producing higher test scores and possibly also improving civic values. In this section, we dig deeper into these average effects, and characterize the heterogeneous quality of public and private schools. Guided by the fact that value-added specifications deliver similar estimates of the effect of attending a private school to the child fixed effects and instrumental variables strategies, in the first subsection, we construct a value-added measure for each school in the data. In the second subsection, we validate that these measures are forecast unbiased. In the third subsection, using these estimates, we report new facts about the dispersion in school quality and the correlates of school quality.

4.1 Estimating School Value-Added

Our specification for estimating school value-added is very similar to the specification for estimating teacher value-added used by Bau and Das (forthcoming) in Pakistan and is very similar to our value-added specification for estimating private school effects (equation (1)). Specifically, for each subject, as well as for mean test scores, we estimate

$$y_{igst} = \beta_0 + \lambda_g y_{igs,t-1} + \alpha_s + \alpha_g + \alpha_t + \varepsilon_{igst}, \tag{5}$$

where α_s is a school fixed effect. Intuitively, α_s captures the average test score gains of students in a school s after accounting for observable factors such as past achievement. It will be an unbiased estimate of the predicted test score gains of a child in a school as long as the controls account for the sorting of children into schools.[13] We note that α_s includes both the average effect of the teachers in a school, as well as any independent school effect. Indeed, since we do not observe teachers changing schools, we cannot separately estimate school and teacher effects. However, for our purposes, this is not a problem. For a policy-maker or household considering the effects of shifting a child into a different school, the effect of the school on test scores (including the teacher effect) is the object of interest.

Even if α_s is unbiased, it is estimated with error. That means that taking the variance of α_s for public and private schools will over-estimate the total variance of school quality because this variance will include both the true variance of school quality and the estimation error. Similarly, if we want to use the school value-added as an explanatory variable in a regression, estimation error attenuates its coefficient. To address these issues, we use our fixed effect estimates α_s to construct empirical Bayes estimates of school value-added, as is common in the teacher value-added literature. This process is described in Appendix B.

4.2 Validating SVA

Event Study Approach. Our SVA estimates are only unbiased under the assumption that the controls in equation (5) account for the sorting of students into schools. The fact that our alternative identification strategies deliver very similar results to the value-added strategy already suggests that this is the case. However, to further evaluate this assumption, we turn to an out-of-sample prediction test similar to validations of teacher value-added by Chetty et al. (2014) and Bau and Das (forthcoming). Specifically, we focus on the sample of students that switch schools and evaluate (1) whether the SVA of a student's new school predicts test score gains before they enter the school (indicating sorting), and (2) whether the SVA predicts test score gains with a coefficient of

[13] Chetty et al. (2014) shows that lagged test score controls control for most sorting of students to teachers in a large U.S. school district.

approximately 1 after students enroll in the new school (indicating forecast unbiasedness). To do so, we estimate

$$y_{igst} = \lambda_g y_{igs,t-1} + \sum_{k\in\{-2,..,2\}} \tau_k event\ k_{ist} \widehat{SVA_s} + \mathbf{\Gamma X_{igst}} + \alpha_g + \alpha_t + \varepsilon_{igst}, \quad (6)$$

where s here denotes the school that a child switches into, so that $\widehat{SVA_s}$ is the estimated SVA of the school into which the child switches.

Figure 4 graphs the τ_k values from equation (6). While we have limited pre-periods, there is no trend prior to a school switch. Following a switch, however, $\widehat{SVA_s}$ is highly predictive of test score gains with a coefficient of approximately 1, consistent with forecast unbiasedness.

Comparison to Child FE Estimates. As an alternative test of whether SVA are forecast unbiased, we also test if our SVA measures can be used to replicate the estimated private school effect from the child fixed effect specification (equation (2)). This specification is more stringent than the value-added specification in the sense that it controls for child specific unobserved characteristics. If these characteristics bias the SVA estimates, then we would expect to estimate very different private school effects using the child fixed effect specification and the SVA estimates.

To see if this is the case, for each child who switches between a public and private school, we use the difference between the empirical Bayes SVA measures for the two schools to calculate the effect of a public to private switch (or vice versa). If the SVA are forecast unbiased, the average of the differences between the public and private school quality measures for the switchers should closely approximate the estimated effect of attending a private school on test scores from the child fixed effect regressions. Table 11 reports the two sets of estimates, and shows that the two methods do deliver very similar estimates. For mean test scores, the child FE estimates indicate that private schooling would increase test scores by 0.168 sd. The predicted effect using the SVA estimates is 0.164 sd.

4.3 Distribution of SVA & Correlates of SVA

Distribution of SVA. Having estimated and validated SVA in the private sector, we can now use our estimates to characterize the distribution of school quality in both sectors. Figure 5 graphs the empirical Bayes estimates of SVA for math, English, Urdu, and mean test scores in the public and private sector. Consistent with the estimates in the first half of the paper, on average, private schools are higher quality in every subject. Moreover, the private advantage is especially pronounced in English. However, in both sectors, quality is also highly variant. Using a formula derived in Appendix B, which accounts for estimation error in the SVA measures, to estimate the variances of the SVA distributions, Table 12 reports the effects of attending a 1 sd better school on test scores by subject in the private and public sectors. Attending a school that is 1 sd better in the private

school distribution increases mean test scores by 0.208 student test score standard deviations, while attending a 1 sd better public school increases them by 0.316 sd. Moving from the average public school to a 1 sd better public school would almost double annual test score gains.

The benefits of moving to private schooling are highly dependent on where in the distribution a student's public and potential private school are. As Figure 5 illustrates, moving a student from a public school at the 5th percentile in the public distribution to a private school at the 5th percentile in the private distribution would only result in mean test score gains of 0.056 sd. However, moving from the 95th percentile of the public to private distribution results in gains of 0.283sd. Moving from a public school at the 20th percentile to a private school at the 20th percentile would result in gains of 0.111 sd. Thus, the benefits of policies that move students between sectors will be highly dependent on what public and private schools they move students between.

Explaining Variation in SVA. Next, to further assess the variance in students' choice sets, we examine what percent of school quality is explained by cross-village as opposed to within-village differences. For both the private and public sectors, we regress the SVA estimates on village fixed effects. The resulting R^2 tell us what percent of the variation in SVA is within- vs. across-villages. We find that 43% of the variation in both mean private school and public school quality is across villages rather than within-villages.

We next consider what observable school characteristics are predictive of school quality in the public and private sectors and to what extent these characteristics explain the variation in quality. Appendix Table A6 regresses SVA on school facilities, a student teacher ratio measure, and the number of private and public schools in a village conditional on village size separately for the public and private sectors. Generally, characteristics are not highly predictive of school quality, though they are more predictive in the private sector. In the public sector, school facilities and student-teacher ratios alone explain only 1-2 percent of the variation in SVA, while in the private sector, they explain 6-8 percent. Interestingly, in the private sector, SVA is also positively associated with the number of public schools, suggesting that private school quality may respond to students' choice set. Altogether, information on schools' observable characteristics, particularly in the public sector may not be sufficient for parents to identify high SVA schools.

Who Attends High SVA Schools? Given that it is difficult to observe which schools are high SVA, a natural next question is whether more advantaged families attend higher SVA schools *within* the public and private sectors. Appendix Table A7 reports the results from regressing school SVA's on students' socioeconomic characteristics. Since public education is not co-educational, we run these regressions separately for male and female students in both the public and private sectors. The regressions indicate that even within the public sector, students from more advantaged backgrounds attend higher SVA schools. However, the inclusion of village FE eliminates this

relationship, suggesting that it is entirely explained by variation in SVA across villages. In the private sector, on the other hand, both within- and across-villages, both male and female students from more advantaged backgrounds attend better schools.

4.4 Voucher Counterfactuals

Given the large dispersion in school quality in both sectors – and the fact that students' ability to select a better school may be limited since much of the variation in quality is across rather than within villages – we next evaluate the potential range of treatment effects for programs that move students from the public to the private sector. Taking school quality as fixed, we calculate the upper and lower bound gains from instituting voucher systems. In the lower bound case, all public school students are reassigned to the worst private school in their village, while in the upper bound, all public school students are reassigned to the best private school in their village. Table 13 reports the partial equilibrium average and distributional effects of such a program on test scores in each subject. In the upper bound case, mean test scores increase by 0.23 sd. In the lower bound case, mean test scores increase by only 0.07 sd.

