FOR ONLINE PUBLICATION

APPENDIX FOR "HETEROGENEITY IN SCHOOL VALUE-ADDED AND THE PRIVATE PREMIUM"

APPENDIX A: SUBJECT AND CIVICS TEST IN LEAPS

Test scores are based on exams in English, Urdu, and mathematics. 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.

Performance on items of the cognitive tests over time is detailed in Appendix Table A1. 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, 17% (22%) of public (private) school students could answer 384/6 correctly in round 1 (2003), which increased to 48% (66%) in round 4 (2007).

The civics test was divided into questions designed to elicit civic knowledge and civic dispositions. Appendix Table A2 reports all the civics questions by the index to which they were assigned, as well as the fraction of students who correctly answered them. In the civic knowledge section, we ask about the political structure of the state and its history, the 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. 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 civic disposition index, and (4) a gender bias index.

Students appear to have a poor grasp of civic knowledge (for example, 33% of students in public schools and 41% in private schools knew that India neighbors Pakistan, with U.S., Saudi Arabia, and Kuwait as other choices). For comparison, in the U.S., the 1998 NAEP results show that 45% of fourth graders knew that both citizens and non-citizens are legally protected by US laws, and 43% knew the president's role in making laws is to sign bills passed by Congress into laws (Johnson and Vanneman, 2001).

Pakistani students also mistrust the government and voting as a choice mechanism: 68% preferred to donate money in the case of disasters to private entities or nonprofits rather than the government, and 14% thought that voting was the best way to decide what to eat for lunch relative to handing the decision over to a central authority.

APPENDIX B: EMPIRICAL BAYES ESTIMATES OF SVA AND CORRECTED VARIANCE MEASURES

Let

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

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:

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

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

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

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

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

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

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

Therefore, the variance of the school effects uncontaminated by estimation error is

$$\text{(B1)} \qquad Var(\delta_s) = Var(\hat{\delta}_s) - E\Big(\frac{\sum_{jt} N_{jt}^2}{N_s^2}\sigma_C^2 + \frac{1}{N_s}\sigma_\varepsilon^2\Big).$$

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

$$\text{(B2)} \qquad h_s = \frac{\sigma_S^2 + \frac{\sum_j N_j^2}{N_s^2}\sigma_T^2}{\sigma_S^2 + \sum_j \frac{N_j^2}{N_s^2}\sigma_T^2 + \sum_C \frac{N_{jt}^2}{N_s^2}\sigma_C^2 + \frac{1}{N_s}\sigma_\varepsilon^2}$$

Note that σ_s^2, σ_{jt}^2, σ_j^2 and σ_ε^2 are all calculated in Bau and Das (2020) separately for private and public schools in the same data, so we can substitute these values into equation (B1) to get the variances of school quality in the public and private sectors and into (B2) to get the scaling value for calculating the empirical Bayes estimates of SVA.

Appendix C: Additional Validations of SVA

Validating the SVA Estimates with AKM

Relationship Between AKM and Standard SVA Estimates. As an additional validation of the SVA estimates, we re-estimate SVA with an AKM-style regression (Abowd, Kramarz and Margolis, 1999; Card, Heining and Kline, 2013), which exploits children who switch schools to control for time-invariant child-level heterogeneity in test score growth. We add a control for child fixed effects to equation (1), so that the estimating equation is:

$$\text{(C1)} \qquad y_{igst} = \lambda_g y_{igs,t-1} + \theta_s^{AKM} + \phi_i + \alpha_g + \alpha_t + \varepsilon_{igst},$$

where ϕ_i is a child fixed effect. To focus on estimating primary school effects (which is what our main SVA estimates identify), we exclude switches between schools that occur due to entering 6th grade from the estimating sample. Otherwise, since switchers are key for identification, and these are by far the most common type of switches, our school effect estimates would mainly capture secondary school quality (which may differ from primary school quality even within the same school).

Importantly, equation (C1) only identifies SVAs within a set of schools that are connected by student switches. Because students do not switch schools across villages, the maximum size of a connected set is all the schools in the village. As a result, these estimates do not allow for cross-village comparisons in school quality. Nonetheless, they are useful as a robustness check since, if the identifying assumptions of equation (1) are satisfied, the original SVA estimates should predict the AKM estimates with a coefficient of 1 within connected sets.

In Appendix Figure A8, we report the binscatter plot of the AKM-style estimates of school effects against the empirical Bayes SVA estimates (controlling for connected set fixed effects). The figure indeed shows that the estimates are highly correlated. The regression coefficient is 0.79 (se=0.16), and an F-test cannot reject that the coefficient is equal to 1.

Validity of AKM Estimates. We also test whether the assumptions of equation (C1) are satisfied. One potential concern is that students differentially sort into schools with which they have an especially good match, and this could bias our school effects estimates. To assess the degree to which this is an issue, we use two tests from Card, Heining and Kline (2013). First, to evaluate whether there are important match effects that may bias our school effects estimates, we compare the R^2 and the RMSE of the equation above to those from a fully saturated model that includes child-by-school fixed effects. The R^2 and RMSE are identical up to 3 decimal points (R^2=0.922 and RMSE=0.311). Second, we plot the mean residuals from the AKM-style estimating equation by deciles of school and child effects (normalizing the school effects to be mean 0 within a set and dropping schools that are in singleton sets) in Appendix Figure A9. As the figure shows, there are no systematic patterns in the residuals. The smallest value for the mean residuals is -0.004, and the largest is 0.008.

Another concern is that students who switch into different quality schools have different trends in their test scores. To assess whether this is an issue, we replicate the symmetric mobility tests in Card, Heining and Kline (2013) using the AKM-style school effects estimates in Appendix Figure A10 to ensure that students who switch from good to bad vs. bad to good schools are on similar trends. After normalizing the AKM estimates of SVA within connected sets, we divide schools into terciles and plot mean test scores for students who switch between bottom and top tercile schools (in both directions) and for those who switch but remain in bottom/top tercile schools. We note that this is a conservative test of our estimating model because it doesn't condition on lagged test scores. The results in Appendix Figure A10 are consistent with symmetric mobility. Across all groups, test scores grow by approximately the same amount between $t = -2$ and $t = -1$. Students who switch into (roughly) same quality schools do not experience changes in their test score trajectory. Students who switch from bottom to top-tercile schools experience gains, while students who switch from top to bottom quality schools experience declines.

Multi-dimensionality & Civic Values

In this section, we explore the relationship between test score SVA and another important outcome in this context – civic values. While our results suggest that the SVA measures are unbiased, there is an important concern that SVA measures based on test scores capture only part of what schools are supposed to provide. If this is the case, even if a school's SVA is an unbiased predictor of a student's test score gains within the school, a focus on SVA may lead us to miss important variation in school quality. While this concern is important for any unmeasured component of school quality, it is particularly relevant for civics knowledge and disposition in the context of private schools precisely because these subjects are thought to be non-contractible and to have large positive externalities. The importance of civics, particularly in weaker states like Pakistan, is one of the main rationales for support of public education.[1]

Our results show that high SVA schools, defined as schools where children gain more in the tested subjects, are also schools where children have higher levels of civic values and knowledge, both of which are excluded from the original SVA computation.[2] We report the relationship between schools' SVA and students' civic values scores in Appendix Figure A11 and Appendix Table A10, noting that we do not have value-added measures for civics as the test was only administered in one year. We aggregate performance into average scores on three indices designed to capture different civic skills (Pakistan Knowledge, Government Disposition, and Gender Bias) in addition to a "Full Index" that includes all the questions.[3] A higher score on all indices is better with the exception of gender bias, where a higher score indicates greater gender bias. For all the indices (Appendix Figure A11) and within both the public and private sectors (Appendix Table A10), we observe a strong association between SVA and civic values. Schools that are better at producing cognitive skills are also better at producing civic values. Moreover, the correlations are strong for questions around civic disposition (preference for democratic processes, trust in the government), which arguably are not simply affected by a school being better at teaching the curriculum (including civic knowledge) in general. This suggests that any trade-off between a focus on cognitive skills and civics measures, if present at all, will be small.[4]

Appendix D: Details of Identification Strategies for Estimating the homogeneous private school premium

Value-Added Strategies

Our first strategy uses the value-added approach to calculate the private school premium. In our most parsimonious specification, we include controls for gender, age, age squared, and the interaction between gender and the age control and year fixed effects. We then layer on additional controls for socioeconomic status (mother and father education, and time-varying household assets, as well as all of their interactions with child gender). Finally, to account for unobserved but time-invariant child characteristics that are correlated with test score improvements, we exploit variation due to students switching schools by controlling for child fixed effects. In this specification, β_1 is identified by comparing the change in test scores over time for children who switch into or out of private schools to those who do not.

