Decomposing the Government-Private School Performance Differential: Village Ethnic Politics and School Sorting

Nicholas Eubank*
Stanford University

July 21, 2016

Click here to download the most recent version of this paper.

Abstract

The emergence of rural, secular, affordable private schools across South Asia is one of the most promising recent developments in the education sector. Yet whether private schools provide superior educations remains unclear. Observational studies consistently show private school students outperform government students even when controlling for demographic characteristics and some unobservable heterogeneity using Value-Added models. Nevertheless, it remains unclear whether this is because (a) private schools provide students with a better education, or (b) students attending private schools are more academically inclined in unobservable ways. Using data from the Learning and Educational Attainment in Punjab Schools (LEAPS) survey, this paper leverages variation in sorting on academic potential caused by village caste politics to isolate the component of private school performance caused by sorting rather than superior teaching. It concludes that even the most sophisticated observational techniques – lagged Value-Added models – overstate private school quality by at least half.

*Post-Doctoral Fellow in Political Science. nicholaseubank@stanford.edu. This project would not have been possible without the exceptional support numerous parties, including Jishnu Das, Tahir Andrabi, Alex Lee, Kate Casey, Neil Malhotra, Meredith Startz, and Paul Novosad. Code and replication directions can be found at `www.github.com/nickeubank/pk_schooling`

1 Introduction

The rapid rise of affordable and purportedly high quality private schools in South Asian rural communities is one of the most exciting developments in the education sector in decades. Private schools account for an ever rising share of children attending school – in 2005, 33% of Pakistani primary school students and 20-24% of Indian rural primary school students attended a private school, and the students in these private schools consistently outperform their government school counterparts, even when controlling for observable student characteristics (Jimenez et al., 1991; Jimenez and Lockheed, 1995; Pratham, 2005; Andrabi et al., 2011; Desai et al., 2009; Tooley and Dixon, 2003; Alderman et al., 2003, 2001). This has given rise to the hope that private schools may someday circumvent reform-resistant government schools and finally deliver quality education to the hundreds of millions of children in the region.

Despite the promise of these developments, however, the true significance of the rapid emergence of private schools hinges critically on the question of whether (a) these private schools are actually delivering superior educations, or whether (b) they just attract students who are more academically inclined or come from families that prioritize educational attainment. In other words, are observed differences in school performance due to better teaching or student sorting on unobservable academic potential.

To help answer this question, this paper takes advantage of an observable source of variation in student sorting – caste politics. In particular, this paper shows that in caste-homogeneous villages, students tend to sort on perceived academic potential, with parents sending their more academically gifted children to private schools. In caste-heterogeneous villages, however, school choice is also shaped by a desire to keep children in caste-homogeneous schools. High-status families tend to send their children to private schools to keep them in homogeneous social settings *regardless of their perceived academic abilities*, leading to less academic sorting in these villages.

This paper argues that as a result, differences in the private school / government school test-score gap between heterogeneous and homogeneous villages can be attributed to differences caused by differential sorting on unobservable academic potential, offering an opportunity to estimate the impact of this otherwise difficult-to-measure phenomenon.

To take advantage of this variation, this analysis proceeds in three steps. First, this analysis estimates the performance differential between government and private schools using lagged-value-added models applied to a four-year panel of child test scores with demographic controls. This technique is currently considered to be the most rigorous method of studying observational education data (Gordon et al., 2006; McCaffrey et al., 2003; Hanushek, 2003). It estimates the test-score gap between government and private school students that remains after controlling not only for observable differences in child demographics, but also some non-observable differences (a value often referred to as each school-type's "value-added").[1] This constitutes a baseline estimate of the government-private school performance differential using non-experimental data.

This analysis then compares differences in the government-private test-score gap in homogeneous villages with the test-score gap in heterogeneous villages. Because school choice is primarily based on academic potential in homogeneous villages and not in heterogeneous villages, under a mild set of assumptions detailed below, differences in the test-score gaps found in homogeneous and heterogeneous villages can be attributed to differences in academic sorting.

This analysis finds that while private schools outperform government schools in all villages, the

[1] Lagged-value-added models control for unobservable differences that affect test score *levels*, although they cannot control for unobserved heterogeneity in learning rates. These issues are discussed in more detail in Section 4.1.

amount they outperform government schools falls by half when moving from homogeneous villages to heterogeneous villages. This implies that *at least half* of the estimated superior performance of private schools is due not to better teaching, but rather to unobservable differences in the quality of students in private schools that cannot be accounted for by lagged-value-added models.

This conclusion is supported by two other sets of results presented in this analysis. First, and most importantly, this analysis is unable to find any other differences between homogeneous and heterogeneous villages which might account for changes in test-score gaps. As shown in Section 2.2, for example, caste heterogeneity does not appear to be well correlated with village median wealth, adult literacy, land inequality, the number of schools per household, or number of households. Moreover, other factors often cited as explanations for the government-private performance differential – like performance pay in private schools or differences in school resources – do not appear to vary systematically with village heterogeneity 5.

Second, at the level of villages there is no evidence that *overall* learning outcomes vary with village caste composition. As detailed in Section 4.3, the decrease in the government-private school test-score gap is the result of off-setting convergence in test scores between the two school types, not a change in overall learning outcomes. This is consistent with a re-distribution of student talent rather than actual differences in teaching quality.

There are two nuances to the conclusions drawn here that are worth noting. First, the difference between homogeneous and heterogeneous villages is best interpreted as a *lower-bound* on the contribution of sorting to estimates of the government-private performance differential. There is likely still *some* sorting on potential in heterogeneous villages. Thus the comparison between heterogeneous and homogeneous villages is best understood as a comparison between villages with sorting and villages with *less* sorting, not a difference between villages with and without sorting.

Second, this analysis is motivated by the assumption that sending high-caste children to private schools and low-caste children to government schools does not constitute sorting on ability. For this to be true, it must be the case that residual academic potential – potential that cannot be accounted for by factors like parental education and wealth that enter into lagged-value-added models – must be equally distributed across different castes (or be distributed slightly in favor of lower status *biraderis*). As shown in Section 4.4, however, there is no evidence that those from higher social status *biraderis* have higher residual talent than those from low status *biraderis*; student caste does not appear to have any consistent effect on test scores.

It is difficult to overstate the potential importance of the answer to this question for education policy in the developing world. Not only do private schools constitute a substantial portion of current enrollments in South Asia, but enrollment is also growing explosively. From 2000 to 2005 in rural Pakistan, for example, the number of private schools in Pakistan rose from 32,000 to 47,000. (Andrabi et al., 2007, p. vi), and evidence suggests this growth continues today. Moreover, private schools deliver educations at a fraction of the cost of government schools by not requiring formal teacher training and by hiring local, secondary-educated women as teachers rather than college-educated teachers who have to move to the villages where they teach (Andrabi et al., 2007). Thus the question of whether these lower-cost educations are of similar quality to government school educations also speaks to the importance of different teacher training and employment practices.