5 Conclusion

In this paper, we provide causal estimates of the impact of private schooling on test scores in a low-income country. We show that the average causal impacts of private schooling on test scores are positive and meaningful. On average, the private school quality distribution is to the right of the public school distribution in every topic. Furthermore, we are able to show that there are no associated costs with these learning gains in terms of civic values. If anything, private schools appear to improve civic values.

A natural question is why public and private school students may have different civic value outcomes, particularly since private schools use the same textbooks as public schools. Given evidence in favor of the significant differences in civic value outcomes, it is likely that civic values are learned experientially, as in Otsu (2001). The experience of a public school in Pakistan – with high absenteeism and little reward for better performance – may be counterproductive for the instillation of civic values.

Additionally, we demonstrate that with only a few years of data, measures of SVA can be credibly constructed in low-income contexts. These measures allow us to characterize the distribution of school quality in both the public and private sectors. Large dispersion in quality means that the effects of policies that shift students between the sectors will heavily depend on the identity of the marginal student and the marginal school. Thus, we conclude that the design and implementation of policies such as voucher systems matters greatly for their effects.

Finally, estimating SVA allows us to document several additional, new facts about education in

a low-income context. First, We find that much of the variation in school quality is across rather than within-villages. Second, we find that observed school characteristics are not predictive of school quality, particularly in the public sector. Finally, even *within* the public sector, wealthier families attend better public schools, as more advantaged families appear to live in villages with better public schools.

References

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of political economy*, 2005, *113* (1), 151–184.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan," *Comparative Education Review*, 2008, *52* (3), 329–355.

_ , _ , and _ , "Students today, teachers tomorrow: Identifying constraints on the provision of education," *Journal of public Economics*, 2013, *100*, 1–14.

_ , _ , _ , and Tristan Zajonc, "Religious School Enrollment in Pakistan: A Look at the Data," *Comparative Education Review*, 2006, *50* (3), 446–477.

_ , _ , _ , Tara Vishwanath, and Tristan Zajonc, "Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to Inform the Education Policy Debate," *World Bank, Washington, DC*, 2007.

Barrett, Angeline, Sajid Ali, John Clegg, J Enrique Hinostroza, John Lowe, Jutta Nikel, Mario Novelli, George Oduro, Mario Pillay, Leon Tikly et al., "Initiatives to improve the quality of teaching and learning: A review of recent literature," *Background paper for Global Monitoring Report*, 2007.

Bashir, Sajitha, "Achievement performance at the primary level in public and private schools of Tamil Nadu," *Indian Educational Review*, 1994, *29* (3-4), 1–26.

Bassey, Magnus O, *Western education and political domination in Africa: A study in critical and dialogical pedagogy*, Greenwood Publishing Group, 1999.

Bau, Natalie, "Estimating an Equilibrium Model of Horizontal Competition in Education," *Working Paper*, 2019.

_ and Jishnu Das, "Teacher Value-Added in a Low-Income Country," *American Economic Journal: Economic Policy*, forthcoming.

Carneiro, Pedro Manuel, Jishnu Das, and Hugo Reis, "The value of private schools: Evidence from Pakistan," 2016.

Chetty, Raj, John Friedman, and Jonah Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 2014, *104* (9), 2593–2632.

Cohn, Bernard S and David Scott, *Colonialism and its forms of knowledge: The British in India*, Princeton University Press, 1996.

Das, Jishnu and Tristan Zajonc, "India Shining and Bharat Drowning: Comparing Two Indian States to the Worldwide Distribution in Mathematics Achievement," *Journal of Development*

Economics, 2010, *92* (2), 175–187.

Dean, Bernadette L, "Citizenship education in Pakistani schools: Problems and possibilities," *International Journal of citizenship and Teacher education*, 2005, *1* (2), 35.

Filling the Gaps: Private Education is Stepping In Where the State Leaves Off

***Filling the Gaps: Private Education is Stepping In Where the State Leaves Off**, Economist, 2019.*

Habib, Masooma, *"Education in Pakistan: All Demand and Little Supply," in "Pakistan Economic and Social Development Conference" 1998.*

Hoxby, Caroline M and Sonali Murarka, *"Charter schools in New York City: Who enrolls and how they affect their students' achievement," Technical Report 2009.*

Inoue, Atsushi and Gary Solon, *"Two-sample instrumental variables estimators," The Review of Economics and Statistics, 2010, 92 (3), 557–561.*

Jimenez, Emmanuel, Marlaine E Lockheed, and Vicente Paqueo, *"The Relative Efficiency of Private and Public Schools in Developing Countries," The World Bank Research Observer, 1991, 6 (2), 205–218.*

_ , Marlaine Lockheed, and Nongnuch Wattanawaha, *"The relative efficiency of private and public schools: The case of Thailand," The World Bank Economic Review, 1988, 2 (2), 139–164.*

Kingdon, Geeta, *"The quality and efficiency of private and public education: a case-study of urban India," Oxford Bulletin of Economics and Statistics, 1996, 58 (1), 57–82.*

Kingdon, Geeta G, *"Private schooling in India: Size, nature, and equity-effects," Economic and Political Weekly, 1996, pp. 3306–3314.*

Kremer, Michael and Karthik Muralidharan, *"Public and Private Schools in Rural India," School Choice International, 2008.*

Low Cost Private Schools: Learning Unleashed

Low Cost Private Schools: Learning Unleashed, *Economist*, 2015.

Meyer, John W, David Tyack, Joane Nagel, and Audri Gordon, "Public education as nation-building in America: Enrollments and bureaucratization in the American states, 1870-1930," *American journal of Sociology*, 1979, *85* (3), 591–613.

Muralidharan, Karthik and Venkatesh Sundararaman, "Contract Teachers: Experimental Evidence From India," *NBER Working Paper #19440*, 2013.

_ **and** _ , "The aggregate effect of school choice: Evidence from a two-stage experiment in India," *The Quarterly Journal of Economics*, 2015, *130* (3), 1011–1066.

NCAER, "National Council of Applied Economic Research Annual Report," 2005.

Neal, Derek A, "What have we learned about the benefits of private schooling?," *Economic Policy Review*, 1998, *4* (1).

Otsu, Kazuko, "Civics education in transition: the case of Japan," *International Journal of Educational Research*, 2001, *35* (1), 29–44.

Pakistan Bureau of Statistics, "Pakistan Integrated Household Survey," 1991.

_ , "National Education Census, Pakistan," 2005.

Paustian, Paul William et al., "Canal irrigation in the Punjab," 1930.

Pratham, "Annual Status of Education Report," 2005.

_ , "Annual Status of Education Report," 2010.

Pritchett, Lant and Martina Viarengo, "The state, socialisation, and private schooling: When will governments support alternative producers?," *The Journal of Development Studies*, 2015, *51* (7), 784–807.

Shapiro, Jacob N and C Christine Fair, "Understanding support for Islamist militancy in Pakistan," *International Security*, 2010, *34* (3), 79–118.

Singh, Abhijeet, "Private School Effects in Urban and Rural India: Panel Estimates at Primary and Secondary School Ages," *Journal of Development Economics*, 2015, *113*, 16–32.

Tooley, James and Pauline Dixon, "Private schools for the poor: A case study from India," *Reading, UK: Centre for British Teachers*, 2003.

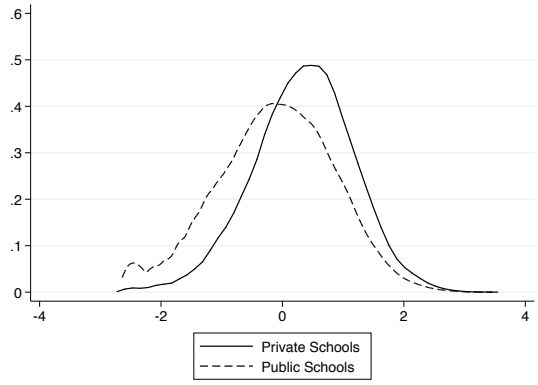
Wolf, Patrick J, "Civics Exam Schools of choice boost civic values," *Education Next*, 2007, *7* (3), 66.

_ **and Stephen Macedo**, *Educating citizens: International perspectives on civic values and school choice*, Brookings Institution Press, 2004.