[1]In most post-colonial countries, nation-building was one of the key aims of the public schooling system (see Cohn and Scott (1996) on India and Bassey (1999) on sub-Saharan Africa). For Pakistan, 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, which has explicitly called for training in citizenship. In the influential first education conferences in 1947, the Minister of Education stated that, "*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*" (Dean (2005), page 36).

[2]The focus on civics is motivated by a literature that micro-founds the provision of state schooling in terms of the non-contractibility of civic values, starting from Meyer et al. (1979) and with recent contributions, for instance, by Pritchett (2013) and Bandiera et al. (2019). The scores are from a specially designed test of civics administered once in 2005-2006 and therefore represent a cross-section rather than value-added.

[3]Appendix Table A2 reports the components of each index.

[4]The base civics curricula that all schools follow is the same, with textbooks vetted and prescribed by a textbook board. According to a principle of 'additionality,' conditional on satisfying that base, schools may choose to add on additional subjects or allocate more/less time to a given subject.

The regression specification is

$$y_{igst} = \beta_0 + \lambda_g y_{igs,t-1} + \beta_1 private_{ist} + \alpha_g + \alpha_t + \mathbf{\Gamma X_{igst}} + \varepsilon_{igst}, \quad \text{(D1)}$$

where $private_{ist}$ is an indicator variable equal to 1 if student i attends a private school in year t, and $\mathbf{X_{igst}}$ is a vector of controls. In the most conservative specification, these controls include child fixed effects. As before, λ_g, the coefficient on the lagged test score, is allowed to vary at the grade-level. Then, β_1 estimates the effect of a year of private schooling. There is a natural connection between β_1 and our SVA estimates, as β_1 should approximate the difference in the means of the student-weighted SVA distributions between the public and private sectors.[5]

Even after controlling for child fixed effects and lagged test scores, β_1 may still be biased if students whose test scores are on a better trajectory switch into private schools. We assess whether this is a likely source of bias by plotting an event study graph for switching to a private school in Appendix Figure A13. Reassuringly, there is no evidence that students who switch into private schools are on different trends than those who do not, and following a switch, students' test scores appear to rise by exactly the estimated private premium.

School Closure IV

The value-added estimate of the private premium β_1 may still be biased if school-switching is associated with time-varying shocks that also affect test scores. To address this concern, similar to our validation strategy in Section III.B, we exploit exogenous 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

$$private_{ivst} = \beta_0 + \lambda_g y_{igs,t-1} + \mu_1 closure_{it} + \mathbf{\Gamma X_{igst}} + \alpha_g + \alpha_t + \alpha_v + \varepsilon_{igst}. \quad \text{(D2)}$$

The second stage is equation (D1) except that it now controls for village fixed effects. Recall that we already assessed whether $closure_{it}$ is associated with children's characteristics in Appendix Table A3.

Column 1 in Appendix Table A15 shows the first stage of our closure instrument: a closure reduces the subsequent probability that a student attends a private school by a statistically significant 25 percentage points. The remainder of the table reports the (highly positive) estimates of the private premium for math, English, Urdu, and mean test scores.

We also use the closure IV strategy to estimate the 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 then be interpreted not as the magnitude of the effect of one year of private

[5]This would be mechanically true if we included the exact same controls across specifications and imposed that the coefficients on controls were identical in equation (D1) to those estimated by equation (1). In practice, however, we include more stringent controls in the private premium specification. Furthermore, even if the controls were the same, the inclusion of school fixed effects in equation (1) and not in equation (D1) may lead to differences in these coefficient estimates.

schooling but as the net effect of attending private school across multiple years. Appendix Table A16 reports the instrumental variables estimates for the civics outcomes. There is again no evidence that private schooling reduces civic values and some evidence that it significantly improves them.

Distance Instrument

Our next strategy exploits the fact that Pakistani households are extremely distance sensitive in their choice of schools. The recognized challenge with a distance-based instrument is that households with a greater demand for quality education may choose to locate closer to private schools (and private schools may endogenously locate to be closer to these households). To address this concern, our identification strategy exploits two historical developments that led to villages where the rich are located in the center, but the public schools are located on the peripheries.

First, settlement patterns arising from 1880 onward resulted in richer households locating in village centers. In his discussion of settlement patterns in Punjab, Paustian (1930) details how the British built water canals in one of the world's largest irrigation projects and leased land in order to settle previously uninhabited regions. Planned villages were built 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.[6] 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 Faisalabad and the majority in Rahim Yar Khan (the districts in our study that are in the center and the South).

Second, many public schools were constructed in the 1980s and 1990s through programs that required the village to provide land for the school. This land often came from village common property and was thus easier to donate compared to the private land in the center of the village, which would have to be purchased. As a result, a significant fraction of public schools are 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.

These settlement and school location patterns suggest that using distance to the closest private school as an instrument will be problematic as richer households are likely to live closer to a private school. Instead, we use the relative distance to the closest private versus closest public school. Additionally, we calculate and directly control for distance to the center, as well as village fixed effects, in all our specifications. Consequently, the variation

[6]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.

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.

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 Figure A12, 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.

Finally, 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 score 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, we use the two sample 2SLS methodology of Inoue and Solon (2010), which allows us to estimate the relationship between the instrument and the endogenous variable using the full sample of children for whom we observe distance 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. Summary statistics for the sample of 1,269 students for whom both distance and test score data are available are reported in Appendix Table A11 with observations at the child-year level.

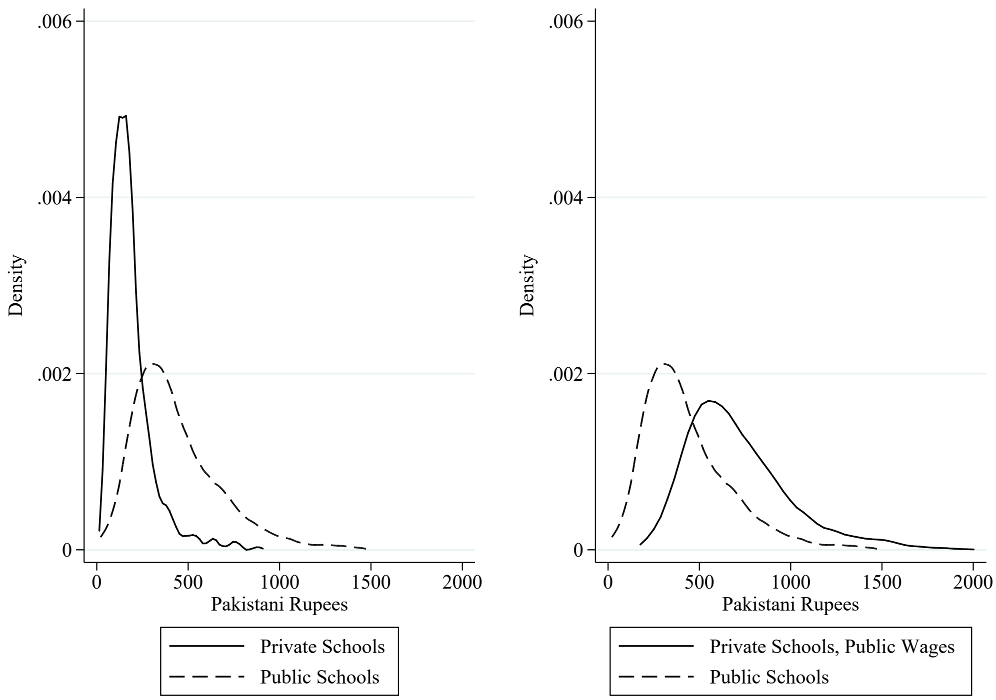
The first stage is

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

where v indexes a village, $years\,private_{ist}$ 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, $female_i$ is an indicator variable equal to 1 if the child is female, and α_v is a village fixed effect. The controls $\mathbf{X_{igst}}$ are the same as in equation (D1). We focus on $years\,private_{ist}$ instead of private schooling as our variable of interest for two reasons. First, the instrument varies little over time. Second, we cannot control for lagged test scores in the two-sample IV strategy since we do not observe test scores for non-tested children, and we use the sample of non-tested children to estimate the first-stage. 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 $years\,private_{igst}$.