The findings of this paper are unlikely to put to rest the debate over whether private schools are superior to government schools. Indeed, private school students continue to outperform government school students even in the most heterogeneous villages, just by a dramatically smaller margin.

But even after using some of the most sophisticated econometric methods available, this paper still establishes (as a "lower bound") that sorting explains a very large portion of the government-private test-score gap. This should give analysts pause when examining other empirical results that claim to fully control for sorting.[2]

This paper is organized as follows: Section 2 provides an overview of the rural Pakistan context from which data for this analysis is drawn. Section 3 details how the determinants of school choice vary with village caste heterogeneity. Section 4 then shows how the government-private test-score gap varies with caste heterogeneity, and presents evidence this is due to differences in student sorting. Section 5 then examines and rules out a number of alternative possible explanations for this empirical regularity. Finally, Section 6 discusses the strengths and weaknesses of these findings and their interpretation.

2 Study Context

The focus of this paper is the rural, secular, primary school educational ecosystem of 112 randomly selected villages (*mauzas*) in the Punjab districts of Attock, Faisalabad, and Rahim Yar Khan in Pakistan. These 112 villages were the subject of the Learning and Educational Attainment in Punjab Schools (LEAPS) panel survey. LEAPS villages were selected through probability proportional to population sampling (stratified at the district level) from the universe of all villages in these districts with at least one private school. The survey ran from to 2003-2007, and includes detailed surveys of households, students, teachers, and both government and private schools in these villages. In addition, some data used in this analysis comes from a listing census conducted in sample villages prior to the start of the LEAPS survey that include basic demographic information on all households, allowing for the computation of accurate village level statistics.

The LEAPS survey is organized around two panels of students – one (initiated in 2003 with an initial population of 12,110 children) which followed students for four years, and one (initiated in 2005 with an initial population of 11,852 students) which followed students for two years. Each panel represents the universe of enrolled students in sample villages in Class 3 at the time of panel initiation in both government and private schools. Students and their teachers in both panels were administered annual exams in English, math, and Urdu by the LEAPS team, and these tests were subsequently standardized using Item Response Theory (IRT) methods.

2.1 Caste in Punjab

This paper focuses on how village-level caste politics influence the selection of students into private and government schools. Before directly examining this dynamic, however, some background on the nature of caste in Pakistan is warranted.

[2]It also bears noting that experimental studies of government-private school test differentials are, as currently implemented, not an empirical silver bullet. Randomizing school assignments is untenable, but in two major cases private school vouchers have been randomly assigned. Even these randomizations have proven problematic, however. Angrist et al. (2002) examines a voucher lottery system in Colombia and finds a small positive effect of vouchers, but inference is clouded by the fact that voucher students who performed poorly were at risk of losing their vouchers, making it impossible to separate this incentive effect from the private school effect. And several studies have been conduced of a voucher system in Chile, but as Bellei (2008) notes, the slight private-school advantage these studies show may be down to the fact that private school admissions are selective and poorly performing students can be expelled from private schools, making it difficult to disentangle selectivity from school effects.

Caste – known variously as *biraderi*, *zaat*, or *qaam* – is a central aspect of rural social identity in Pakistan, especially in Punjab. While *biraderi* is a somewhat distinct concept from the idea of "caste" in India, "it retains a very important feature of the [Indian subcaste] – that of an inherent, inbuilt hierarchy that governs social interactions. Society is hierarchically ordered with the Syeds at the top, followed by the landowning castes, then by the service castes or kammis, and finally by the Musallis, who occupy the lowest rung of the social ladder. This ordering dictates much of the social life in a Punjabi village and is most profound in the notions of community cooperation, where solidarity is strongest within a biraderi."(Gazdar and Mohmand, 2007, p. 29).

Biraderi is correlated with wealth, land holdings, and education, but it is not synonymous with economic class. As a related report observes, "while economic power is required to reinforce biraderi-based dominance, membership of a dominant biraderi can help mitigate some of the effects of being economically poor. As one respondent put it, 'the poorest Jatt is still better off than the richest kammi."' (Gazdar and Mohmand, 2007, p. 13)

2.2 Village Caste Composition

While the centrality of caste politics is relatively universal across villages in Punjab, however, there is significant variation in village *biraderi* composition. Figure 1 below shows the density plot of villages of different levels of caste fractionalization, measured as one minus a herfindahl index.[3] As the figure shows, there is significant variation in the degree of fractionalization both within and across the three districts of the LEAPS survey.

As shown in Table 1, this variation in heterogeneity is not clearly related to village wealth, land inequality, or adult education. There is some relationship to village size, but on the whole fractionalization appears to be relatively independent of other village characteristics.

Table 1: Village Characteristics and Fractionalization

	(1) Median Wealth	(2) Adult Literacy	(3) Land Gini	(4) Enrollment Pct	(5) Schools per HH	(6) Log Num HH
Fractionalization	-314.1	-6.66	0.10	-14.1	0.0010	0.64*
	(146.8)	(15.0)	(0.074)	(13.3)	(0.0073)	(0.21)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112	112	112	112	112	112

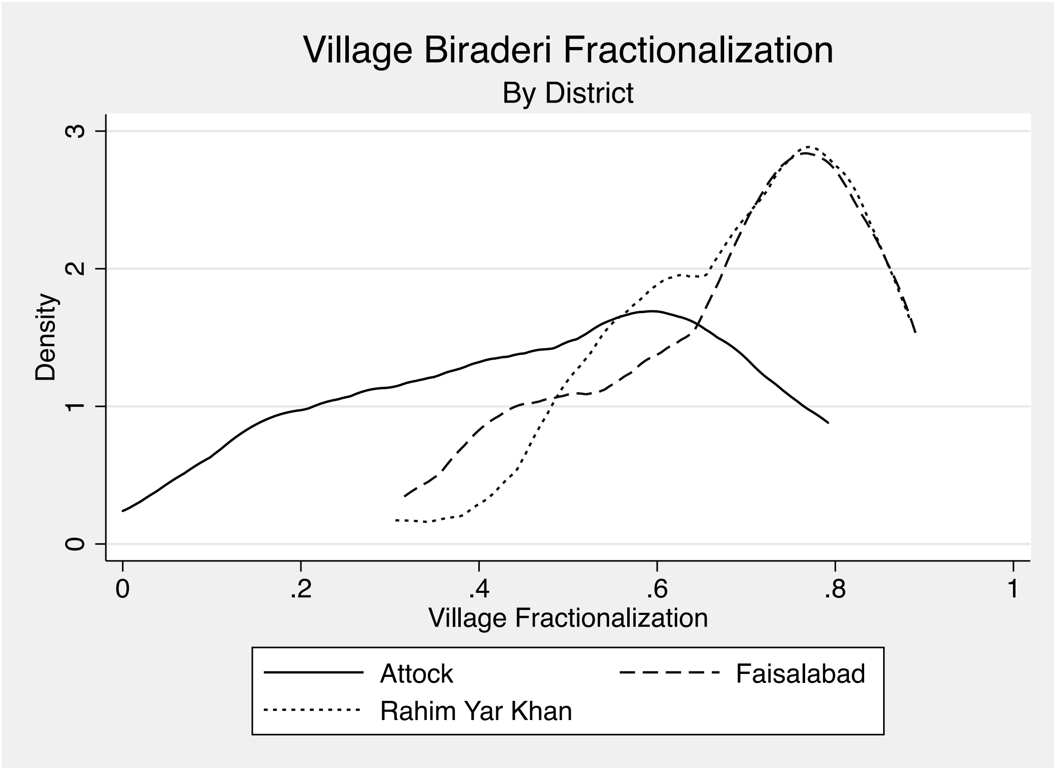
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors clustered by District.