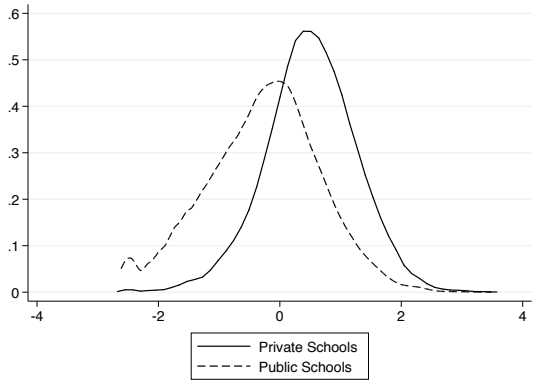
World Bank, "World Bank Development Data Platform," 2019.

Figures

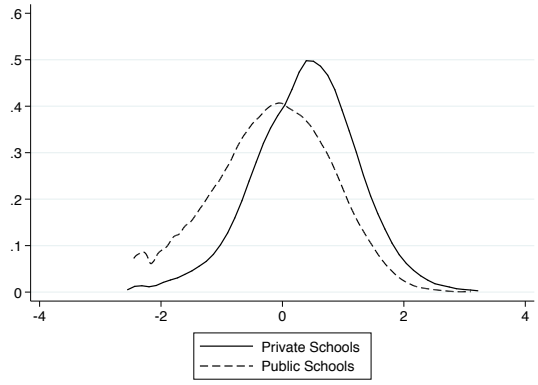
Figure 1: Test Scores in Public and Private Schools



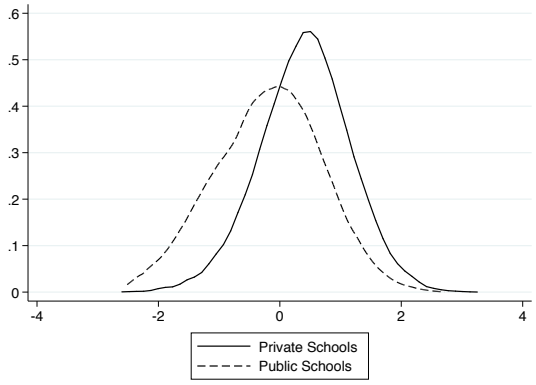
(a) Mathematics



(b) English



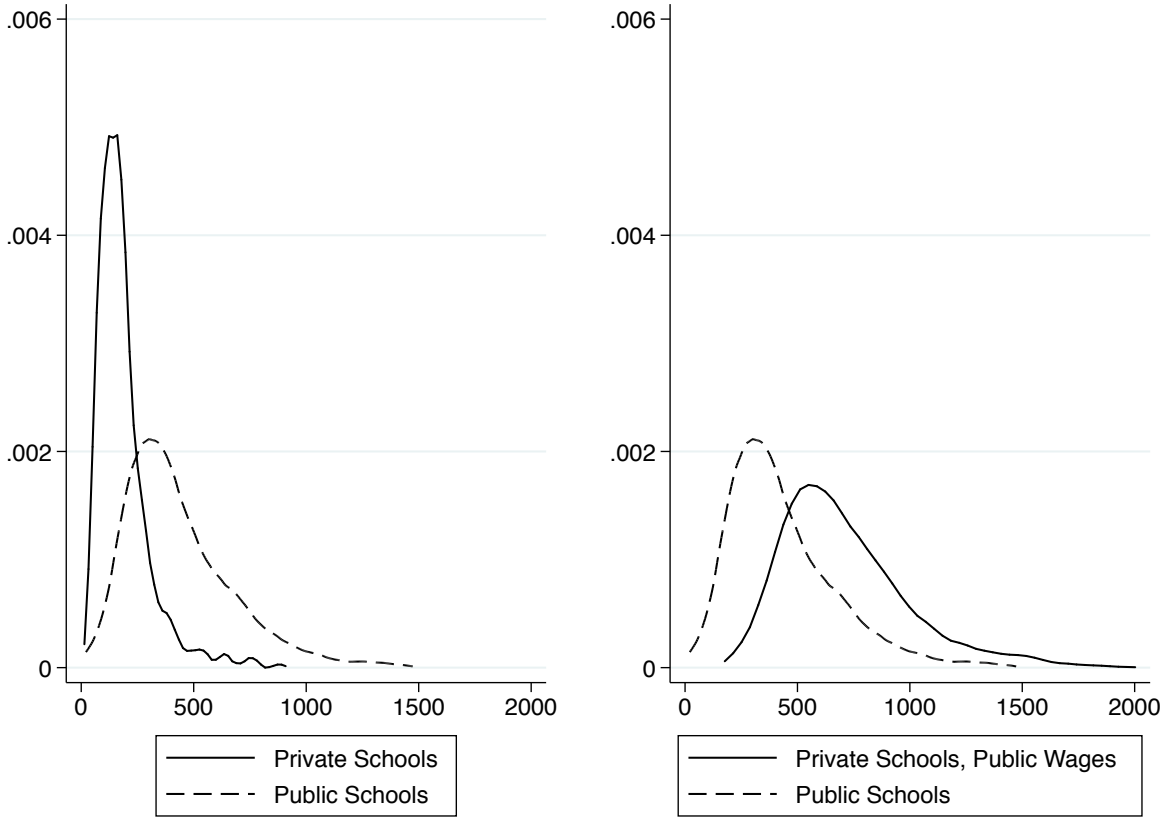
(c) Urdu



(d) Mean

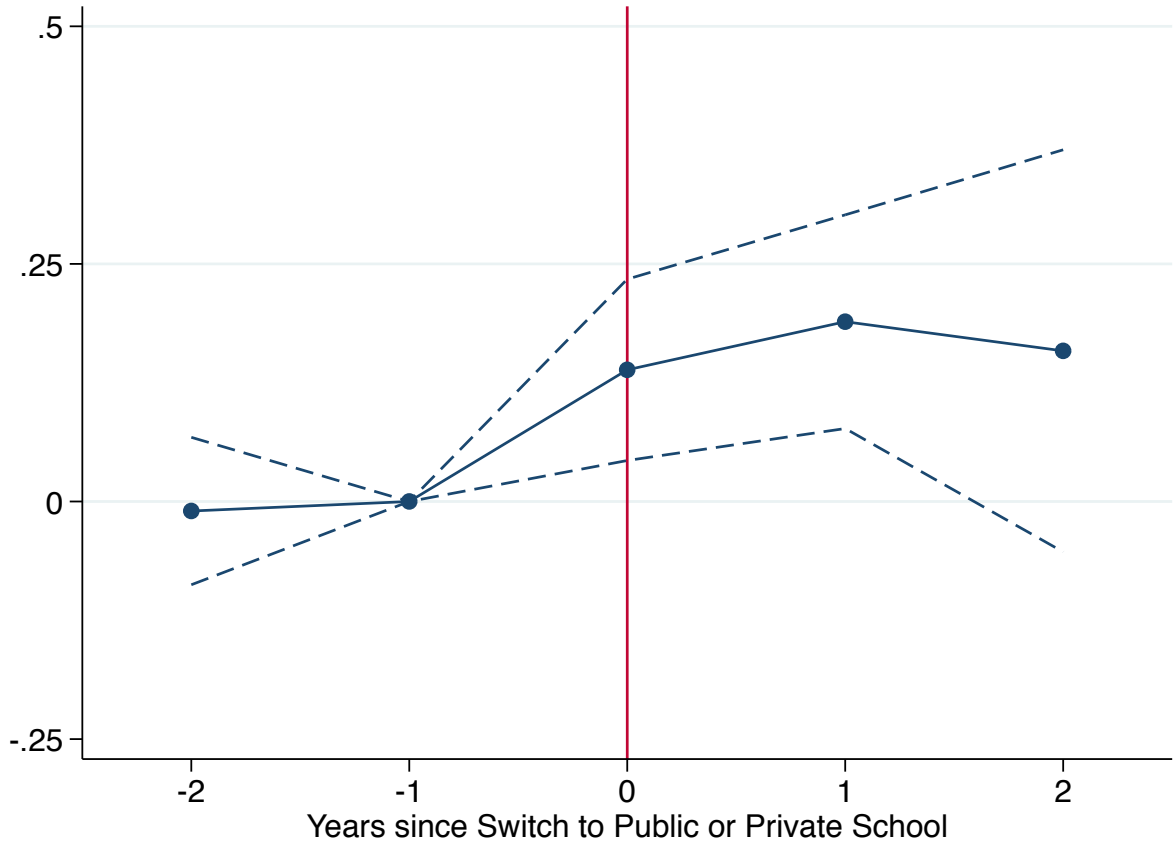
Notes: This figure plots the distribution of test scores for students enrolled in private and public schools, respectively. The mean test score is the average over test scores in mathematics, Urdu and English.