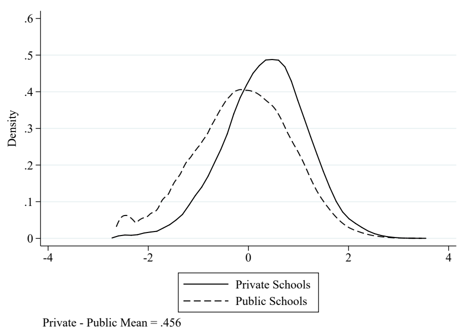
Appendix Table A12 confirms that, conditional on controlling for distance to the center, relative distance to a private school is indeed uncorrelated with parental education, assets, consumption, or family size. In a regression of the instrument on the individual characteristics in the table, the F-statistic from a joint test of those characteristics is 0.66 (p=0.76). Interestingly, enrollment in any school is not correlated with our instrument. This suggests both that a very important marker of demand is uncorrelated with our modified distance instrument and that we need not worry that the instrument induces additional selection into the sample of test-taking students on the extensive margin. Finally, the first column in Appendix Table A13 reports the first stage, showing 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 remaining columns report estimates of the private premium, which are sizable and positive for English, Urdu, and mean test scores.

Figure A1. : Monthly School Expenditures per Student in Public and Private Schools

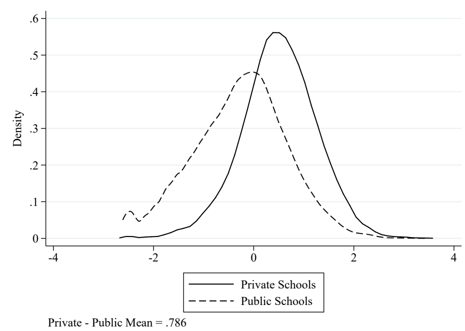


Notes: The left panel shows the distribution of public and private schools' total costs per student in the data. The right panel shows public and private schools' total costs per student if private schools were to pay their teachers at the reported village average public school teacher wage. Total expenditures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

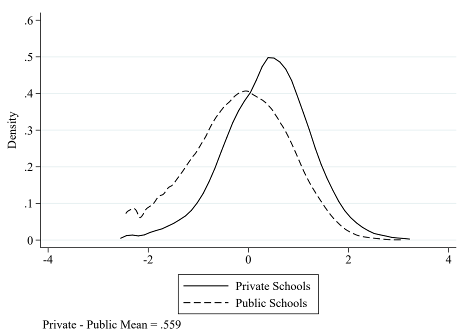
Figure A2. : Test Scores in Public and Private Schools



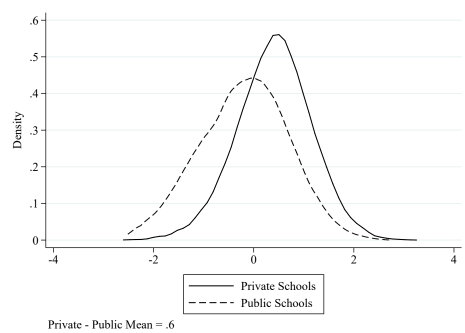
(a) Mathematics



(b) English



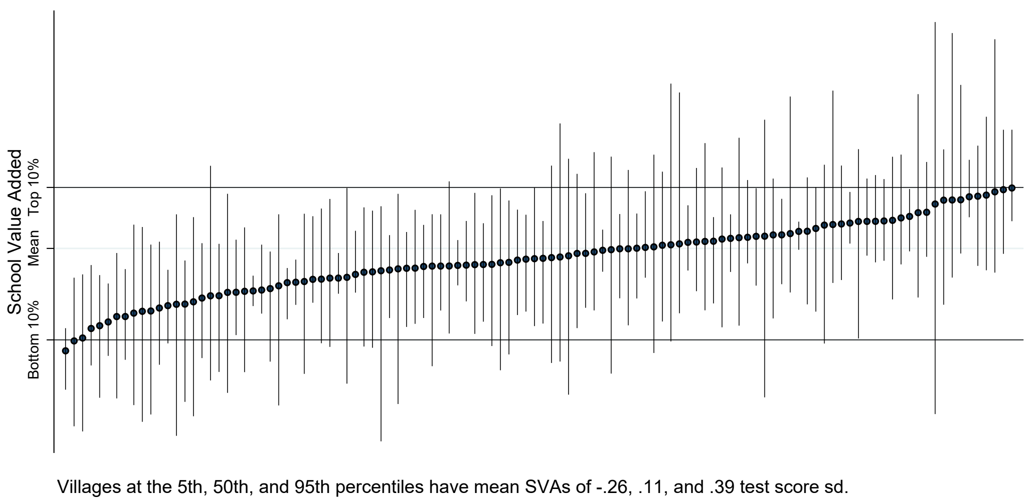
(c) Urdu



(d) Mean

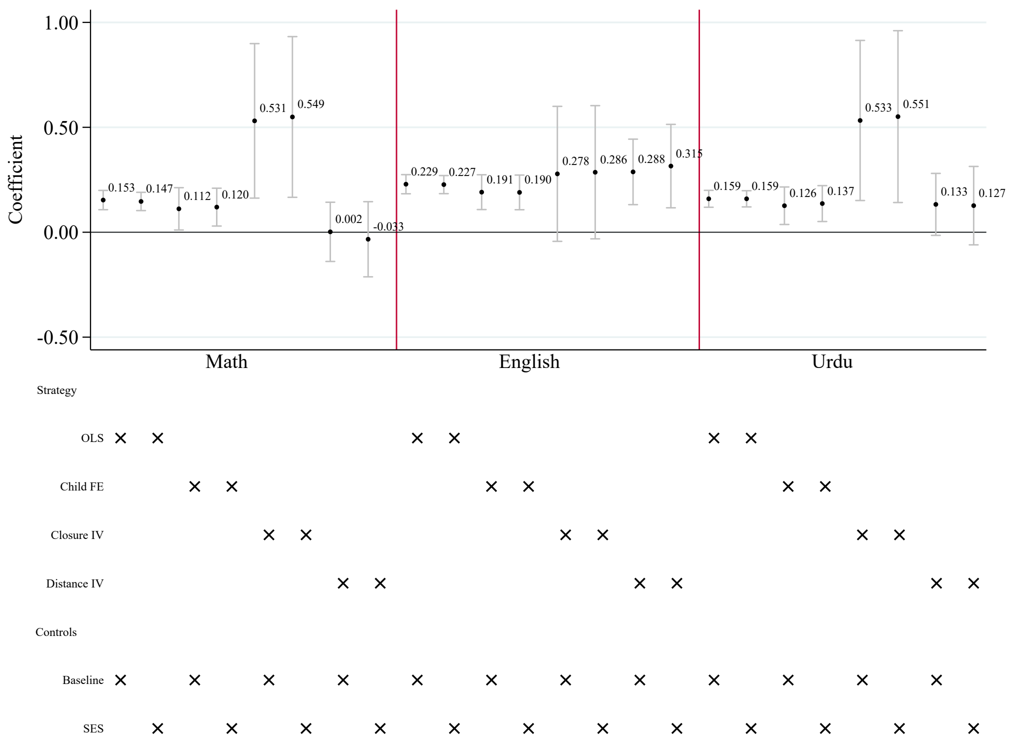
Notes: This figure plots the distribution of test scores for students enrolled in private and public schools, respectively. The mean test score is the average over test scores in mathematics, Urdu, and English.