[3]One minus the herfindahl index is a common measure of fractionalization equal to the probability that any two randomly selected individuals belong to the same group. Village fractionalization is computed using data from a village census conducted in 2002 to facilitate household sampling for the LEAPS survey which includes data on the *biraderi* of all households in LEAPS villages.

Figure 1: Village Caste Fragmentation by District



3 Caste Politics and School Sorting

The key to determining whether the government-private test-score gap is caused by differences in quality or simply student sorting is understanding what motivates parents to pick one type of school over the other. This Section provides an overview of how households make these choices, and how those choices vary by village caste composition.

3.1 Selection in Homogeneous Villages

Parents in the LEAPS survey shown a marked tendency to invest preferentially in the children they view as having the most potential. As noted by the original LEAPS survey authors, "through their choices of whether to enroll a child, through the choice of school ([government] or private) and finally through the amount they chose to spend, households pick "winners" and try to carry them through." (Andrabi et al., 2007, p. 103)

One manifestation of this is that conditional on sending their children to school, they are more likely to send their children to a private school if they perceive them as being more intelligent. This is illustrated in Table 2 below, in which an indicator for whether an enrolled child attends a private school is regressed on parental perceptions of child intelligence and a number of demographic controls. The results show clearly that child intelligence are a strong predictor of whether a parent will send their child to a private school. Most notably, these results show this pattern holds even *within individual households*. As shown in Column 2, which includes household fixed effects, many parents send the child they perceive to be more intelligent to private school and the child they perceive to be less intelligent to government schools.

Table 2: Probability Enrolled Child Attends Private School

	(1) Village FE	(2) HH FE
Mom Reports Child Above Average Intelligence	0.058***	0.043**
	(0.021)	(0.021)
Mom Has Some Schooling	0.077	-0.029
	(0.056)	(0.12)
Dad Has Some Schooling	0.080***	0.083
	(0.026)	(0.12)
PCA Wealth Index	-0.029	.
	(0.022)	.
Age	-0.0049	0.0066
	(0.023)	(0.029)
Age Squared	-0.00060	-0.0011
	(0.0012)	(0.0015)
Female	0.032	0.0016
	(0.022)	(0.030)
Observations	3346	3346

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors clustered at village level.

The implications of this tendency for understanding the government-private test-score gap is clear: if parents are choosing to send their more academically-inclined children to private schools, and if parents have more information about student quality than researchers are able to measure in surveys and control for statistically, then standard analyses are likely to systematically overstate the quality of private school educations.

3.2 Selection in Heterogeneous Villages

While the tendency for parents to invest in "winners" rather than distribute resources uniformly is a consistent tendency in the LEAPS data, it is not the only factor that shapes school choice. In more caste-heterogeneous villages, the *social* composition of schools becomes increasing salient determinant. As shown below, as villages become more diverse, "high status" *biraderis* become more likely to send their children to private schools, creating socially segregated schools where sorting is based on social factors in addition to perceived academic potential.

To illustrate this, it is necessary to first grouping *biraderis* into "high" and "low" social status groupings allows for a better understanding of segregation patterns. The crudeness of these categorizations is unfortunate, but necessary – although *biraderis* are associated with strict hierarchies within villages, there does not exist an explicit global hierarchy of *biraderis* in Pakistan as in with the more familiar *varna* caste designations India. As a result, these hierarchies may vary somewhat from village to village, and as noted previously, this variation may not perfectly follow economic position.

To estimate the social status of different *biraderis*, Punjabi Pakistanis recruited on *oDesk.com* were asked to classify *biraderis* as having either "high" or "low" social status. Details of classifications can be found in Appendix B.[4]

Using these classifications, Table 3 examines how school choice varies with student social standing and village composition. The results show that in villages with higher caste fractionalization, a larger share of private school students come from higher status *biraderis* and a larger share of government school students come from low *biraderis*. Private schools, in other words, become reservoirs of the social elite.

Table 3: Student Body Social Composition

	(1) Pct of Students High Status	(2) Pct of Students High Status
Private School	-0.097**	-0.11*
	(0.045)	(0.060)
Biraderi Fractionalization	-0.080**	-0.025*
	(0.037)	(0.014)
Fractionalization * Private	0.19**	0.22**
	(0.083)	(0.11)
Median Village Expenditure	-0.0000032	
	(0.0000027)	
Village: Pct Adults Literate	-0.00011	
	(0.00024)	
Log Village Size	-0.0058	
	(0.0050)	
Village: Pct High Status	1.03***	
	(0.024)	
District Fixed Effects	Yes	No
Village Fixed Effects	No	Yes
Observations	772	772