Figure 2: School Expenditures per Student in Public and Private Schools



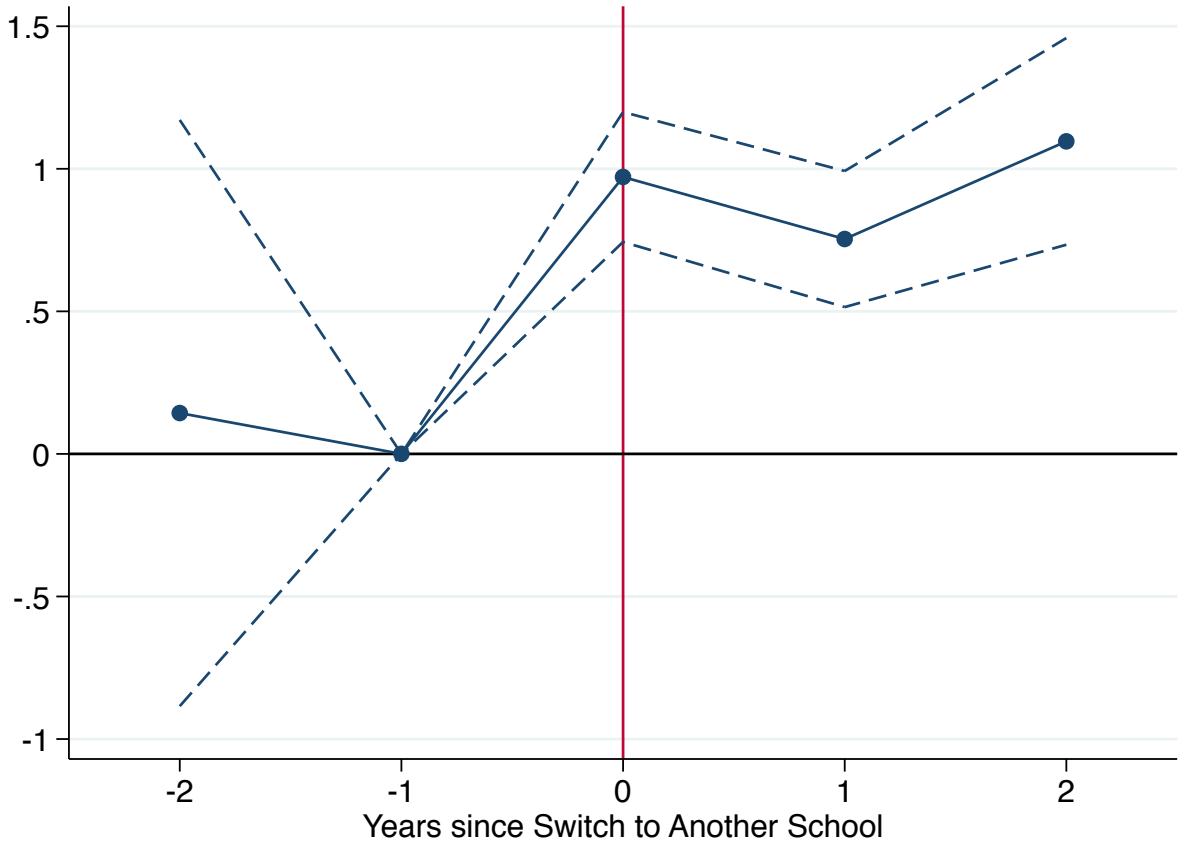
Notes: The left panel shows the distribution of public and private schools total costs per student in the data. The right panel shows public and private schools total costs per student if private schools were to pay their teachers at the reported village average public school teacher wage. Total expenditures are converted to 2010 dollars using the consumer price index for Pakistan. Top and bottom 1 percent values are excluded.

Figure 3: Event Study Graph for Private School Effect (Child FE Estimates)



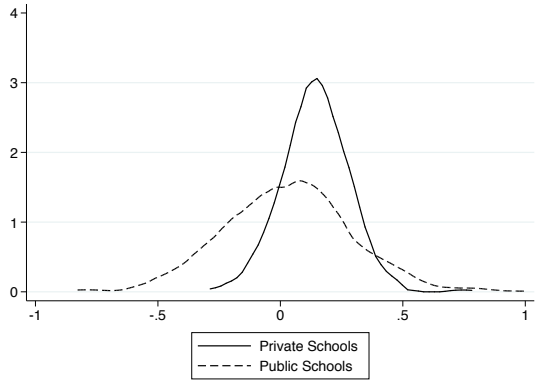
Notes: This graph reports estimates of the effect of being in a private school t years after a switch to a public or private school for $t \in \{-2, ..., 2\}$. A switch is coded as taking the value 1 if a child switches from a public to private school and -1 if she switches from a private to public school. The vertical red line at $t = 0$ identifies the year in which a child switches the type of school. The sample consists of students enrolled in school who ever switch between the public and private sector during primary school and excludes multiple switchers. The regressions control for child fixed effects, lagged test scores, whose effects are allowed to depend on the grade, and grade fixed effects, as well as female, age, age squared, year fixed effects, and their interactions with gender. The outcome is the mean of test scores in math, English and Urdu.

Figure 4: Event Study Graph for SVA Validation: Mean Test Scores

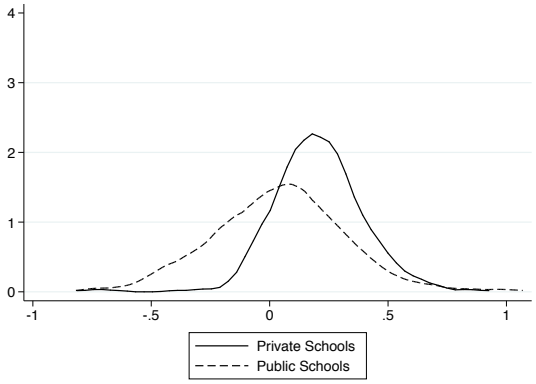


Notes: This graph reports estimates of the effect of the empirical Bayes SVA of the primary school a child switches into on her test scores t years after the switch for $t \in \{-2,..,2\}$. $t = 0$ is the year the switch occurs (denoted by a the red line). The sample consists of all students enrolled in school who ever switched schools during primary schools. It excludes multiple switchers. The regressions control for child fixed effects, lagged test scores, whose effects are allowed to depend on the grade, and grade fixed effects, as well as female, age, age squared, year fixed effects, and their interactions with gender. The outcome is the mean of test scores in math, English and Urdu.

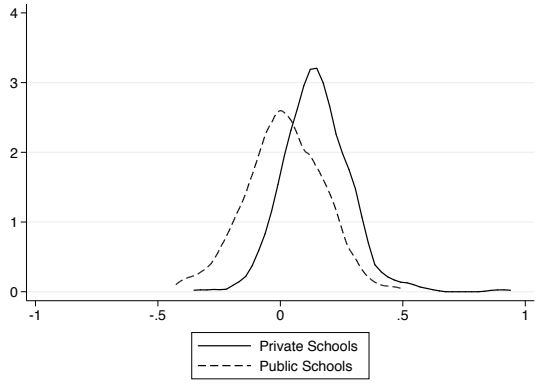
Figure 5: Distribution of Empirical Bayes SVA in the Public and Private Sectors



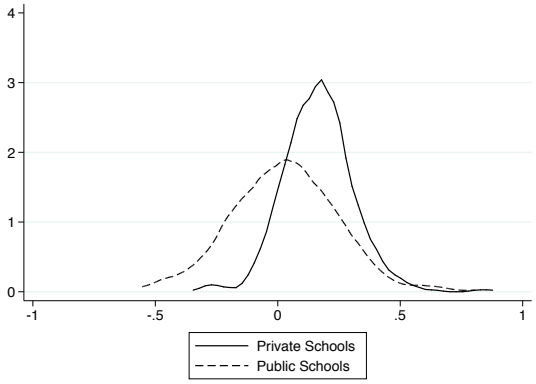
(a) Mathematics



(b) English



(c) Urdu



(d) Mean

Notes: These graphs report the distributions of the empirical Bayes SVA estimates for public and private schools separately.