Figure A3. : Range of SVA Estimates Within Each Village



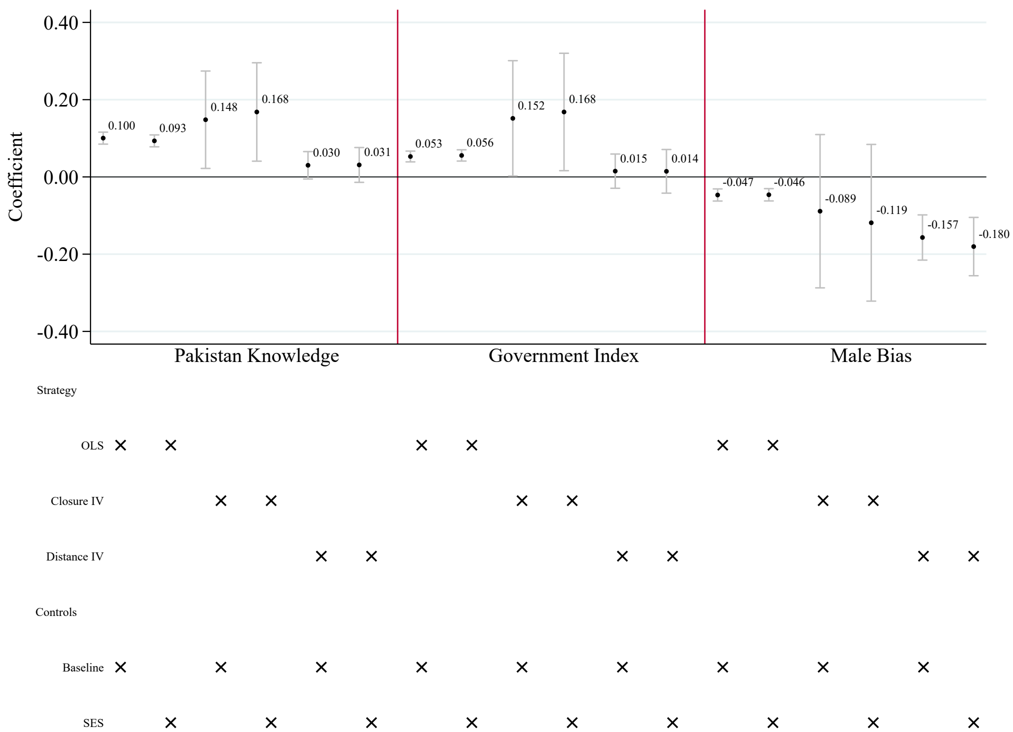
Notes: This graph plots the range of SVA estimates by village. Dots denote the average SVA in a village, villages are sorted in order of the average (non-population weighted) SVA, and the vertical lines denote the range between the minimum and maximum values of SVA in the village. The SVA are empirical Bayes estimates.

Figure A4. : Private Premium Estimates by Subject



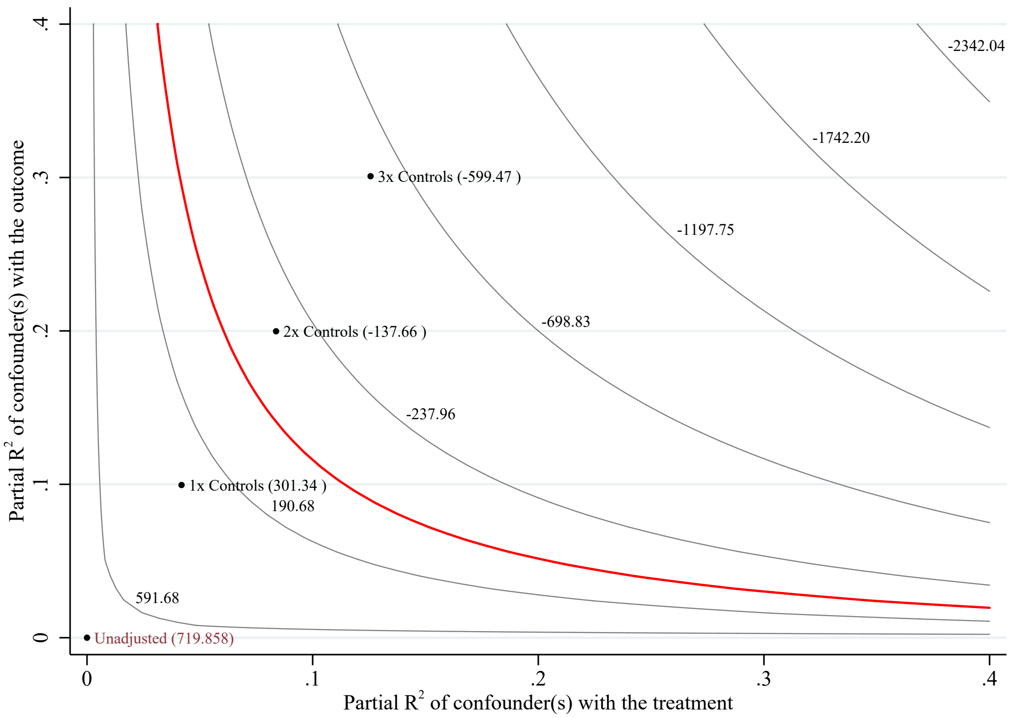
Notes: This graph plots the point estimates and 95 percent confidence intervals of the effects of private schooling on math, English, and Urdu test scores from the different identification strategies. The top panel shows the estimated coefficient, and the bottom panel shows the estimation strategy. For example, the first coefficient for each subject corresponds to an OLS specification with baseline value-added controls, while the second uses the OLS specification with additional SES controls.

Figure A5. : Private Premium Estimates by Civics Sub-Indices



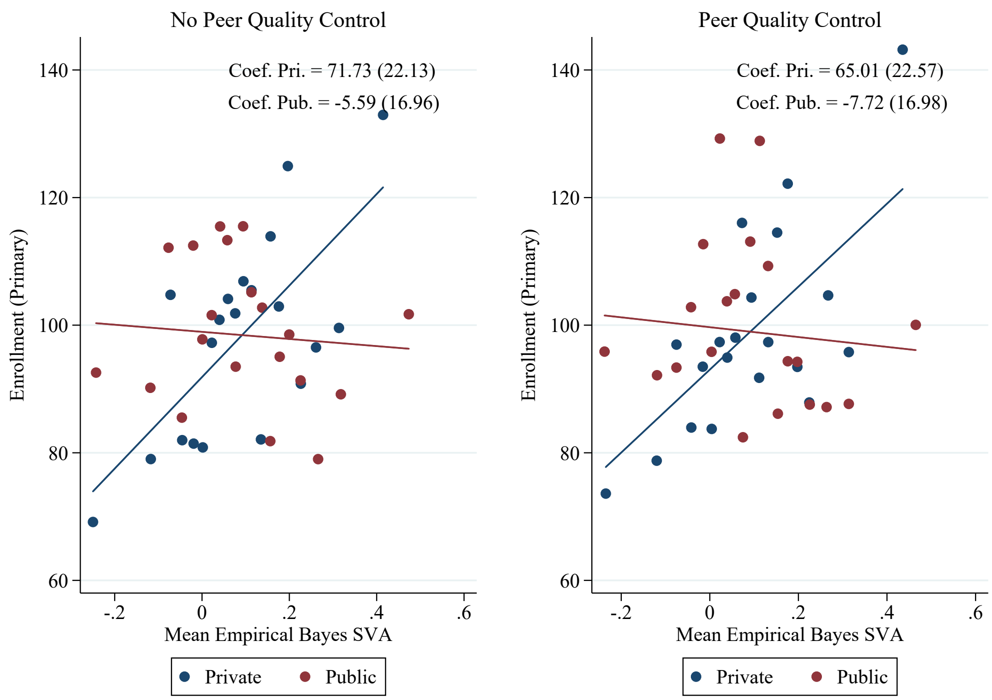
Notes: This graph plots the point estimates and 95 percent confidence intervals of the effects of private schooling on civics scores from the different identification strategies. The top panel shows the estimated coefficient, and the bottom panel shows the estimation strategy. For example, the first coefficient of each subject corresponds to an OLS specification with baseline variables, while the second uses the OLS specification with added SES controls.

Figure A6. : Sensitivity Analysis for the Relationship Between Fees and SVA



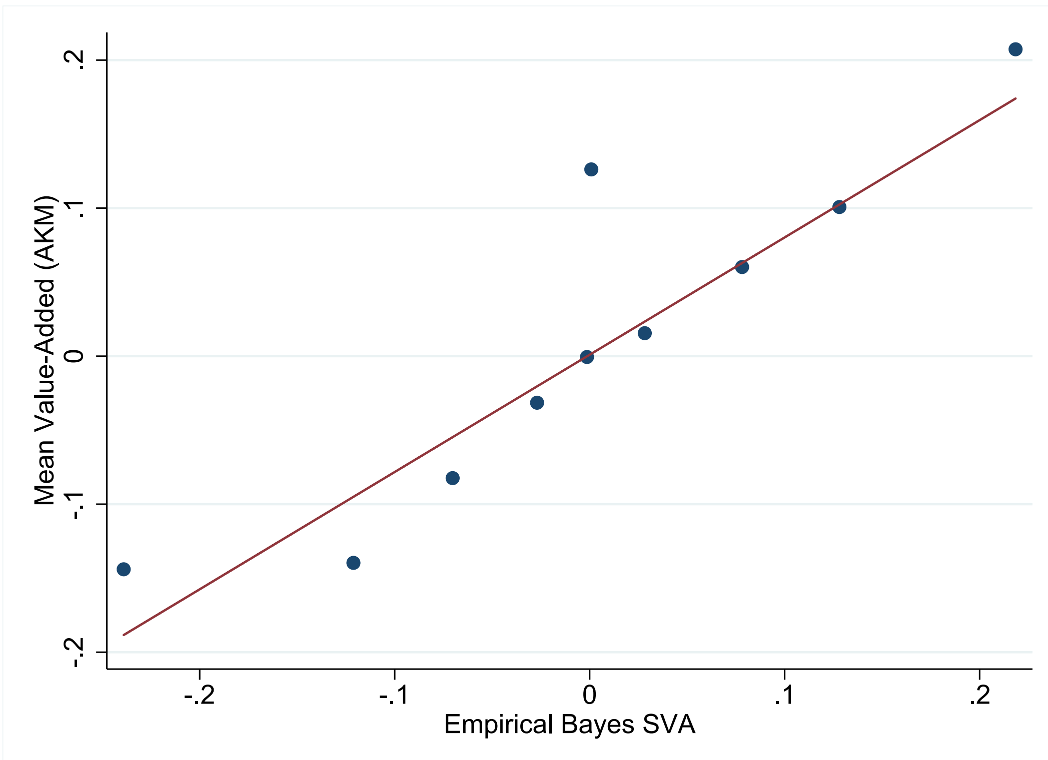
Notes: This graph is a contour plot for the true relationship between fees and SVA under different assumptions about the partial R^2's for the relationship between unobservables and the treatment and outcome variables. The dots benchmark the results to the effects of controls for mother education, father education, and household assets. The graph is produced using the sensemakr command of Cinelli and Hazlett (2020).