Standard errors in parentheses

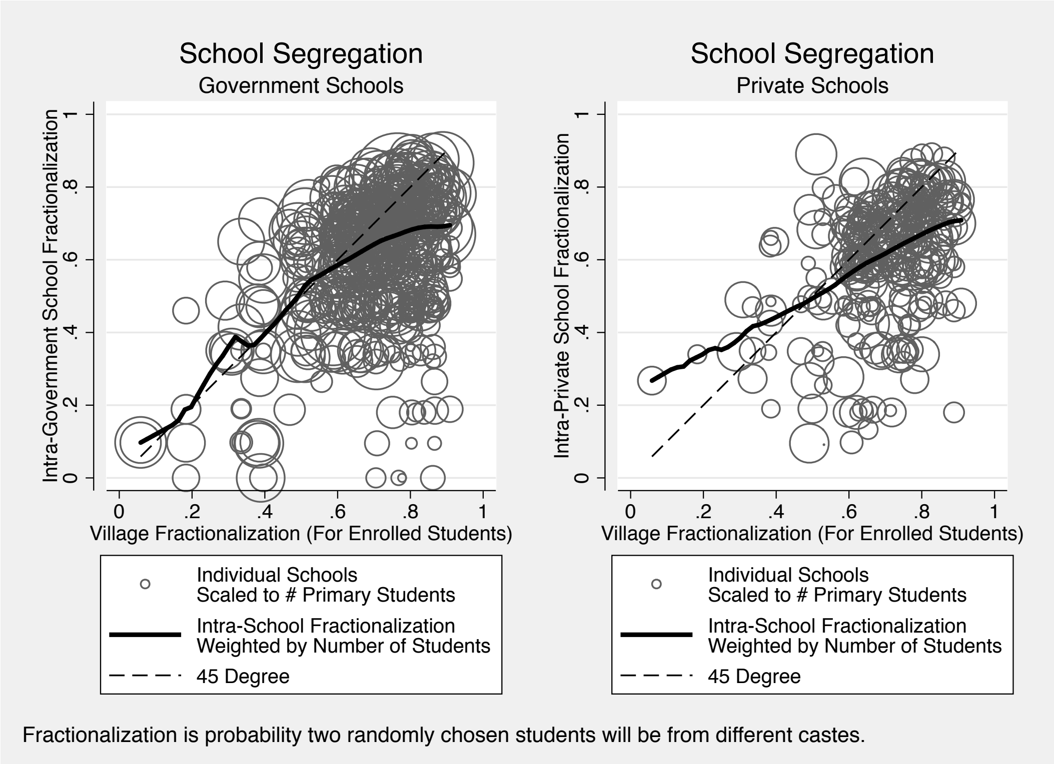
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors clustered at village level. Weighted by number of students.

Further evidence of social segregation in schooling is provided by Figure 2, which plots the relationship between a village-level fragmentation index and intra-school fragmentation indices. If schools were unsegregated, then we would expected the herfindahl indices computed *within* each school to track closely with herfindahl indices computed at the village level. Yet as shown in Figure 2, this is far from the case. Almost all schools are below the 45 degree line that would indicate school and village diversity moving one for one, and many are well below.

[4]This work has avoided the Jacoby and Mansuri (2011) methodology – where castes are ranked on the basis of their land holding – due to input from numerous sources that social standing and land holding are not equivalent, and in this exercise *social*-status is of substantially more importance than *socio-economic* status.

Figure 2: School Versus Village Fragmentation



Demand for caste segregation is also manifest in the dramatically higher prices charged by segregated private schools. As shown in Table 4 below, moving from a perfectly non-fractionalized village to a perfectly fractionalized village is associated with an average 392 Rupees increase in annual school fees. Given that the median annual fee for all private schools in the LEAPS survey is 1130 Rupees, this is a very significant amount.[5]

Table 4: Annual Private School Fees

	(1) Weighted by School	(2) Weighted by School	(3) Weighted by Primary Students
Biraderi	375.7	390.2	409.2
Fractionalization	(251.5)	(257.7)	(307.5)
Village: Median		51.8	34.0
Expenditures		(65.5)	(58.9)
Expenditure Gini		87.3	95.2
		(224.2)	(244.1)
District Fixed Effects	Yes	Yes	Yes
Observations	296	296	295

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Further, as shown in Table 5, none of these changes are driven by a change in the share of students

[5]Fees above the 95th percentile – 2750 Rupees – were adjusted down to 2750 Rupees. Without this adjustment, the coefficient on village fractionalization is even larger.

in private schools. The percentage of students in private schools remains quite stable, even when controlling for numerous village characteristics.

Table 5: Share of Enrolled Students in Private Schools

	(1) Share Students in Private School	(2) Share Students in Private School
Biraderi Fractionalization	0.077	0.085
	(0.060)	(0.032)
Median Village Expenditure		0.000029**
		(0.0000040)
Village Land Gini		0.022
		(0.087)
Village: Pct Adults Literate		0.0019
		(0.0019)
Log Num HHs		0.019
		(0.022)
District Fixed Effects	Yes	Yes
Observations	112	112

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results clustered at district level.

4 School Sorting and Test Scores

Having established that the determinants of school choice vary dramatically between caste-homogeneous and caste-heterogeneous villages in Section 3, this section now turns to an analysis of how differences in sorting impact the test-score gap between government and private schools. If private schools outperform government schools primarily due to differences in the quality of instruction, then the government-private test-score gap should be relatively stable across villages with different caste compositions. If, however, private schools appear to outperform government schools primarily due to differences in the composition of their students, then villages that are subject to different sorting processes should also see differences in the government-private test-score gap. Specifically, the degree to which private schools scores exceed those in government schools should decline as one moves from caste-homogeneous villages (where students sort on academic potential) to caste-heterogeneous villages (where sorting is driven by considerations of caste politics).

4.1 Measuring Learning

To measure learning, this analysis employs a lagged-value-added model. Lagged-value-added models have increasingly become the norm in the education research (Gordon et al., 2006; McCaffrey et al., 2003; Hanushek, 2003) due to their potential to take into account not only observational differences between students, but also the potential to control for some unobserved differences, and the fact that learning is not entirely persistent (things learned in the past are often forgotten).

The lagged-value-added model incorporates the assumption that current knowledge is an additive function of all current and past inputs and an i.i.d. stochastic error term, and can be expressed as:

$$Y_{i,t} = X_{i,t}\alpha + Y_{i,t-1}\beta + \epsilon_{i,t} \tag{1}$$

where $Y_{i,t}$ is child i's test scores at time t and $X_{t,i}$ is a vector of child, school, and village controls at time t (a full discussion of the lagged-value-added model, specifications employed, and identifying assumptions can be found in Appendix A).

Note that while the inclusion of a lagged dependent variable effectively controls for unobserved heterogeneity that affect differences in test score *levels*, it cannot control for unobserved heterogeneity that affects learning *rates*. It is for this reason that while superior to other available methods, value-added analyses can not fully overcome selection issues.[6] In the lagged-value-added model, coefficients on independent variables are interpreted as the contribution of each variable to learning.

4.2 Convergence in Government-Private Test Scores

Table 6 presents lagged-valued-added estimates of learning as a function of various demographic controls and village fractionalization. It shows that the effect of caste fractionalization on the government-private test-score gap (row 2) is negative and significant for English and negative (albeit insignificant) for math and Urdu. Further, as shown in columns (2), (5), and (8) of Table 6, the inclusion of various demographic controls such as a child wealth index and dummies for parental education along with the village fixed effects has no significant effect on the results.

To aid in interpretation, Figure 3 plots the government-private performance differential as a function of caste fractionalization (these plots correspond to columns (2), (5), and (8) respectively). In all three cases, the rise in fractionalization is associated with a near 50% decline in the private school premium, although this is by far most striking in the case of English.

The differential impact of village fragmentation on English and other subjects is not surprising. English is generally considered the path to upward mobility in Pakistan, and is often the focus of private schools in Punjab. For example, while only 6% of government schools use English as one of their languages of instruction, this is the case in 28% of private schools (Andrabi et al., 2007, p. 49). Indeed, a consistent pattern in the data is that – possibly as a result of this specialization – the English test-score gap is consistently the largest among the three subjects tested.