Tables

Table 1: Learning Dynamics Over Time for the First Cohort of Tested Students

	Public Schools				Private Schools			
	(1) Year 1	(2) Year 2	(3) Year 3	(4) Year 4	(5) Year 1	(6) Year 2	(7) Year 3	(8) Year 4
Match picture with English word, Banana	0.518	0.648	0.773	0.823	0.824	0.897	0.938	0.944
Fill missing letter for picture, Cat	0.556	0.632	0.744	0.800	0.916	0.914	0.950	0.948
Fill missing letter for picture, Flag	0.182	0.197	0.358	0.470	0.508	0.521	0.722	0.728
Fill missing word in sentence	0.227	0.262	0.358	0.465	0.374	0.483	0.638	0.697
Construct sentence with word 'deep'	0.004	0.006	0.020	0.071	0.024	0.028	0.071	0.177
Construct sentence with word 'play'	0.006	0.010	0.052	0.154	0.065	0.070	0.219	0.357
Count number of moons, write number	0.563	0.618	0.749	0.724	0.693	0.740	0.852	0.783
Add 3 + 4	0.884	0.885	0.929	0.927	0.913	0.929	0.962	0.971
Multiply 4 x 5	0.534	0.551	0.686	0.760	0.690	0.755	0.868	0.885
Add 36 + 61	0.810	0.842	0.897	0.916	0.897	0.926	0.955	0.963
Add 678 + 923	0.477	0.505	0.647	0.700	0.666	0.732	0.826	0.827
Subtract 98 - 55	0.647	0.691	0.782	0.837	0.772	0.829	0.892	0.901
Multiply 32 x 4	0.448	0.466	0.622	0.700	0.620	0.712	0.839	0.852
Divide 384/6	0.172	0.183	0.368	0.464	0.224	0.345	0.603	0.661
Cost of necklace, simple algebra	0.075	0.119	0.211	0.241	0.144	0.192	0.341	0.335
Convert 7/3 into mixed fractions	0.020	0.032	0.052	0.101	0.011	0.063	0.091	0.209
Match picture with word, Book	0.675	0.749	0.876	0.920	0.803	0.876	0.958	0.966
Match picture with Urdu word, Banana	0.669	0.747	0.866	0.918	0.802	0.876	0.952	0.969
Match picture with word, House	0.464	0.510	0.612	0.698	0.633	0.724	0.830	0.859
Combine letters into word # 1	0.666	0.729	0.821	0.859	0.831	0.854	0.920	0.946
Combine letters into word # 2	0.286	0.350	0.450	0.542	0.518	0.598	0.700	0.731
Antonyms, Chouta	0.380	0.416	0.604	0.738	0.520	0.615	0.789	0.861
Antonyms, Khushk	0.321	0.401	0.548	0.627	0.420	0.584	0.749	0.775
Complete passage for grammar	0.248	0.296	0.476	0.618	0.369	0.511	0.678	0.765

Notes: This table reports summary statistics on learning over time on selected test items for the first cohort of students tested in LEAPS.

Table 2: Summary Statistics for the Tested Sample of Students

	All			Public Schools			Private Schools			Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SE	P-Value
Math Score	0.011	0.952	71467	-0.122	0.967	49917	0.327	0.822	19459	-0.449***	0.008	0.000
English Score	-0.056	0.952	71467	-0.282	0.932	49917	0.497	0.743	19459	-0.779***	0.007	0.000
Urdu Score	-0.012	0.969	71467	-0.177	0.970	49917	0.375	0.844	19459	-0.552***	0.008	0.000
Mean Score	-0.019	0.885	71467	-0.194	0.881	49917	0.400	0.735	19459	-0.593***	0.007	0.000
Change in Math	0.366	0.714	38336	0.370	0.721	27029	0.393	0.653	9936	-0.023***	0.008	0.005
Change in English	0.365	0.677	38336	0.381	0.693	27029	0.346	0.591	9936	0.034***	0.008	0.000
Change in Urdu	0.427	0.662	38336	0.433	0.677	27029	0.435	0.592	9936	-0.002	0.008	0.843
Change in Mean Score	0.386	0.554	38336	0.394	0.562	27029	0.391	0.495	9936	0.003	0.006	0.601
Female	0.447	0.497	71459	0.445	0.497	49910	0.443	0.497	19458	0.002	0.004	0.622
Age	10.453	1.827	71455	10.494	1.804	49917	10.228	1.791	19458	0.266***	0.015	0.000
Mom Some Education	0.392	0.488	53249	0.332	0.471	35146	0.527	0.499	16187	-0.195***	0.005	0.000
Dad Some Education	0.677	0.468	53247	0.626	0.484	35142	0.792	0.406	16189	-0.166***	0.004	0.000
Household Asset Index	0.021	1.727	53250	-0.255	1.708	35146	0.614	1.620	16188	-0.869***	0.016	0.000

Notes: This table reports summary statistics for all tested children in years 1-4 in grades 3 to 6. Since only a random sub-sample of tested students were surveyed in school, socioeconomic characteristics are only available for a subset of the observations. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 3: Summary Characteristics for the Combined Household and Tested Sample

	All			Public Schools			Private Schools			Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SE	P-Value
Math Score	-0.011	0.891	3383	-0.113	0.882	2419	0.288	0.820	738	-0.401***	0.036	0.000
English Score	-0.197	0.915	3383	-0.390	0.860	2419	0.397	0.738	738	-0.787***	0.035	0.000
Urdu Score	-0.066	0.913	3383	-0.211	0.884	2419	0.271	0.870	738	-0.483***	0.037	0.000
Mean Score	-0.091	0.825	3383	-0.238	0.793	2419	0.319	0.742	738	-0.557***	0.033	0.000
Change in Math	0.346	0.723	2001	0.354	0.691	1424	0.475	0.680	417	-0.121***	0.038	0.002
Change in English	0.314	0.712	2001	0.349	0.675	1424	0.346	0.663	417	0.003	0.037	0.940
Change in Urdu	0.402	0.667	2001	0.408	0.663	1424	0.456	0.627	417	-0.049	0.036	0.184
Change in Mean Score	0.354	0.567	2001	0.370	0.537	1424	0.426	0.544	417	-0.056*	0.030	0.064
Female	0.455	0.498	3382	0.448	0.497	2418	0.454	0.498	738	-0.006	0.021	0.788
Age	10.487	1.914	3351	10.427	1.671	2419	9.967	1.581	738	0.460***	0.069	0.000
Mom Some Education	0.289	0.453	3348	0.241	0.428	2394	0.463	0.499	730	-0.222***	0.019	0.000
Dad Some Education	0.649	0.477	2991	0.627	0.484	2160	0.733	0.443	636	-0.105***	0.021	0.000
Household Asset Index	-0.053	1.916	3383	-0.227	1.766	2419	0.504	2.275	738	-0.731***	0.080	0.000
Distance to Center	0.550	0.869	3383	0.582	0.861	2419	0.463	0.941	738	0.119***	0.037	0.001
Distance to Closest Private	0.629	0.814	3283	0.699	0.852	2334	0.404	0.652	735	0.295***	0.034	0.000
Distance to Closest Public	0.444	0.611	3371	0.462	0.627	2409	0.408	0.595	736	0.054**	0.026	0.040

Notes: This table reports summary statistics for all tested children in years 1-4 in grades 3 to 5 who also appear in the household survey sample. Students in grades 3 to 5 is the relevant sample for the effect of private schooling on test scores using the distance to primary school instrument. Primary schools ends in grade 5. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 4: Value-Added Estimates of the Effect of Private Schooling

	Math		English		Urdu		Mean	
	(1) Baseline	(2) SES	(3) Baseline	(4) SES	(5) Baseline	(6) SES	(7) Baseline	(8) SES
Private	0.153***	0.147***	0.229***	0.227***	0.159***	0.159***	0.129***	0.130***
	(0.023)	(0.022)	(0.023)	(0.022)	(0.021)	(0.020)	(0.020)	(0.019)
Adjusted R Squared	0.528	0.523	0.572	0.569	0.590	0.589	0.653	0.648
Number of Observations	37432	29394	37432	29394	37432	29394	37432	29394
Number of Clusters	969	968	969	968	969	968	969	968

Notes: This table reports value-added estimates of the effect of private schooling on the sample of tested students. All regressions include grade fixed effects, gender, lagged test scores interacted with grade level, and controls for age, age squared, year, and their interaction with gender. Even columns also include controls for whether the mother has some education, the father has some education, an index of household assets, and their interaction with gender. In odd columns, the sample consists of all tested children enrolled in school in years 1-4 in grades 3 to 6. In even columns, the sample consists of all students tested who were also surveyed about their socioeconomic background. Standard errors are clustered at the school level.* denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 5: OLS Estimates of the Effect of Private Schooling on Civic Values

