

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 (????????). 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 these private schools are actually delivering superior educations, or whether they just attract students who are more academically inclined or come from families that prioritize educational attainment.

This paper leverages variation across villages in the factors that drive school choice to provide new insight into this question. In particular, it takes advantage of the fact that school choice in rural Pakistan is motivated by different factors in caste-homogeneous and caste-heterogeneous villages. This paper shows that in caste-diverse villages, school choice is driven by social considerations – high-status families send their children to private schools to keep them in homogeneous social settings. In caste-homogeneous villages where caste sorting is unnecessary, by contrast, school choice is driven by parental perceptions of academic potential – families send their more academically gifted children to private schools.

To take advantage of this variation, this analysis proceeds in three steps. First, this analysis estimates the test-score gap 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 (???), and is able to control not only for observable differences in child demographics, but also some non-observable differences.<sup>1</sup> This constitutes a baseline estimate of the government-private school performance differential using non-experimental data.

This analysis then compares estimates of the government-private performance differential in homogeneous villages with the performance differential 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 performance differential between 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 amount they outperform government schools falls by half when moving from homogeneous villages to heterogeneous villages. This implies that at least half of the 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 performance differential. As shown in Section ??, for example, caste fragmentation 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 per-

---

<sup>1</sup>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 ??.

formance differential – like performance pay in private schools or differences in school resources – do not appear to vary systematically with village heterogeneity ??.

Second, at the level of villages there is no evidence that overall learning outcomes vary with village caste composition. As detailed in Section ??, the decrease in government-private school performance differential is the result of convergence between the two school types, consistent with a re-distribution of students 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 in by things like parental education and wealth – must be equally distributed across different castes (or be distributed slightly in favor of lower status *biraderis*). As shown in Section ??, 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. (?, 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 (?). 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 public-private test score gap. This should give analysts pause when examining other empirical results that claim to fully control for sorting.<sup>2</sup>

---

<sup>2</sup>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. ? 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 conducted of a voucher system in Chile, but as ? 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.

This paper is organized out as follows: Section ?? provides an overview of the rural Pakistan context from which data for this analysis is drawn. Section ?? details how the determinants of school choice vary with village caste heterogeneity. Section ?? 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 ?? then examines and rules out a number of alternative possible explanations for this empirical regularity. And finally, Section ?? discusses the strengths and weaknesses of these findings and their interpretation.