4.3 Decomposition of Convergence

Further evidence that the convergence in government-private test scores is driven differences in student sorting – not differences in actual learning outcomes – comes from the fact that while the government-private test-gap decreases, overall learning remains relatively unchanged across

[6] Some analysis have turned to second-differencing the data and focusing on students who change schools (Andrabi et al., 2011), but these analyses have their own limitations, among them limited sample sizes (given that changes between types of school are relatively infrequent in most surveys) and the assumption that school changes are not the result of some unobserved shock (i.e. that school switches are not accompanied by contemporaneous with other changes – a potentially problematic assumption given the relative infrequency with which students change schools).

Table 6: Child Test Scores

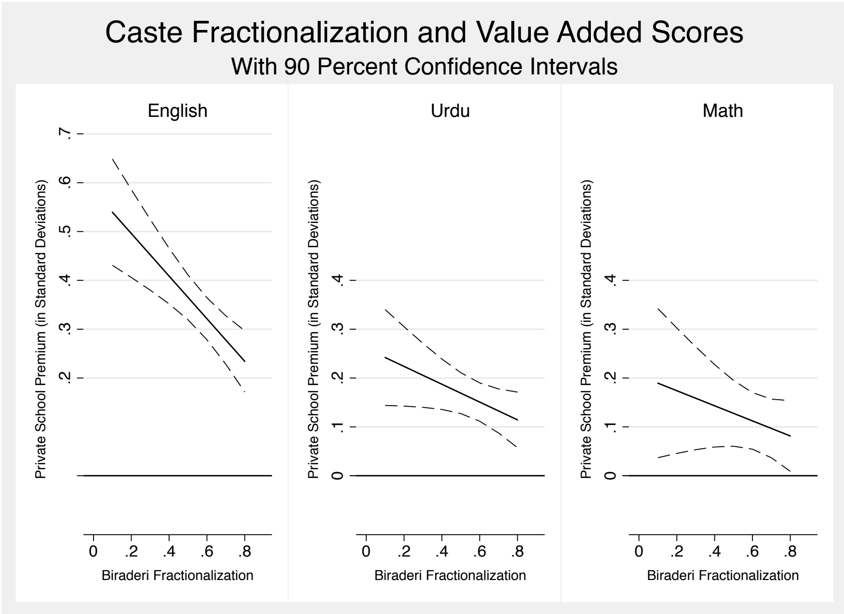
	English		Urdu		Math	
	(1)	(2)	(3)	(4)	(5)	(6)
Private School	0.62***	0.58***	0.27***	0.26***	0.21*	0.20*
	(0.072)	(0.078)	(0.073)	(0.070)	(0.12)	(0.11)
Fractionalization * Private	-0.46***	-0.44***	-0.18	-0.18	-0.12	-0.15
	(0.12)	(0.13)	(0.12)	(0.11)	(0.18)	(0.16)
Lagged English Scores	0.37***	0.36***	0.15***	0.14***	0.16***	0.15***
	(0.018)	(0.018)	(0.012)	(0.012)	(0.015)	(0.015)
Lagged Math Scores	0.069***	0.072***	0.12***	0.12***	0.37***	0.38***
	(0.0080)	(0.0080)	(0.0083)	(0.0086)	(0.013)	(0.014)
Lagged Urdu Scores	0.15***	0.15***	0.38***	0.38***	0.23***	0.22***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.013)	(0.013)
Child's Wealth Index		0.017***		0.0073**		0.014***
		(0.0033)		(0.0030)		(0.0040)
Educated Parent		0.057***		0.052***		0.045***
		(0.012)		(0.011)		(0.014)
Constant	0.25	0.40*	0.55**	0.69***	0.12	0.31
	(0.26)	(0.22)	(0.22)	(0.25)	(0.34)	(0.32)
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37147	26141	37147	26141	37147	26141

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Controls for age, age squared, gender, and class omitted from table. Standard errors clustered at village level.

Figure 3: Private School Test Score Premium with Lagged Scores



villages. As shown in Table 7 – which examines the relationship between overall test scores and village heterogeneity – test scores do not vary with caste composition. English scores are slightly higher in more fractionalized villages in Column 1, but the magnitude of this difference is relatively small, and once more demographic controls are added in Column 2 this effect disappears. No relationship exists for other subjects. Government scores increase and private scores decline with fractionalization, in other words, but those changes are almost perfectly offsetting.

4.4 Caste and Residual Academic Potential

For it to be the case that sorting by caste reduces the degree to which private schools enroll disproportionately academically-inclined students, residual academic potential – potential that cannot be explained by things like parental education and wealth – must be equally distributed across different castes (or be distributed slightly in favor of lower status *biraderis*). If not, and even the least talented "high status" students were more talented than the most talented "low status" students, then the concentration of "high status" students in private schools would result in *divergence*, rather than *convergence*, of test scores. As shown in Table 8, however, there is no evidence that those from higher social status *biraderis* have higher residual talent than those from low status *biraderis*. As evident from the top row of coefficients, after controlling for other observational factors, student caste does not appear to have any consistent effect on test scores.

5 Alternate Explanations

Observational studies pointing to the apparently superior performance of private schools are not the only reason that education reform advocates have voiced excitement about the potential of private schools; private schools also operate in fundamentally different ways from government

Table 7: Child Test Scores and Fractionalization

	English		Urdu		Math	
	(1)	(2)	(3)	(4)	(5)	(6)
Private School	0.31***	0.29***	0.14***	0.14***	0.11***	0.088**
	(0.028)	(0.028)	(0.025)	(0.024)	(0.035)	(0.034)
Biraderi Fractionalization	0.14*	0.10	0.080	0.057	0.13	0.12
	(0.075)	(0.073)	(0.066)	(0.064)	(0.094)	(0.089)
Lagged English Scores	0.40***	0.39***	0.16***	0.14***	0.17***	0.16***
	(0.018)	(0.018)	(0.012)	(0.012)	(0.016)	(0.016)
Lagged Math Scores	0.067***	0.070***	0.12***	0.12***	0.39***	0.40***
	(0.0082)	(0.0083)	(0.0083)	(0.0086)	(0.013)	(0.014)
Lagged Urdu Scores	0.15***	0.15***	0.39***	0.40***	0.23***	0.22***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.013)	(0.013)
Village: Pct Adults Literate	0.00058	0.00017	-0.00027	-0.00060	0.00031	0.00025
	(0.00097)	(0.00098)	(0.00086)	(0.00085)	(0.0014)	(0.0013)
Log Number of Households	0.017	0.019	0.012	0.015	0.0083	0.011
	(0.013)	(0.013)	(0.015)	(0.014)	(0.019)	(0.017)
Village Land Gini	-0.0043	0.036	0.0079	0.057	-0.29**	-0.26*
	(0.13)	(0.12)	(0.11)	(0.098)	(0.13)	(0.13)
Child's Wealth Index		0.015***		0.0068**		0.016***
		(0.0035)		(0.0030)		(0.0042)
Educated Parent		0.053***		0.049***		0.043***
		(0.013)		(0.011)		(0.014)
Constant	0.45	0.64**	0.68***	0.81***	0.34	0.62*
	(0.28)	(0.25)	(0.25)	(0.28)	(0.37)	(0.35)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37147	26141	37147	26141	37147	26141

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Controls for age, age squared, gender, and class omitted from table. Standard errors clustered at village level.