Figure A7. : SVA and Enrollment



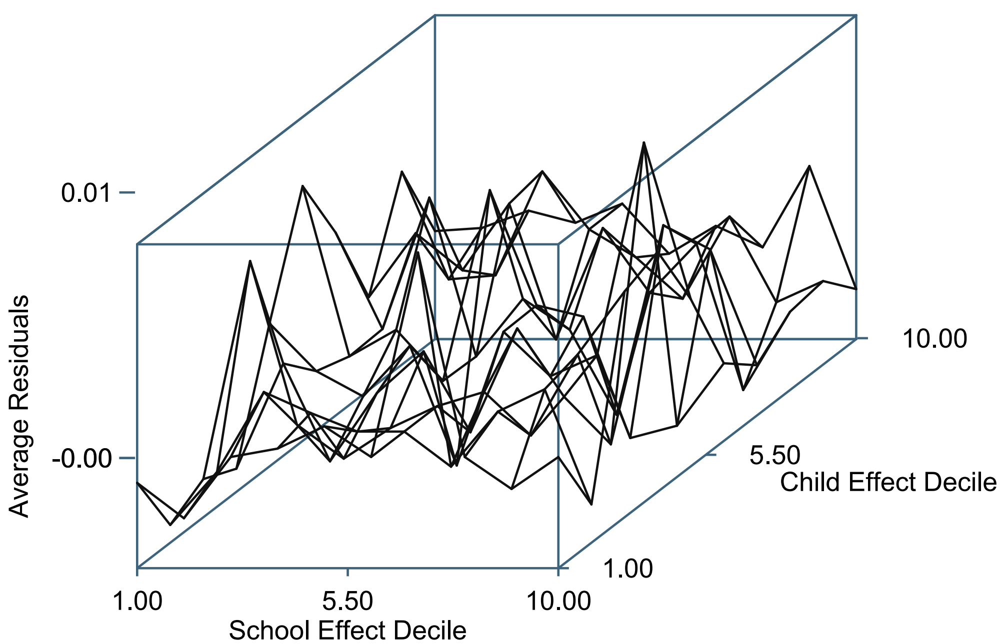
Notes: This figure plots binscatters of primary enrollment against empirical Bayes SVAs in private and public schools separately. Each observation is a school-year. Both panels control for year and district fixed effects, which are allowed to be different in the public and private sectors. The right panel also controls for the school-year level peer quality proxy, which again is allowed to have different effects across sectors. We also report the coefficient on SVA from the analogous OLS regressions. The standard error reported in parentheses is clustered at the school-level.

Figure A8. : Relationship Between Standard and AKM SVA Estimates Within Connected Sets



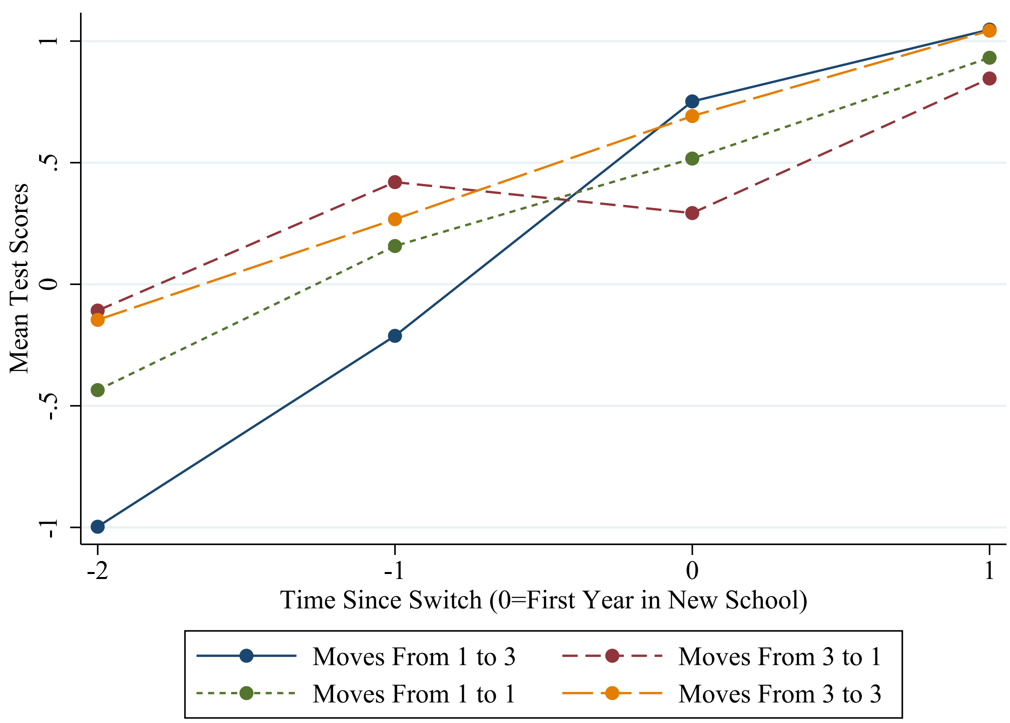
Notes: This figure reports a binscatter plot of the relationship between AKM SVA estimates (equation (C1)) and the standard SVA estimates (equation (1)).

Figure A9. : AKM Validation: Plot of Mean Residuals by Child and School Effects Deciles



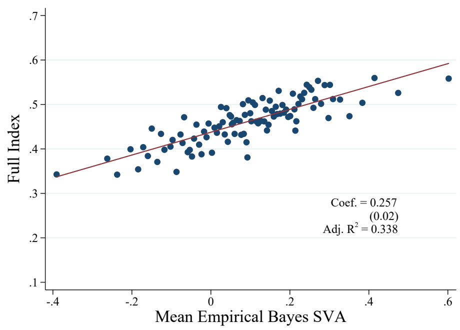
Notes: This figure plots the mean of student-year residuals from equation (C1) by deciles of the estimated school and student effects.

Figure A10. : AKM Validation: Evidence on Symmetric Mobility

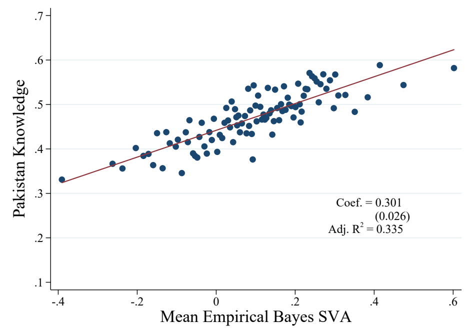


Notes: This figure plots raw trends in mean test scores for students who switch between schools of different terciles according to the AKM SVA estimates (after normalizing the estimates so that they are mean 0 within a connected set). A '1' refers to a bottom tercile school, and a '3' refers to a top tercile schools.