	Full Index		Pakistan Knowledge		Government Index		Male Bias	
	(1) Baseline	(2) SES	(3) Baseline	(4) SES	(5) Baseline	(6) SES	(7) Baseline	(8) SES
Private	0.083***	0.077***	0.100***	0.093***	0.053***	0.056***	-0.047***	-0.046***
	(0.006)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)
Adjusted R Squared	0.238	0.227	0.249	0.236	0.124	0.119	0.098	0.097
Number of Observations	23959	17341	23959	17341	23959	17341	21332	15713
Number of Clusters	792	792	792	792	792	792	790	790

Notes: This table reports OLS estimates of the effect of private schooling on the sample of students with civic values test scores. All regressions include controls for grade fixed effects, gender, and controls for age, age squared, and their interaction with gender. Even columns also include controls for whether the mother has some education, the father has some education, an index on household assets, and their interaction with gender. In odd columns, the sample consists of all children with civic values scores in year 3 in grades 3 to 6. In even columns, the sample is restricted to students who were also surveyed about their socioeconomic background. Standard errors are clustered at the school level.* denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 6: Value-Added Estimates of the Effect of Private Schooling on Yearly Test-Score Gains Controlling for Child FE

	Math		English		Urdu		Mean	
	(1) Baseline	(2) SES	(3) Baseline	(4) SES	(5) Baseline	(6) SES	(7) Baseline	(8) SES
Private	0.112**	0.120***	0.191***	0.190***	0.126***	0.137***	0.148***	0.154***
	(0.051)	(0.046)	(0.042)	(0.042)	(0.046)	(0.044)	(0.044)	(0.042)
Adjusted R Squared	0.780	0.774	0.788	0.785	0.817	0.816	0.845	0.842
Number of Observations	37432	29395	37432	29395	37432	29395	37432	29395
Number of Clusters	969	968	969	968	969	968	969	968

Notes: This table reports value-added estimates of the effect of private schooling on the sample of tested students, controlling for child fixed effects. All regressions include grade fixed effects, gender, lagged test scores interacted with grade level, and controls for age, age squared, year and their interaction with gender. Even columns also include controls for an index on household assets, and its interaction with gender. In odd columns, the sample consists of all tested children enrolled in school in years 1-4 in grades 3 to 6. In even columns, the sample consists of all students tested who were also surveyed about their socioeconomic background. Standard errors are clustered at the school level.* denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 7: Effect of Private Schooling on Contemporaneous Test Scores With School Closure IV

	(1) First Stage	(2) Math	(3) English	(4) Urdu	(5) Mean
School Closure IV	-0.253***				
	(0.063)				
Private		0.531***	0.278*	0.533***	0.380**
		(0.188)	(0.164)	(0.195)	(0.157)
F-Statistic		133.46	126.91	133.10	129.59
Number of Observations	10695	10695	10695	10695	10695
Number of Clusters	603	603	603	603	603

Notes: This table reports IV estimates of the effect of private schooling on test scores. All regressions include controls for village FE, grade FE, grade FE interacted with lagged test scores, gender, age, age squared, year and their interaction with gender. The instrument is an indicator variable equal to 1 if a student attended a private school that has been closed. The sample consists of students enrolled in a private school the first year they are observed in the data. Standard errors are clustered at the school level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 8: Effect of Private Schooling on Civic Values With School Closure IV

	(1) First Stage	(2) Full Index	(3) Pakistan Knowledge	(4) Government Index	(5) Male Bias
School Closure IV	-0.314***				
	(0.076)				
Private		0.145***	0.148**	0.152**	-0.089
		(0.056)	(0.064)	(0.076)	(0.101)
F-Statistic		181.22	181.22	181.22	165.43
Number of Observations	7045	7045	7045	7045	6711
Number of Clusters	459	459	459	459	458

Notes: This table reports IV estimates of the effect of private schooling on civic values. All regressions include controls for village FE, grade FE, gender, age, age squared, and the interaction between the age controls and gender. The instrument is an indicator variable equal to 1 if a student attended a private school that has been closed. The sample consists of students enrolled in a private school the first year they are observed in the data. Standard errors are clustered at the school level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 9: Effect of Private Schooling on Contemporaneous Test Scores With Distance IV using Two Sample 2SLS

	(1) First Stage	(2) Math	(3) English	(4) Urdu	(5) Mean
Difference Distance IV	-0.312**				
	(0.150)				
Years in Private		0.013	0.304***	0.148*	0.155**
		(0.073)	(0.081)	(0.077)	(0.068)
F-statistic	10.34				
Number of Observations	5963				
Number of Obs. 1st Stage		5963	5963	5963	5963
Number of Obs. 2nd Stage		3102	3102	3102	3102

Notes: This table reports the two sample 2SLS results of the effect of private schooling on test scores. All regressions include controls for village FE, grade FE, gender, age, age squared, distance to the village center, year FE, and the interaction between the age controls, distance to the center, year, and gender. The instrument is the difference between the distance to the closest private and closest government schools. The first stage sample consists of children aged 6-13 in the household survey enrolled in primary school. The second stage sample consists of enrolled children who were both tested and appear in the household survey. Standard errors are estimated following Inoue and Solon (2010). * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 10: Effect of Private Schooling on Civic Values Using the Distance IV

	(1) First Stage	(2) Full Index	(3) Pakistan Knowledge	(4) Government Index	(5) Male Bias
Difference Distance IV	-0.315**				
	(0.150)				
Years in Private		0.027*	0.031*	0.015	-0.159***
		(0.015)	(0.018)	(0.023)	(0.030)
F-statistic	10.34				
Number of Observations	5963				
Number of Obs. 1st Stage		5963	5963	5963	5963
Number of Obs. 2nd Stage		1037	1037	1037	968

Notes: This table reports the two sample 2SLS estimates of the effect of private schooling on civics scores. All regressions include controls for village FE, grade FE, gender, age, age squared, distance to the village center, and the interaction between the age controls, distance to the center, and gender. The instrument is the difference between the distance to the closest private and closest government schools. The first stage sample consists of children aged 6-13 in the household survey enrolled in primary school. The second stage sample consists only of students who were both tested and appear in the household survey. Standard errors are estimated following Inoue and Solon (2010). * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 11: Comparison of Private School Effect Estimated with SVAs and Child Fixed Effects

	(1) Math	(2) English	(3) Urdu	(4) Mean
Child FE Estimates	0.157	0.201	0.124	0.168
SVA difference	0.177	0.174	0.140	0.164

Notes: This table compares the private school effect estimates from the child fixed effects estimation strategy (see Table 6) and from taking the difference between the SVAs of the public and private schools attended by switchers.

Table 12: Effect of a 1 Standard Deviation Better School

	(1) Public	(2) Private
Math	0.321	0.223
English	0.358	0.250
Urdu	0.269	0.153
Mean	0.316	0.208

Notes: This table reports the effect of attending a 1 standard deviation better private school or public school on test scores in math, English, Urdu and on mean test scores.

Table 13: Partial Equilibrium Effects of Different "Voucher Programs"

	Public to Best Private			Public to Worst Private		
	(1)	(2)	(3)	(4)	(5)	(6)
	p(10)	Mean	p(90)	p(10)	Mean	p(90)
Math	0.032	0.212	0.395	-0.074	0.066	0.199
English	0.016	0.278	0.570	-0.069	0.054	0.246
Urdu	0.036	0.209	0.327	-0.042	0.070	0.179
Mean	0.027	0.227	0.431	-0.037	0.072	0.188

Notes: This table uses the empirical Bayes SVA estimates to calculate the effect of moving all public school students to the best private school (columns 1-3) or the worst private school (columns 4-6) in their village. Columns 1 and 4 report the effect on students in the 10th percentile of test score gains. Columns 2 and 5 report the average effects, and columns 3 and 6 report the effects on students in the 90th percentile of test score gains.

Appendix A: Identifying Village Centers

To identify the population weighted center for each village, we first defined a two-dimensional space with the horizontal axes running from east to west and the vertical axes running from south to north. We then identified the north, south, east, and west boundaries of the village (the households that were located at the most extreme coordinates along each of these dimensions). Using our data on GPS coordinates, we divided the village into a grid with its bottom left corner at the combination of the most extreme south and east coordinates, its top left corner at the combination of the most extreme north and east coordinates, and so on. Each square in the grid was .002 decimal GPS coordinates by .002 decimal GPS coordinates. We then counted the number of households in each square and assigned a new weighted count to each square equal to the number of households in the square plus one-third times the number of households in each adjacent square. The center coordinate of the square with the highest weighted count was then determined to be the village centroid.