Table 8: Child Social Status and Residual Talent

	English			Urdu			Math		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High Status Zaat	-0.040	-0.072	-0.11**	-0.037	-0.10	-0.081	0.0014	-0.025	0.037
	(0.045)	(0.070)	(0.046)	(0.058)	(0.071)	(0.052)	(0.066)	(0.085)	(0.067)
Private School	0.32***	0.22*	0.21*	0.28**	0.26*	0.14	-0.037	-0.060	-0.11
	(0.12)	(0.13)	(0.12)	(0.12)	(0.14)	(0.13)	(0.11)	(0.14)	(0.15)
Fractionalization * Private	-0.040	0.081	0.048	-0.21	-0.19	-0.067	0.28	0.33	0.29
	(0.20)	(0.21)	(0.20)	(0.19)	(0.23)	(0.20)	(0.19)	(0.24)	(0.24)
Lagged English Scores	0.32***	0.31***	0.41***	0.16***	0.16***	0.16***	0.18***	0.19***	0.18***
	(0.042)	(0.046)	(0.040)	(0.031)	(0.036)	(0.029)	(0.045)	(0.046)	(0.040)
Lagged Math Scores	0.060**	0.048	0.045	0.11***	0.076*	0.089**	0.30***	0.27***	0.37***
	(0.030)	(0.035)	(0.032)	(0.038)	(0.042)	(0.040)	(0.050)	(0.050)	(0.048)
Lagged Urdu Scores	0.17***	0.19***	0.17***	0.34***	0.35***	0.41***	0.25***	0.26***	0.26***
	(0.038)	(0.046)	(0.042)	(0.039)	(0.045)	(0.046)	(0.046)	(0.053)	(0.052)
Child's Wealth Index		-0.0067	-0.011		0.0020	0.0010		-0.014	-0.0094
		(0.013)	(0.012)		(0.010)	(0.0090)		(0.017)	(0.014)
Educated Parent		0.14***	0.14***		0.14***	0.14***		0.17***	0.18***
		(0.043)	(0.039)		(0.044)	(0.035)		(0.055)	(0.046)
Biraderi Fractionalization			-0.029			-0.063			-0.042
			(0.10)			(0.13)			(0.17)
Village: Pct Adults Literate			0.00037			-0.0032**			-0.0013
			(0.0018)			(0.0016)			(0.0027)
Log Number of Households			0.021			0.025			0.050
			(0.036)			(0.033)			(0.061)
Village Land Gini			0.027			0.11			-0.15
			(0.18)			(0.21)			(0.32)
Constant	0.57	0.71	0.75	1.18*	2.35***	2.22***	1.57**	2.83***	2.19**
	(0.56)	(0.71)	(0.60)	(0.64)	(0.74)	(0.67)	(0.74)	(0.92)	(0.94)
Village Fixed Effects	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
District Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes
Observations	1790	1323	1323	1790	1323	1323	1790	1323	1323

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Controls for age, age squared, gender, and class omitted from table.

Standard errors clustered at village level.

schools. For that reason, it is important to examine not only whether the government-private performance differential varies with caste heterogeneity, but also whether other factors some suspect explain school differences can be ruled out as alternate explanations.

5.1 Differences in Teacher Incentives

Advocates of private schools argue that not only are observational studies able to control for many factors, but there is also evidence to explain *why* private schools outperform government schools. In particular, they point out that private schools appear to address the biggest problem in government schools: low effort. High absenteeism and low accountability in government schools has been well documented (Muralidharan and Kremer, 2008; Chaudhury et al., 2006), but appears less prevalent in private schools. The reason, many argue, is that in private schools good teachers are better paid, and poor teachers are let go. This line of reasoning is also buoyed by a growing body of literature that suggests that what matters for success is not the availability of educational "inputs" (like qualified, well paid teachers or good facilities), but incentive schemes that reward effort on the behalf of teachers (Hanushek, 1997, 2003; Banerjee et al., 2007).

The conclusion of the LEAPS survey authors (full disclosure: this author was a research assistant on the LEAPS project, though not an author) is that this is the case in Pakistan. Private schools deliver better educational outcomes despite hiring only secondary-educated local women with no training and providing them with relatively low wages because private schools incentize good teaching by paying good teachers more, improving effort. This differentiates them from government schools, which offer salaries which are higher but unresponsive to performance. As a result, the authors argue, government school teachers exert less effort (Andrabi et al., 2007, p. 72-74).

Tables 9 and 10 reiterate these findings from the original LEAPS survey and adds a measure of caste heterogeneity. If it is the case that differences in incentive schemes are driving test score convergence, then we should see (a) the private-school salary penalty for absenteeism decline in fractionalization in Table 9, and (b) the compensation premium (compensation above what is otherwise predicted) for performance decline with fractionalization in Table 10. As shown in the tables, however, the absenteeism penalty actually *increases* in fractionalization , and no relationship exists between performance and compensation. In sum, variation in compensation schemes cannot explain the convergence in the test-score gap in heterogeneous villages.

5.2 Differences in School Inputs

A second potential explanation for variation in the government-private test-score gap is that the availability of school resources varies with village caste heterogeneity. To test this, this paper examines the relationship between village fragmentation and different school inputs.

Table 11 presents a series of teacher-level and school-level regressions in which various inputs are regressed against village fractionalization and controls. The table shows that there is little or no evidence that – at least in terms of measurable inputs – the decline in private school performance in fractionalized villages can be explained by differences in inputs. These results show no difference in school inputs across villages.