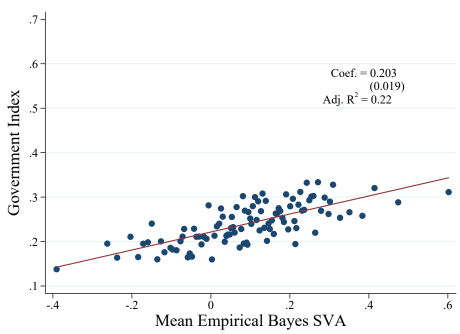
Figure A11. : Association Between School Value-Added and Civic Values



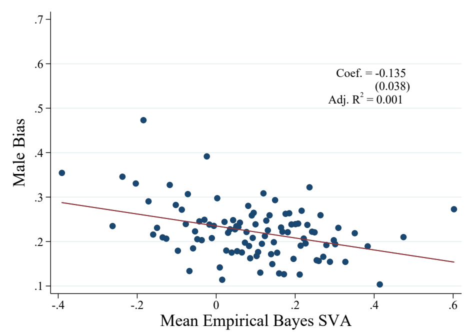
(a) Full Index



(b) Pakistan Knowledge



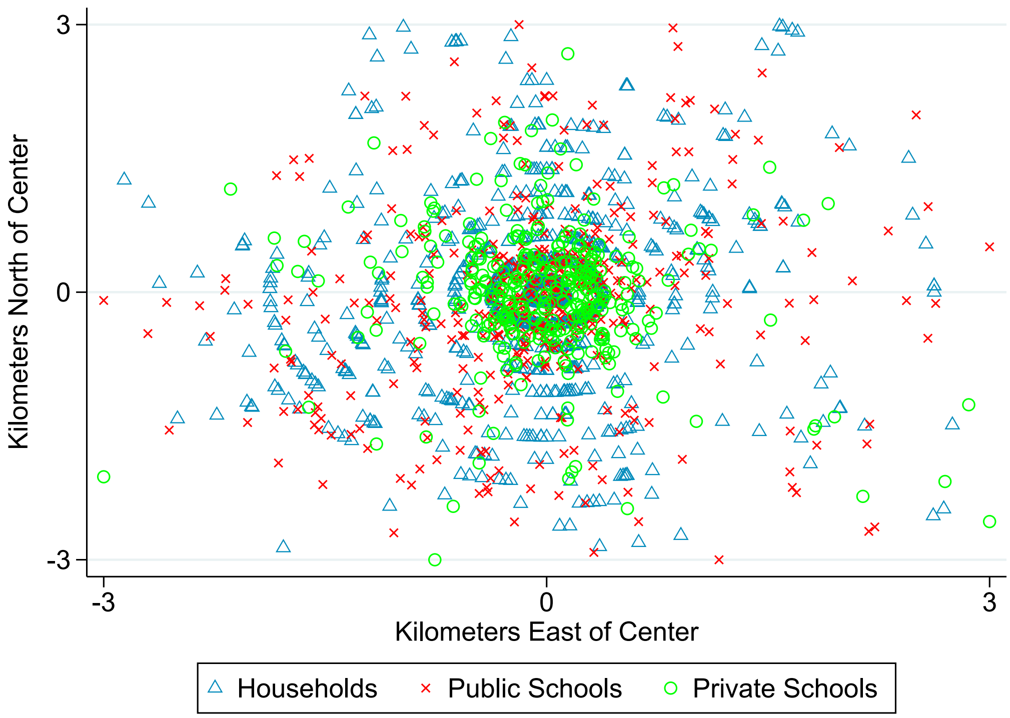
(c) Government Index



(d) Male Bias

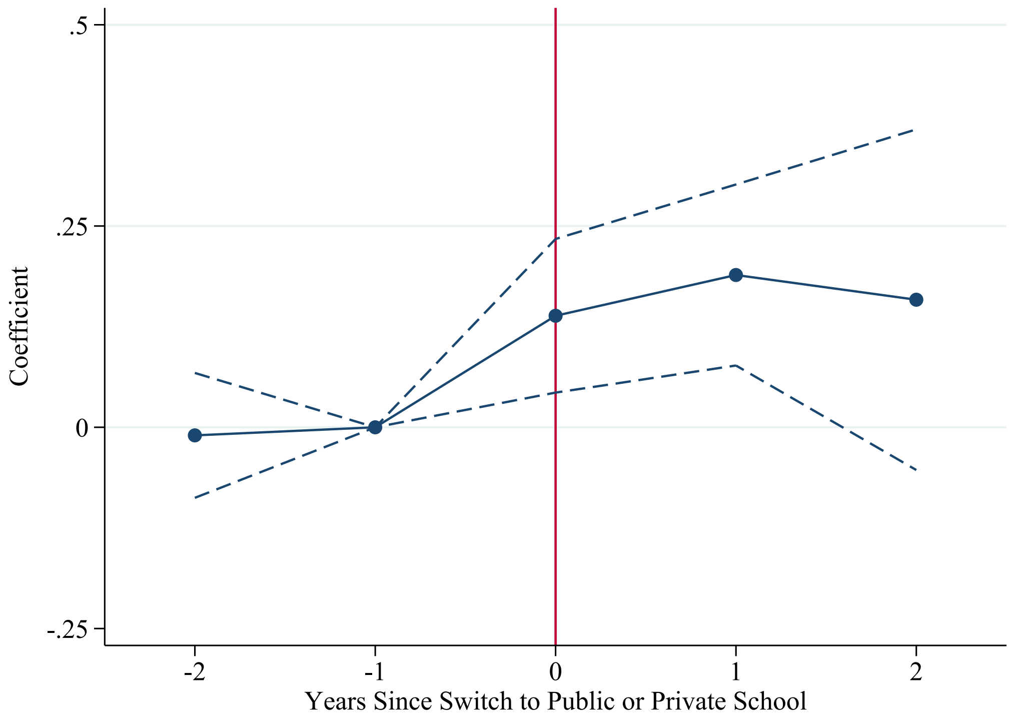
Notes: These graphs plot schools' mean scores on different civics measures against school value-added, controlling for village fixed effects. A more positive score on the full index, Pakistan knowledge, and government index indicates higher civic values. A more negative score on the male bias index indicates less gender bias. We report the estimated coefficient from the OLS regression and standard error, as well as the R^2 adjusted for village fixed-effects for each correlation. Standard errors are clustered at the village-level.

Figure A12. : 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.

Figure A13. : Event Study Graph for Private School Effect (Child FE Estimates)



Notes: This graph reports estimates of the effect of being in a private school k years after a switch to a public or private school for $k \in \{-2,...,2\}$. 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. The solid line denotes the coefficient estimates, and the dashed lines denote the 95% confidence interval.

Table A1—: 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	0.829	0.824	0.897	0.938	0.942
Fill missing letter for picture, Cat	0.556	0.632	0.744	0.805	0.916	0.914	0.950	0.942
Fill missing letter for picture, Flag	0.182	0.197	0.358	0.471	0.508	0.521	0.722	0.720
Fill missing word in sentence	0.227	0.262	0.358	0.463	0.374	0.483	0.638	0.689
Construct sentence with word 'deep'	0.004	0.006	0.020	0.067	0.024	0.028	0.071	0.174
Construct sentence with word 'play'	0.006	0.010	0.052	0.150	0.065	0.070	0.219	0.353
Count number of moons, write number	0.563	0.618	0.749	0.714	0.693	0.740	0.852	0.775
Add 3 + 4	0.884	0.885	0.929	0.931	0.913	0.929	0.962	0.970
Multiply 4 x 5	0.534	0.551	0.686	0.772	0.690	0.755	0.868	0.882
Add 36 + 61	0.810	0.842	0.897	0.916	0.897	0.926	0.955	0.962
Add 678 + 923	0.477	0.505	0.647	0.706	0.666	0.732	0.826	0.826
Subtract 98 - 55	0.647	0.691	0.782	0.838	0.772	0.829	0.892	0.899
Multiply 32 x 4	0.448	0.466	0.622	0.710	0.620	0.712	0.839	0.850
Divide 384/6	0.172	0.183	0.368	0.478	0.224	0.345	0.603	0.657
Cost of necklace, simple algebra	0.075	0.119	0.211	0.243	0.144	0.192	0.341	0.331
Convert 7/3 into mixed fractions	0.020	0.032	0.052	0.106	0.011	0.063	0.091	0.205
Match picture with word, Book	0.675	0.749	0.876	0.925	0.803	0.876	0.958	0.966
Match picture with Urdu word, Banana	0.669	0.747	0.866	0.923	0.802	0.876	0.952	0.969
Match picture with word, House	0.464	0.510	0.612	0.704	0.633	0.724	0.830	0.860
Combine letters into word # 1	0.666	0.729	0.821	0.861	0.831	0.854	0.920	0.942
Combine letters into word # 2	0.286	0.350	0.450	0.544	0.518	0.598	0.700	0.727
Antonyms, Chouta	0.380	0.416	0.604	0.748	0.520	0.615	0.789	0.857
Antonyms, Khushk	0.321	0.401	0.548	0.630	0.420	0.584	0.749	0.772
Complete passage for grammar	0.248	0.296	0.476	0.624	0.369	0.511	0.678	0.758

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

Table A2—: 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 civics 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 but not when results for male bias are reported separately.