We do not simply use the centroid of the square with the highest unweighted count because there is a tradeoff in this algorithm between precision (the closeness of the approximation of the centroid using the center of the square to the "true village center") and the accuracy of the choice of the highest count square. A very small square will give higher "precision" but could lead the estimate to be easily biased by very small dense settlements far from most of the village or even by randomly occurring density generated by the random sampling design. To compromise between precision and accuracy, we instead use this weighted count.

Appendix B: Empirical Bayes Estimates of SVA

Let

$$y_{ijst} = \beta X_{ijt} + \theta_s + \theta_j + \theta_{jt} + \varepsilon_{ijt}, \tag{7}$$

where y_{ijst} is the test score, X_{ijt} is the set of controls, θ_s is the school effect (not including the teacher shock), θ_j is the teacher effect, θ_{jt} is the classroom effect, and ε_{ijt} is an idiosyncratic student-specific shock. The variances of these shocks are σ_S^2, σ_T^2, σ_C^2, and σ_ε^2 respectively, and they are assumed to be independent and homoskedastic.

Our object of interest is the expected test score gains a child will experience in a school:

$$\delta_s = \theta_s + \sum_{j \in s} \frac{N_j}{N_s} \theta_j, \tag{8}$$

where N_j is the number of students taught by teacher j and N_s is the number of students in school s. Note that this is just the independent school effect plus the weighted average of the teacher effects of the teachers who teach in a school. To calculate $Var(\delta_s)$, use the fact that $Var(\delta_s) = E(\delta_s^2) - E(\delta_s)^2$. Noting that $E(\delta_s)^2 = 0$ by construction, the variance of δ is

$$Var(\delta_s) = E((\theta_s + \sum_{j \in s} \frac{N_j}{N_s} \theta_j)^2).$$

Recognizing that θ_j and θ_s are independent by assumption, this can be further simplified to

$$\begin{aligned} Var(\delta_s) &= E(\theta_s^2) + E(\sum_{j \in s} \sum_{j' \in s} \frac{N_j N_{j'}}{N_s^2} \theta_j \theta_{j'}). \\ &= \sigma_S^2 + E(\frac{\sum_j N_j^2}{N_s^2} \sigma_T^2). \end{aligned} \tag{9}$$

Our estimate of δ_s (the school fixed effect) is given by

$$\hat{\delta}_s = \theta_s + \frac{1}{N_s} \sum_{ijt \in s} \left(\theta_j + \theta_{jt} + \varepsilon_{ijt}\right) \tag{10}$$

Then, the variance of $\hat{\delta}_s$ is

$$
\begin{aligned}
Var(\hat{\delta}_s) &= E\Big((\theta_s + \frac{1}{N_s}\sum_{ijt\in s}\big(\theta_j + \theta_{jt} + \varepsilon_{ijt}\big))^2\Big) \\
&= \sigma_S^2 + E\Big(\sum_{j\in s}\sum_{k\in s}\frac{N_j N_k}{N_s^2}\theta_j\theta_k + \sum_{jt\in s}\sum_{kl\in s}\frac{N_{jt}N_{kl}}{N_s^2}\theta_{jt}\theta_{kl} + \sum_{ijt\in s}\sum_{i'j't'\in s}\frac{1}{N_s^2}\varepsilon_{ijt}\varepsilon_{i'j't'}\Big) \\
&= \sigma_S^2 + E\Big(\sum_j \frac{N_j^2}{N_s^2}\sigma_T^2 + \sum_{jt}\frac{N_C^2}{N_s^2}\sigma_C^2 + \frac{1}{N_s}\sigma_\varepsilon^2\Big), \qquad (11)
\end{aligned}
$$

Therefore, the variance of the school effects is

$$
Var(\delta_s) = Var(\hat{\delta}_s) - E\Big(\frac{\sum_{jt}N_{jt}^2}{N_s^2}\sigma_C^2 + \frac{1}{N_s}\sigma_\varepsilon^2\Big). \qquad (12)
$$

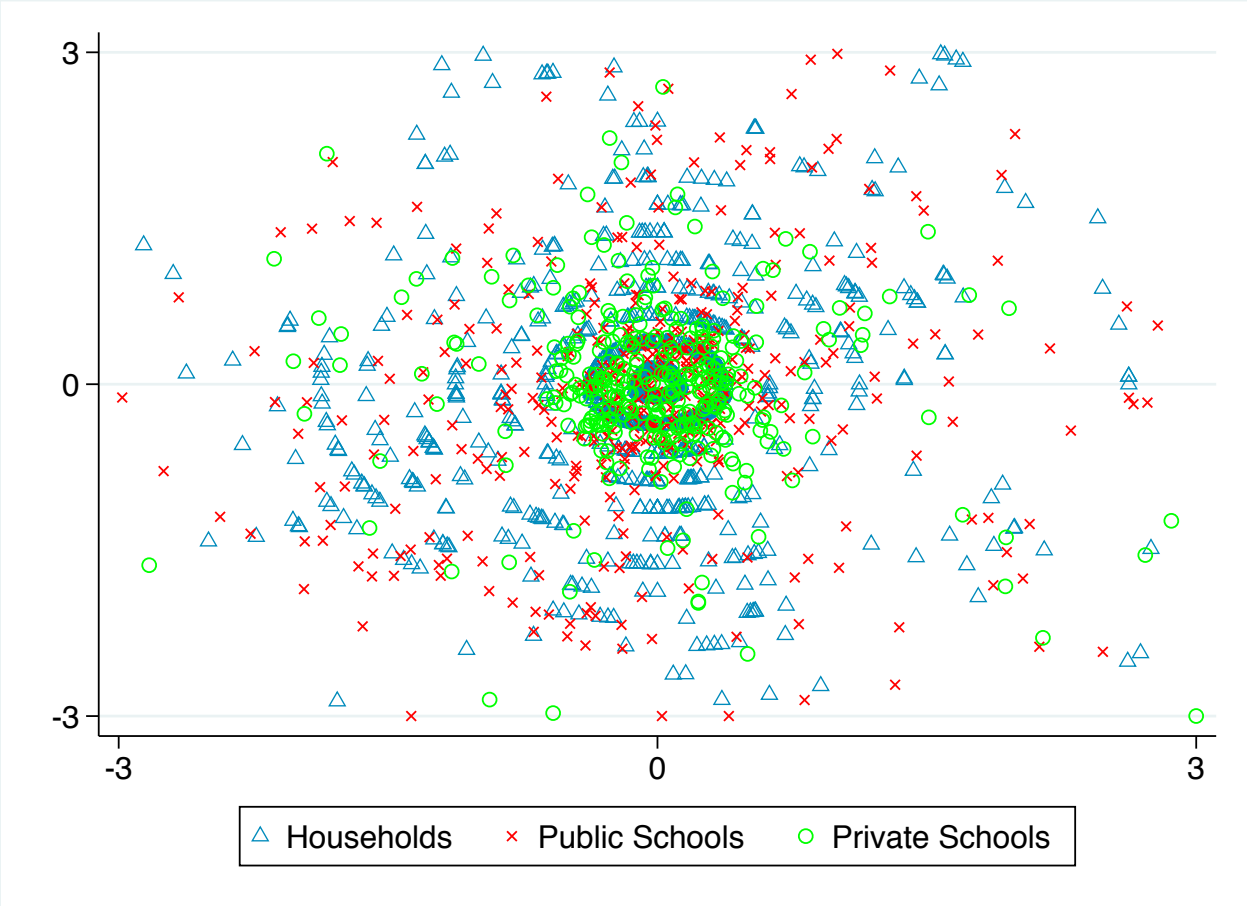
For empirical Bayes, we should then scale $\hat{\delta}_s$ by

$$
h_s = \frac{\sigma_S^2 + \frac{\sum_j N_j^2}{N_s^2}\sigma_T^2}{\sigma_S^2 + \sum_j \frac{N_j^2}{N_s^2}\sigma_T^2 + \sum_C \frac{N_{jt}^2}{N_s^2}\sigma_C^2 + \frac{1}{N_s}\sigma_\varepsilon^2} \qquad (13)
$$

Note that σ_s^2, σ_{jt}^2, σ_j^2 and σ_ε^2 are all calculated in Bau and Das (forthcoming) separately for private and public schools in the same data, so we can substitute these values into equation (12) to get the variances of school quality in the public and private sectors and in (13) to get the scaling value for calculating the empirical Bayes estimates of SVA.