Table 12 repeats this exercise for government schools. It should be noted that government schools

Table 9: Teacher Compensation and Absenteeism

	Government Teachers		Private Teachers	
	(1)	(2)	(3)	(4)
	Log Salary	Log Salary	Log Salary	Log Salary
Days Absent Last Month	0.017***	0.014	-0.0060*	0.0094
	(0.0039)	(0.010)	(0.0033)	(0.012)
Biraderi Fractionalization		-0.0028		0.30
		(0.13)		(0.20)
Days Absent * Fractionalization		0.0055		-0.023
		(0.018)		(0.017)
District Fixed Effects	Yes	Yes	Yes	Yes
Observations	3685	3685	4638	4638

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Errors clustered at village-level

Table 10: Teacher Compensation and Value-Added

	Government Teachers		Private Teachers	
	(1)	(2)	(3)	(4)
	Log Salary	Log Salary	Log Salary	Log Salary
Average Value Added Score	0.025	0.18	0.14**	-0.14
	(0.028)	(0.11)	(0.069)	(0.24)
Biraderi Fractionalization		-0.24*		0.47
		(0.14)		(0.30)
Value-Added * Fractionalization		-0.25		0.43
		(0.15)		(0.34)
District Fixed Effects	Yes	Yes	Yes	Yes
Observations	1232	1232	745	745

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Teacher value-added estimates control for student age, age squared, wealth index, parental education, and class. Teachers with less than 5 students dropped from analysis. Errors clustered at village-level

Table 11: Private Teacher Characteristics and Village Fractionalization

	(1) Days Absent Last Month	(2) Female	(3) Local Teacher	(4) More than Grade School Educ	(5) Teacher's English Score	(6) Basic School Facility Index
Biraderi Fractionalization	-0.13	-0.15	0.21	0.11	0.30	0.45
	(0.34)	(0.096)	(0.16)	(0.080)	(0.19)	(0.68)
Median Village Expenditures	0.0000047	0.000032	0.0000038	0.0000096	0.000038	-0.000081
	(0.000061)	(0.000021)	(0.000019)	(0.000016)	(0.000024)	(0.000096)
Log Number of Households	0.061	-0.0067	-0.072*	-0.030	0.038	-0.24*
	(0.10)	(0.022)	(0.039)	(0.020)	(0.033)	(0.13)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4667	4675	4670	4675	1041	295

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All results clustered at the village level.

in Pakistan are administered at the state level, and are thus relatively insulated from village politics, making any such differences unlikely. Nevertheless, results are presented for thoroughness. If anything, government school teachers appear to be slightly *better* educated in fractionalized villages.

6 Discussion and Policy Implications

Private schools represent a radical departure from the educational status quo in nearly every way. Where government schools hire teachers with college educations, emphasize teacher training, and pay high wages, private schools hire secondary-educated women from the local community for a fraction of the cost, invest nothing in teacher training, and appear to reward performance. If government schools were to emulate their business model, tax payers would save millions. But would educational outcomes improve?

The results presented here cannot conclusively answer this question. Even in the most heterogeneous villages, private schools still outperform government schools (just by dramatically less than in homogeneous villages). But these results do suggest at least two major reasons for caution on the part of policy-makers thinking about the implications of private school growth, and whether recent calls for the widespread distribution of private school vouchers (e.g. Chakrabarti and Peterson (2008), Kelkar (2006), and Panagariya (2008)) are good policy.

First, this analysis shows that even with a lagged-value-added specification, parental education and wealth controls, and panel data, it is still easy to vastly over-estimate the contributions of private schools to learning using observational data. As such, this study illustrates the need for researchers and policy-makers to maintain a health degree of skepticism when reading observational studies that claim to separate sorting from school quality.

Second, while this analysis only shows that 50% of the private school premium can be explained by sorting, it is important to remember that this represents a lower-bound on the contributions of sorting. The difference between homogeneous and heterogeneous villages is not "sorting on intelligence" and "no sorting on intelligence," but rather "sorting on intelligence" and "*less* sorting on intelligence." As such, sorting is still likely contributing to the private school premium in heterogeneous villages. Indeed, perceived intelligence remains an important determinant of school choice even in highly fractionalized villages.

In light of persistent poor performance among government schools, the hope that private schools will reform the South Asian education sector is understandable, and may yet prove to be well founded. But as shown here the superiority of private school is not self-evident, and the government would do well to gather more evidence before embracing private schools as a substitute for government schools.

Table 12: Government Teacher Characteristics and Village Fractionalization

	(1) Days Absent Last Month	(2) Female	(3) Local Teacher	(4) More than Grade School Educ	(5) Teacher's English Score	(6) Basic School Facility Index
Biraderi Fractionalization	-0.028	0.035	0.18*	0.11	0.38**	-0.46
	(0.42)	(0.084)	(0.097)	(0.069)	(0.15)	(0.46)
Median Village Expenditures	0.000013	-0.00000071	0.0000068	0.000012	0.000080***	0.000045
	(0.000054)	(0.000013)	(0.000011)	(0.000010)	(0.000027)	(0.000081)
Log Number of Households	-0.15	-0.050***	-0.055**	0.0067	-0.012	-0.25
	(0.11)	(0.017)	(0.025)	(0.017)	(0.044)	(0.15)
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3687	3691	3689	3691	1170	496

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All results clustered at the village level.

References

Alderman, H., Kim, J., and Orazem, P. F. (2003). Design, Evaluation, and Sustainability of Private Schools for the Poor: The Pakistan Urban and Rural Fellowship School Experiments. *Economics of Education Review*, 22(3):265–74.

Alderman, H., Orazem, P. F., and Paterno, E. M. (2001). School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan. *The Journal of Human Resources*, 36(2):304–326.

Andrabi, T., Das, J., Khwaja, A., Vishwanath, T., and Zajonc, T. (2007). *Pakistan: Learnings and Educational Achievements in Punjab Schools*. The World Bank.

Andrabi, T., Das, J., Khwaja, A. I., and Zajonc, T. (2011). Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics*, 3(3):29–54.

Angrist, J. D., Bettinger, E., Bloom, E., King, E., and Kremer, M. (2002). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review*, 92(5):1535–1558.

Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics*, 122(3):1235–1264.

Bellei, C. (2008). The Private-Public School Controverersy: The Case of Chile. In Chakrabarti, R. and Peterson, P. E., editors, *School Choice International: Exploring Public-Private Partnerships.*, page 272. The MIT Press, Cambridge, Massachusetts.

Chakrabarti, R. and Peterson, P. E. (2008). *School Choice International: Exploring Public-Private Partnerships.* MIT Press, Cambridge, Massachusetts.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., and Rogers, F. H. (2006). Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1):91–116.

Chay, K. Y., McEwan, P. J., and Urquiola, M. (2005). The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools. *American Economic Review*, 95(4):1237–1258.

Currie, J. and Thomas, D. (1995). Does Head Start Make a Difference? *American Economic Review*, 85(3):341–364.

Desai, S., Dubey, A., Vanneman, R., and Banerji, R. (2009). Private schooling in india: A new educational landscape. *India Policy Forum*.

Gazdar, H. and Mohmand, S. K. (2007). Social Structure in Rural Pakistan.

Glewwe, P., Ilias, N., and Kremer, M. (2010). Teacher Incentives. *American Economic Journal: Applied Economics*, 2(3):205–227.

Gordon, R., Kane, T. J., and Staiger, D. O. (2006). Identifying Effective Teachers Using Performance on the Job. The Hamilton Project Policy Brief No. 2006-01. *Brookings Institution*.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, 19(2):141–164.

Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies. *The Economic Journal*, 113(485):F64–F98.

Jacoby, H. and Mansuri, G. (2011). Crossing boundaries: gender, caste and schooling in rural Pakistan. *World Bank Policy Research Working Paper No. 5710*.

Jimenez, E. and Lockheed, M. E. (1995). Public and Private Secondary Education in Developing Countries: A Comparative Study. *World Bank Discussion Papers*.

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

Kane, T. J. and Staiger, D. O. (2002). The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, 16(4):91–114.

Kelkar, V. (2006). Let Every Parent be a Consumer. *India Today*.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., and Hamilton, L. (2003). *Evaluating value-added models for teacher accountability*. Rand Corp.

Muralidharan, K. and Kremer, M. (2008). Public and Private Schools in Rural India. In Chakrabarti, R. and Peterson, P. E., editors, *School Choice International: Exploring Public-Private Partnerships*. MIT Press.

Panagariya, A. (2008). *India: The Mergeing Giant*. Oxford University Press, New York.

Pratham (2005). *Annual Status of Education Report*. Pratham Documentation Center, New Delhi.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics*, 125(1):175–214.

Tooley, J. and Dixon, P. (2003). *Private schools for the poor: A case study from India*. Centre for British Teachers, Reading, UK.

A Value-Added Test Scores

The lagged-value-added model incorporates the assumption that current knowledge is an additive function of all current and past inputs and an i.i.d. stochastic error term. This can be written formally as:

$$Y_{i,t} = \alpha_t X_{i,t} + \alpha_{t-1} X_{i,t-1} + \cdots + \alpha_1 X_{i,1} + \epsilon_{i,t} \tag{2}$$

where $Y_{i,t}$ is child i's test scores at time t and $X_{t,i}$ is a vector of child, school, and village controls at time t. As data on all past inputs are usually unavailable, however, they are generally subsumed into a lagged dependent variable included as a control. In this case the lagged-value-added model can be re-written as:

$$Y_{i,t} = X_{i,t}\alpha + Y_{i,t-1}\beta + \epsilon_{i,t} \tag{3}$$

Where $Y_{i,t-1}$ is assumed to capture all past inputs and unobservable heterogeneity across students. The specification given by Equation 3 is the model primarily employed in this paper.

The primary interest of this analysis is on the differential "value-added" by government and private schools. This parameter is the coefficient on a dummy variable in $X_{i,t}$ for whether the child attends a private school, where positive values suggest a higher "value-added" (a positive test-score gap) between private and government schools.

Value-added estimates for teachers – used in in Section 5 – are generated by adding dummies for each teacher to Equation 3. One dummy variable is added for each teacher, and the coefficient associated with each teacher's dummy is their "value-added."

Three aspects of the lagged-value-added specification are worth emphasizing. First, while the inclusion of a lagged dependent variable effectively controls for unobserved differences that affect differences in test *levels*, it cannot control for unobserved heterogeneity that affects learning *rates*. It is for this reason that while superior to other available methods, value-added analyses can not fully overcome selection issues.[7]

Second, the β term can be interpreted as the "persistence parameter" in that it estimates the degree to which past learning may carry forward. A value of one is equivalent to assuming that children do not forget past lessons, while a value of zero corresponds to students forgetting all past lessons each year. While imposing a persistence parameter of one may seem reasonable – it amounts to regressing the difference in test scores between time t and time $t-1$ on controls – a growing literature has shown that the test score gains of short term interventions often "die out" over time, suggesting that not all that is learned is retained (Banerjee et al., 2007; Glewwe et al., 2010; Currie and Thomas, 1995; Rothstein, 2010). Andrabi et al. (2011) shows that imposing the restriction that $\beta = 1$ biases learning estimates.

Finally, lagged test scores are generally measured with error, which leads to an often significant attenuation bias in the estimate of β and biased estimates of other coefficients (Kane and Staiger, 2002; Chay et al., 2005; Andrabi et al., 2011). Thus keeping with best practices, lagged test scores from all three subjects – English, Urdu, and math – are included in all specifications to instruments for the primary lagged test score of interest.

[7] Some analysis have turned to second-differencing the data and focusing on students who change schools (Andrabi et al., 2011), but these analyses have their own limitations, among them limited sample sizes (given that changes between types of school are relatively infrequent in most surveys) and the assumption that school changes are not the result of some unobserved shock (i.e. that school switches are not accompanied by contemporaneous with other changes – a potentially problematic assumption given the relative infrequency with which students change schools).

Table 13: Social Status Classifications

Biraderi	Status	Notes
Abbasi	High*	“Section, both of [Syeds] and [Sheikhs]” (?, p. 341)
Ansari	Low*	“[Sheikh] section”
Arain	High*	
Awan	High*	
Baloch	Low*	
Butt	High*	
Charchar	???	“Kharwar” are among “Non-functional castes of low position” (?, p. 106
Dhobi / Naich / Mochi / Lohar	???	Professional castes: laundry washers (Dhobi), blacksmiths (Lohar), cobb
Gujjar	High	“Non-functional castes of respectable position” (?, p. 106)
Jat	High	“Non-functional castes of respectable position” (?, p. 106)
Kharar	???	
Lar	???	
Mohana	???	
Mughal	Low*	
Muslim Sheikh	Low*	
Non-Muslim	???	
Pathan	High	“High caste” (?, p. 268)
Qureshi / Hashmi	Low*	“[Sheikh] section” (?, p. 353)
Rajput / Bhatti	High	“High caste” (?, p. 268); Warrior and land owning caste (?, p. 353)
Rehmani	Low*	
Samejha	???	
Sheikh	High	“High caste” (?, p. 268)
Solangi	Low*	
Syed	High	“High caste” (?, p. 268)

B Biraderi Classification

Table 13 shows classifications and sources.

Land holdings are sufficient but necessary condition.

Note: Mohana in 15_make_hh_child_cross_section;

Kharar & Charchar duplicates? (?, p. 344) says Charchar may be Charghar – 4 houses. But doesn’t say what it is besides caste subdivision.

(?, p. 10) defines a “Section” or “Sept.” as the largest exogamous group within a “subcaste”, which he defines as the smalled endogamous groups within a caste.

Working with a local anthropologist, we constructed a caste-status identifier that categorizes dozens of distinctly named caste/clan (zaat/biradari) groups into high and “low” caste. High-caste includes all such groups that self-identify on the basis of traditional access to land (zamindars). The low-caste group comprises zaats that were historically considered either out-castes (similar to the dalits in India) or were in clientalist relationships with zamindars as providers of services in the village economy; i.e. barbers, metalworkers, clothes washers, etc. Based on this definition, around 25% of the population from which we draw our sample consists of low-caste households, with the highest proportion (35%) found in Sindh province.