Table A3—: Correlation Between School Closure and Students' Characteristics

	(1) Mom Education	(2) Dad Education	(3) Household Assets	(4) Child High Ability	(5) Female
School Closure	-0.040	-0.057*	-0.082	-0.053*	-0.043
	(0.031)	(0.031)	(0.138)	(0.028)	(0.032)
Adjusted R^2	0.10	0.07	0.09	0.04	0.05
Number of Observations	16902	16904	16903	14852	20276
Number of Clusters	634	634	634	633	641

Notes: This table reports the coefficients from regressions of student characteristics in the survey conducted in schools on the school closure indicator variable. The regressions include controls for village fixed effects, grade fixed effects, year fixed effects, 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 A4—: Association Between School Inputs and SVA

	Dep. Var.: School Value-Added	
	(1)	(2)
	Public	Private
Library	-0.098**	0.065
	(0.043)	(0.051)
Computer	0.114	0.054
	(0.114)	(0.049)
Sports	0.034	0.080
	(0.056)	(0.051)
Hall	-0.042	-0.122
	(0.092)	(0.075)
Wall	0.018	-0.078
	(0.036)	(0.106)
Fans	0.051	0.074
	(0.052)	(0.088)
Electricity	-0.040	-0.043
	(0.052)	(0.085)
Teacher Ratio	-0.001	0.002
	(0.001)	(0.002)
Teacher Absenteeism	-0.023***	-0.018
	(0.007)	(0.011)
Adjusted R^2	0.333	0.163
Within Adj. R^2	0.011	0.037
Number of Observations	474	319
Number of Clusters	112	108

Notes: This table regresses the (non-shrunk) fixed effect estimates of schools' mean SVA's on other school characteristics measures. Inputs measures are the means across all four years to account for variation in facilities over time. Teacher absenteeism is the average number of days teachers are absent in a month across all four years. All regressions include district fixed effects. 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 A5—: Effect of a One Standard Deviation Better School

	(1) Public	(2) Private
Math	0.321	0.223
English	0.358	0.250
Urdu	0.269	0.153
Average	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, and Urdu, as well as the average effect across the three.

Table A6—: Partial Equilibrium Effects of Reallocating Students (NPEB Estimates)

	Current Pub. to Best Pri.			Current Pub. to Worst Pri.			Worst Pub. to Best Pri.			Best Pub. to Worst Pri.		
	(1) p(10)	(2) Mean	(3) p(90)	(4) p(10)	(5) Mean	(6) p(90)	(7) p(10)	(8) Mean	(9) p(90)	(10) p(10)	(11) Mean	(12) p(90)
Math	-0.307	0.146	0.573	-0.638	-0.108	0.356	-0.250	0.254	0.645	-0.749	-0.272	0.179
English	0.021	0.523	1.101	-0.323	0.209	0.749	0.219	0.722	1.208	-0.577	0.027	0.575
Urdu	-0.135	0.213	0.594	-0.365	-0.016	0.322	-0.040	0.338	0.642	-0.518	-0.159	0.116
Mean	-0.049	0.294	0.648	-0.334	0.028	0.377	0.135	0.438	0.761	-0.446	-0.135	0.202

Notes: This table uses non-parametric 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, as well as the effect of moving students from the worst public school to the best private school in the village (columns 7-9) and from the best public to the worst private (columns 10-12). Columns 1, 4, 7, and 10 report the effect on students in the 10th percentile of test score gains. Columns 2, 5, 8, and 11 report the average effects, and columns 3, 6, 9, and 12 report the effects on students in the 90th percentile of test score gains.

Table A7—: 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^2	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, and year fixed effects, as well as 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 A8—: Value-Added Estimates of the Yearly Effect of Private Schooling, 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^2	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. All regressions include grade fixed effects, child fixed effects, lagged test scores interacted with grade level, and controls for age, age squared, and year fixed effects, as well as 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 A9—: Association Between Private Schooling and 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^2	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 association of private schooling with civic values scores. All regressions include controls for grade fixed effects, gender, and controls for age, age squared, year fixed effects, 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 children with civic values scores in year 3. 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 A10—: Relationship Between Civics and SVA

	Public Schools				Private Schools			
	(1) Full Index	(2) Pakistan Knowledge	(3) Government Index	(4) Male Bias	(5) Full Index	(6) Pakistan Knowledge	(7) Government Index	(8) Male Bias
Mean Emp. Bayes SVA	0.250***	0.293***	0.203***	-0.108**	0.236***	0.258***	0.227***	-0.191***
	(0.025)	(0.029)	(0.025)	(0.054)	(0.038)	(0.055)	(0.050)	(0.071)
Library	0.012	0.002	0.006	-0.037	-0.011	-0.011	0.000	0.010
	(0.017)	(0.022)	(0.017)	(0.033)	(0.017)	(0.022)	(0.017)	(0.023)
Computer	-0.075*	-0.048	-0.125***	0.126	-0.011	0.001	0.008	0.029
	(0.039)	(0.045)	(0.038)	(0.091)	(0.014)	(0.018)	(0.017)	(0.019)
Sports	-0.009	-0.018	0.020	-0.006	-0.005	-0.004	-0.007	0.003
	(0.020)	(0.024)	(0.022)	(0.041)	(0.015)	(0.020)	(0.015)	(0.027)
Hall	-0.013	-0.003	-0.001	0.014	0.025	0.024	0.039	0.016
	(0.020)	(0.025)	(0.020)	(0.054)	(0.019)	(0.025)	(0.029)	(0.032)
Wall	0.035***	0.022*	0.027**	-0.162***	0.013	-0.003	-0.011	-0.110*
	(0.010)	(0.012)	(0.010)	(0.020)	(0.056)	(0.073)	(0.056)	(0.059)
Fans	0.054***	0.048**	0.030	-0.092**	0.026	0.007	-0.033	-0.300***
	(0.018)	(0.022)	(0.019)	(0.039)	(0.059)	(0.070)	(0.077)	(0.082)
Electricity	-0.039*	-0.034	-0.021	0.073*	-0.042	-0.018	-0.010	0.233***
	(0.020)	(0.024)	(0.020)	(0.039)	(0.052)	(0.061)	(0.044)	(0.088)
Teacher Ratio	0.000	0.000	0.000	-0.001***	0.000	0.000	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Teacher Absenteeism	0.004	0.002	0.004	-0.023***	-0.009***	-0.011***	-0.005	0.011
	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)	(0.004)	(0.004)	(0.007)
Adjusted R^2	0.09	0.09	0.05	0.05	0.10	0.10	0.05	0.04
Number of Observations	16861	16861	16861	14610	6777	6777	6777	6446
Number of Clusters	112	112	112	112	108	108	108	108

Notes: This table regresses students' civics scores on the empirical Bayes estimates of schools' mean SVA's and other school characteristics measures. Inputs measures are the means across all four years to account for variation in facilities over time. Teacher absenteeism is the average number of days teachers are absent in a month across all four years. Standard errors are clustered at the village-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A11—: Summary Statistics for the Combined Household and Tested Sample

	All			Public Schools			Private Schools			Difference		
	(1) Mean	(2) SD	(3) N	(4) Mean	(5) SD	(6) N	(7) Mean	(8) SD	(9) N	(10) Mean	(11) SE	(12) 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	10.498	1.931	3359	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	0.628	0.813	3292	0.698	0.851	2342	0.404	0.651	736	0.294	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. This set of students is the relevant sample for the first stage estimation of the effect of private schooling on test scores using the distance to primary school instrument since primary schools ends in grade 5. Column 12 reports the p-value for a t-test of the differences in means between public and private schools.