Appendix Figures

Figure A1: The Global Village



Notes: The global village normalizes all villages to have a center at the coordinates (0,0). The distances are in terms of kilometers. Households are placed on the closest ring radiating outwards from the global village center, with rings spaced at 0.25 km to avoid too much direct overlap with school locations.

Appendix Tables

Table A1: Components of Full Civics Index

	(1) Public Schools	(2) Private Schools
Pakistan Knowledge		
What is a neighboring country of Pakistan?	0.334	0.412
What is the largest province by area?	0.282	0.348
Which city has the largest population?	0.472	0.599
Who is the founder of Pakistan?	0.815	0.922
Who is the prime minister?	0.442	0.576
Who gave independence?	0.432	0.451
Where was the earthquake?	0.639	0.782
Finish the pop song	0.497	0.623
Government Index		
Finish the poem	0.248	0.372
Finish the national slogan	0.147	0.201
Would give money to government or army	0.321	0.329
Vote to choose lunch	0.140	0.158
Male Bias		
Boys are better at studies	0.193	0.143
Boys are better at monitoring	0.263	0.245
Additional Question		
A good scientist observes better	0.266	0.247

Notes: This table reports summary statistics on civic scores in round 3. All items are included in the full civics index. The Male Bias questions are recoded so a higher score is "better" when included in the full index.

Table A2: Correlation Between School Closure Instrument and Students' Characteristics

	(1) Mom Education	(2) Dad Education	(3) Household Assets	(4) Child High Ability	(5) Female
School Closure	-0.040	-0.057*	-0.081	-0.054*	-0.041
	(0.031)	(0.031)	(0.138)	(0.028)	(0.032)
Adjusted R Squared	0.10	0.07	0.09	0.04	0.05
Number of Observations	16901	16903	16902	14851	20275
Number of Clusters	633	633	633	632	640

Notes: This table reports the coefficients from regressions of student characteristics in the survey conducted in schools on the school closure instrument. The regressions include controls for village FE, grade FE, gender, age, age squared, and the interaction between the age controls and gender. The sample consists of tested students who were in private school when they were first observed. Standard errors are clustered at the school level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A3: Correlation Between Distance IV and Household Characteristics

	(1) Mom Edu.	(2) Dad Edu.	(3) HH Assets	(4) Log Expend.	(5) Print Media	(6) Land Area	(7) Enrolled in School	(8) High Ability	(9) Elder Sisters	(10) Elder Brothers
Difference Distance IV	-0.030	-0.050	-0.106	-0.001	-0.027	-0.322	-0.010	-0.013	0.142	0.031
	(0.033)	(0.033)	(0.178)	(0.052)	(0.019)	(1.284)	(0.008)	(0.020)	(0.091)	(0.094)
Adjusted R Squared	0.19	0.15	0.13	0.15	0.07	0.09	0.43	0.04	0.09	0.07
Number of Observations	3214	2884	3246	2374	3246	2371	3246	3209	2063	2063
Number of Clusters	111	111	111	111	111	111	111	111	111	111

Notes: This table reports the coefficients from regressions of student characteristics in the household survey on the relative distance instrument. The instrument is the difference between the distance to the closest private and closest public schools. The regressions include the same controls as the distance IV specifications: village FE, grade FE, gender, age, age squared, distance to the village center, and the interaction between the age controls, distance to the center, and gender. Standard errors are clustered at the village level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A4: Effect of Private Schooling on Contemporaneous Test Scores With Distance IV

	(1) First Stage	(2) Math	(3) English	(4) Urdu	(5) Mean
Difference Distance IV	-0.458***				
	(0.119)				
Years in Private		-0.016	0.205**	0.065	0.085
		(0.109)	(0.099)	(0.103)	(0.094)
F-Statistic		47.39	47.39	47.39	47.39
Number of Observations	2980	2980	2980	2980	2980
Number of Clusters	111	111	111	111	111

Notes: This table reports IV estimates of the effect of private schooling on test scores using the distance IV. The instrument is the difference between the distance to the closest private and closest public schools. All regressions include controls for village FE, grade FE, gender, age, age squared, distance to the village center, year FE, and the interaction between the age controls, distance to the center, year, and gender. The sample consists of primary school students who were both tested and appear in the household survey. Standard errors are clustered at the village level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A5: Effect of Private Schooling on Civic Values Using the Distance IV

	(1) First Stage	(2) Full Index	(3) Pakistan Knowledge	(4) Government Index	(5) Male Bias
Difference Distance IV	-0.605***				
	(0.207)				
Years in Private		0.006	0.007	-0.002	-0.076*
		(0.024)	(0.029)	(0.032)	(0.042)
F-Statistic		17.85	17.85	17.85	16.51
Number of Observations	973	973	973	973	914
Number of Clusters	108	108	108	108	108

Notes: This table reports IV estimates of the effect of private schooling on civics scores using the distance IV. The instrument is the difference between the distance to the closest private and closest public schools. All regressions include controls for village FE, grade FE, gender, age, age squared, distance to the village center, and the interaction between the age controls, distance to the center, and gender. The sample consists of only primary school students who were both tested and appear in the household survey. Standard errors are clustered at the village level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A6: Do School Facilities Predict SVA?

	Public Schools		Private Schools	
	(1)	(2)	(3)	(4)
	Mean SVA	Mean SVA	Mean SVA	Mean SVA
Library	-0.094**	-0.110**	0.085*	0.109**
	(0.043)	(0.049)	(0.045)	(0.052)
Computer	0.067	-0.076	0.052	0.052
	(0.130)	(0.167)	(0.048)	(0.053)
Sports	0.027	0.021	0.079	0.074
	(0.055)	(0.061)	(0.052)	(0.062)
Hall	-0.043	-0.029	-0.128*	-0.139
	(0.089)	(0.107)	(0.074)	(0.091)
Wall	0.001	0.002	-0.111	-0.222*
	(0.036)	(0.037)	(0.103)	(0.114)
Fans	0.051	0.054	0.138*	0.087
	(0.055)	(0.063)	(0.081)	(0.133)
Electricity	-0.043	-0.066	-0.071	0.041
	(0.053)	(0.065)	(0.092)	(0.107)
Student-Teacher Ratio	-0.001	-0.001	0.003	0.003
	(0.001)	(0.001)	(0.002)	(0.002)
Number of Private Schools	0.008		-0.000	
	(0.006)		(0.003)	
Number of Public Schools	-0.003		0.010**	
	(0.006)		(0.004)	
Log Number of Children	0.003		-0.014	
	(0.018)		(0.015)	
Fixed Effects	District	Village	District	Village
Adjusted R Squared	0.33	0.45	0.21	0.46
Within Adj. R Squared	0.01	0.02	0.07	0.08
Number of Observations	1881	1881	1160	1160
Number of Clusters	112	112	108	108

Notes: This table regresses the fixed effect estimates of schools' mean SVA's on school facilities measures. Facility measures are the means across all four years to account for variation in facilities over time. The within R^2 reports the R^2 not including the contribution of the fixed effects. Standard errors are clustered at the village level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A7: Correlation Between Students' Characteristics and Mean SVA

	Dependent Variable: Mean SVA							
	Public Schools				Private Schools			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male		Female		Male		Female	
Assets	0.005***	0.001	0.002	-0.001	0.004**	0.006***	0.008***	0.009***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Mom Education	0.005	0.001	0.025***	0.004	0.004	0.001	0.006	0.001
	(0.007)	(0.005)	(0.006)	(0.004)	(0.007)	(0.005)	(0.009)	(0.006)
Dad Education	0.004	0.006	0.002	0.006	0.020**	0.016**	0.032***	0.017**
	(0.006)	(0.004)	(0.006)	(0.004)	(0.009)	(0.007)	(0.011)	(0.008)
Fixed Effects	District	Village	District	Village	District	Village	District	Village
Adjusted R Squared	0.45	0.71	0.22	0.67	0.21	0.55	0.18	0.56
Number of Observations	18439	18439	15432	15432	8777	8776	7234	7234
Number of Clusters	8105	8105	6582	6582	4150	4149	3422	3422

Notes: This table regresses the fixed effect estimates of schools' mean SVA's on student characteristics measures. An observation is at the student-year level. Standard errors are clustered at the student level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.