Table A12—: 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
Relative Distance IV	-0.029	-0.049	-0.113	-0.001	-0.027	-0.324	-0.010	-0.020	0.143	0.024
	(0.033)	(0.033)	(0.178)	(0.052)	(0.019)	(1.280)	(0.008)	(0.026)	(0.091)	(0.093)
Adjusted R^2	0.19	0.15	0.13	0.15	0.07	0.09	0.44	0.05	0.09	0.07
Number of Observations	3237	2898	3269	2385	3269	2382	3269	2368	2075	2075
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 fixed effects, grade fixed effects, gender, age, age squared, distance to the village center, year fixed effects, and the interaction between the age controls, distance to the center, year fixed effects, and gender. Standard errors are clustered at the village-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A13—: Distance IV Estimates of Effect of Private Schooling on Test Scores

	(1) First Stage	(2) Math	(3) English	(4) Urdu	(5) Mean
Relative Distance IV	-0.317**				
	(0.151)				
Years in Private		0.002	0.288***	0.133*	0.141**
		(0.072)	(0.080)	(0.075)	(0.067)
F-statistic	10.23				
Number of Observations	5987				
Number of Obs. 1st Stage		5987	5987	5987	5987
Number of Obs. 2nd Stage		3118	3118	3118	3118

Notes: This table reports the two sample 2SLS results of the effect of private schooling on test scores. All regressions include controls for village fixed effects, grade fixed effects, gender, age, age squared, distance to the village center, year fixed effects, and the interaction between the age controls, distance to the center, year fixed effects, 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 A14—: Effect of Private Schooling on Civic Values With Two-Sample Distance IV

	(1) First Stage	(2) Full Index	(3) Pakistan Knowledge	(4) Government Index	(5) Male Bias
Relative Distance IV	-0.320**				
	(0.151)				
Years in Private		0.027*	0.030*	0.015	-0.157***
		(0.015)	(0.018)	(0.023)	(0.030)
F-statistic	10.67				
Number of Observations	5987				
Number of Obs. 1st Stage		5987	5987	5987	5987
Number of Obs. 2nd Stage		1039	1039	1039	970

Notes: This table reports the two sample 2SLS estimates of the effect of private schooling on civics values measures. All regressions include controls for village fixed effects, year fixed effects, grade fixed effects, gender, age, age squared, distance to the village center, and the interaction between the age controls, year fixed effects, 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 A15—: 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 instrumental variable estimates of the effect of private schooling on test scores. All regressions include controls for village fixed effects, grade fixed effects, grade fixed effects interacted with lagged test scores, and gender, as well as age, age squared, and year fixed effects, 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 were 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 A16—: 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 instrumental variable estimates of the effect of private schooling on civic values scores. All regressions include controls for village fixed effects, grade fixed effects, grade fixed effects interacted with lagged test scores, and gender, as well as age, age squared, and year fixed effects, 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 were 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 A17—: Predictors of Private School Closure Between 2003 and 2007

	Dep. Var.: Indicator variable equal to 1 if a school closed			
	(1)	(2)	(3)	(4)
Lag Enrollment	-0.038***	-0.036***		
	(0.009)	(0.011)		
Emp. Bayes SVA			-0.004	-0.039
			(0.017)	(0.024)
Fixed Effect	District	Village	District	Village
Adjusted R^2	0.048	0.069	0.020	0.051
Number of Observations	895	892	1188	1188
Number of Clusters	111	108	108	108

Notes: This table regresses private school closure on lagged school enrollment (columns 1 and 2) and the empirical Bayes mean SVA (columns 3 and 4). An observation is a private school-year, and the year after a school closes, it is dropped from the sample. All specifications control for year fixed effects. Lag enrollment is measured in 100s of students for readability. Odd columns include district fixed effects, while even columns include village fixed effects. Standard errors are clustered at the village-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A18—: Relationship Between SVA and School Closures, 2003-2007

	Public SVA	Private - Public SVAs (Percentile)
	(1)	(2)
Village had Private Closure	-0.111**	6.816
	(0.051)	(4.777)
Distance to Closed School (km)	0.048***	
	(0.018)	
Mean Outcome	-0.005	22.892
Adjusted R^2	0.011	0.009
Number of Observations	475	108
Number of Clusters	112	108

Notes: This table examines why the school closure IV may deliver a larger estimate of the private premium than other strategies. The first column regresses public school SVA (non-shrunk) on an indicator variable for whether the school is in a village with a private closure and a measure of the distance between the public school and the closed school. An observation is a public school. The second column regresses the village-level percentile for the difference in mean private and public SVAs (non-shrunk) on whether the village had a private closure. An observation is a village, and standard errors are clustered at the village-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table A19—: Parental Response: School Fees and SVA

	(1)	(2)
	Dep. Var.: School Fees	
Empirical Bayes SVA	991.371***	719.858**
	(320.963)	(280.882)
Mean Mother Education		4.697
		(122.849)
Mean Father Education		218.651
		(142.686)
Mean HH Asset Index		263.074***
		(46.820)
Adjusted R^2	0.22	0.29
Number of Observations	1158	1148
Number of Clusters	318	318

Notes: This table reports the results from regressing a measure of total primary school fees on private schools' characteristics. The peer quality measure is the standardized school-year average of predicted student test scores. All regressions include year and district fixed effects. An observation is a school-year, and standard errors are clustered at the school-level.

Table A20—: School Choice Criteria and Parental SES

	(1)	(2)	(3)
	Distance	Quality	Cost
SES Factor	-0.017	0.055***	-0.055***
	(0.012)	(0.010)	(0.010)
Mean Dep. Var.	0.42	0.20	0.27
Adjusted R^2	0.11	0.03	0.12
Number of Observations	2432	2432	2432
Number of Clusters	978	978	978

Notes: This table reports the results from regressing an indicator for whether parents reported their reason for choosing a school was distance (column 1), quality (column 2), or cost (column 3) on a measure of household socioeconomic status. The measure is the first factor from a pca of mother education, father education, and an asset index. All regressions include district fixed effects (the reasons for school choice question was only asked in the first round of surveys of parents of enrolled children). An observation is a child, and standard errors are clustered at the household-level.

Table A21—: Response to Variation in Parental Quality Responsiveness by Sector

	SVA	Δ Market Share 2006-2011	Closure 2011
	(1)	(2)	(3)
Panel A: Private Schools			
Village Avg. SES	0.103***	-0.035***	0.123
	(0.034)	(0.013)	(0.095)
Emp. Bayes SVA		0.072**	-0.633***
		(0.030)	(0.214)
Village Avg. SES × SVA		0.120**	-0.701*
		(0.054)	(0.355)
Adjusted R^2	0.19	0.01	0.03
Number of Observations	282	274	282
Number of Clusters	84	84	84
Panel B: Public Schools			
Village Avg. SES	0.034	0.002	0.012
	(0.031)	(0.008)	(0.022)
Emp. Bayes SVA		0.033*	-0.081
		(0.017)	(0.062)
Village Avg. SES × SVA		0.030	-0.016
		(0.036)	(0.099)
Adjusted R^2	0.328	-0.002	0.047
Number of Observations	475	471	475
Number of Clusters	112	112	112
P-value: Pri ≠ Pub	0.126	0.290	0.004

Notes: This table analyzes the relationship between village's socioeconomic status and private and public schools' SVA, change in market share, and closure in 2011. The village average SES measure is the round 4 (2006-2007) average of the first factor from a pca of mother education, father education, and an asset index in the household data. Emp. Bayes SVA is the mean school value added measure calculate using empirical Bayes. All regressions include district fixed effects and controls for experimental interventions in the LEAPS villages (the report card treatment studied by Andrabi, Das and Khwaja (2017) and school grants program studied by Andrabi et al. (2021)). An observation is a private school, and standard errors are clustered at the village-level. The sample in Panel A is restricted to private schools in competitive villages in round 4 (villages where there was at least one other private school).

*

REFERENCES

Abowd, John M, Francis Kramarz, and David N Margolis. 1999. "High wage workers and high wage firms." *Econometrica*, 67(2): 251–333.

Andrabi, Tahir, Jishnu Das, and Asim Khwaja. 2017. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." *American Economic Review*, 107(6): 1535–1563.

Andrabi, Tahir, Natalie Bau, Jishnu Das, Naureen Karachiwalla, and Asim Khwaja. 2021. "Crowding in Private Quality: The Equilibrium Effects of Public Spending in Education." *Working Paper*.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2019. "Nation-building through compulsory schooling during the age of mass migration." *The Economic Journal*, 129(617): 62–109.

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

Bau, Natalie, and Jishnu Das. 2020. "Teacher Value-Added in a Low-Income Country." *American Economic Journal: Economic Policy*, 12(1).

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly journal of economics*, 128(3): 967–1015.

Cinelli, Carlos, and Chad Hazlett. 2020. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1): 39–67.

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

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

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

Johnson, Carol, and Alan Vanneman. 2001. "Civics: What Do 4th-Graders Know, and What Can They Do?." *NAEPfacts*, 6(2): n2.

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

Paustian, Paul William. 1930. *Canal irrigation in the Punjab: An Economic Inquiry Relating to Certain Aspects of the Development of Canal Irrigation by the British in Punjab.* New York:Columbia university press.

Pritchett, Lant. 2013. *The rebirth of education: Schooling ain't learning.* CGD Books.