RISE Working Paper 22/091
February 2022

The Value of Private Schools: Evidence from Pakistan

Pedro Carneiro, Jishnu Das, Hugo Reis

Abstract

Using unique data from Pakistan, we estimate a model of demand for differentiated products in 112 rural education markets with significant choice among public and private schools. Families are willing to pay substantially for reductions in distance to school, but in contrast, price elasticities are low. Using the demand estimates, we show that the existence of a low fee private school market is of great value for households in our sample, reaching 2 percent to 7 percent of annual per capita expenditure for those choosing private schools.

JEL Classifications: Differentiated Goods, Demand, Education, School Choice, Pakistan

Keywords: I20, I21

The Value of Private Schools: Evidence from Pakistan

Pedro Carneiro
UCL, IFS, CEMMAP, FAIR-NHH
p.carneiro@ucl.ac.uk

Jishnu Das
Georgetown University
jishnu.das@georgetown.edu

Hugo Reis
Banco de Portugal and Católica Lisbon School
hfreis@bportugal.pt

Acknowledgements:

We thank Phillipe Haile, Lars Nesheim, Petra Todd, Ariel Pakes, Orazio Attanasio, Jeffrey Hammer, Lant Pritchett, Abhijeet Singh, and seminar participants at a CSAE-Oxford seminar, the IZA summer school, the SOLE and the RES conferences, and the NBER/BREAD meetings for valuable comments. We also thank the editor and 3 referees for extensive comments on previous drafts that have substantially improved the paper. Pedro Carneiro gratefully acknowledges the financial support from the Economic and Social Research Council (grant reference RES-589-28-0001) through the Centre for Microdata Methods and Practice, and the support of the European Research Council through ERC-2015-CoG-682349. Hugo Reis acknowledges the financial support of Fundação para a Ciência e Tecnologia through the project number PTDC/EGE-ECO/6055/2020. Jishnu Das thanks the RISE programme for financial support through RISE grant P013412.

This is one of a series of working papers from “RISE”—the large-scale education systems research programme supported by funding from the United Kingdom’s Foreign, Commonwealth and Development Office (FCDO), the Australian Government’s Department of Foreign Affairs and Trade (DFAT), and the Bill and Melinda Gates Foundation. The Programme is managed and implemented through a partnership between Oxford Policy Management and the Blavatnik School of Government at the University of Oxford.

Please cite this paper as:
Carneiro, P., Das, J. and Reis, H. 2022. The Value of Private Schools: Evidence from Pakistan. RISE Working Paper Series. 22/091. https://doi.org/10.35489/BSG-RISE-WP_2022/091

1 Introduction

Rising private school enrollments in low income countries have prompted a range of government responses, from active support through subsidies and partnership arrangements, to onerous regulation, sometimes at the same time.[1] The lack of a coherent response reflects, in part, a limited understanding of how households make schooling choices and how educational markets function in low-income countries. This is an area where approaches from industrial organization (IO) can play a central role, as long as estimation methods developed for product markets can be extended to education. In particular, market boundaries may not be clear, objective functions (for both consumers and firms) can be hard to define and, critical data such as the costs of running a school may not be available.

Our goal here is to assess how a careful understanding of the demand for private schools can be used to inform policy in low-income settings. In order to do so, we use data from the Learning and Education in Pakistan Schools project developed by Andrabi et al. (2007). These data are from 112 villages in Pakistan, where each village is a different education market with an average of 7 public and private schools, allowing us to delineate markets clearly.[2] Private schools are minimally regulated and did not receive public subsidies at the time of data

[1]Private school shares in low-income countries increased from 11% in 1990 to 22% in 2010; in Pakistan it was 39% in 2015 (Baum et al. (2014)). Their market share reflects both low prices and tests scores that are similar or higher compared to public schools (Muralidharan and Sundararaman (2015), Andrabi et al. (2020), and Singh (2015)). In response to the growth of private schools, the Government of Punjab, in Pakistan, provides vouchers for students (Barrera-Osorio et al. (2020)) and has recently outsourced management of some public schools to private organizations (Crawfurd (2018)). At the same time, regulators and the Supreme Court have ordered a cap on school fee increases, potentially limiting investment in these schools (PakistanToday (2019)).

[2]This simplifies the issues that arise when markets are not as clearly defined, or when school nominations are affected by strategic considerations due to assignment mechanisms

collection, and the data include specialized surveys in both schools and households. At the time of data collection, prices in the private sector therefore reflected conditions in the local market; public schools were, and continue to be, free at the point of use. Parents could choose among all schools as long as they could afford the fees of the school they chose.

Using these data, we first estimate models of demand for differentiated products adapted to education markets, accounting for the endogeneity of both school fees and peer attributes (Berry et al. (1995), Berry et al. (2004) and Bayer and Timmins (2007)). We then assess the robustness of our models to alternate specifications and assess the plausibility of our estimated price elasticity–a key component of our model–using a voucher experiment that we implemented in these villages. Finally, we conduct counterfactual experiments to demonstrate the value of this exercise for policy.

Our demand model shows, first, that a central determinant of school choice in this setting is the distance to school. The average distance between home and school (for those enrolled) is 510 meters for girls and 680 meters for boys. A 500 meter increase in distance decreases the likelihood that a school is chosen by 11.1 percentage points for girls, and 6.0 percentage points for boys. For boys, parents are willing to pay more than a full year of private school fees of \$13 for a 500 meter reduction in distance, while for girls this value reaches 74% of annual school fees. These estimates mirror the experimental findings of Burde and Linden (2013) on the importance of distance in similar settings.

Second, own-price elasticities of -1.12 for girls and -0.37 for boys, are low. These reflect the change in demand when a single school increases its price; sectoral price elasticities, which reflect the increase in demand from a reduction in the price of *all* private schools are -0.27 for girls and -0.10 for boys. The low sectoral price elasticities run counter to the belief that prices are the main barrier to private schooling in low-income countries. Therefore, we returned to the same households 14 years later and offered a one-year price discount for children of school-going age if they attended private schools in the village, varying the price discounts

(see Burgess et al. (2015)).

experimentally. Although the experimental and structural estimates are not strictly comparable due to the length of the price discount, we find surprisingly and similarly low price elasticities in the experiment as well. Interestingly, our estimates are also consistent with those reported previously, by Dynarski et al. (2009) who report an elasticity of -0.19 and Arcidiacono et al. (2021), who report elasticities between -0.47 and -0.65.

Third, parents value other school characteristics, notably the test scores of peers and school infrastructure, but their value is lower than that placed on distance. For instance, they are willing to pay 13% to 25% of a full year of private school fees for extra facilities and, they are willing to pay 12% to 31% of average annual tuition in a private school, for a 1 standard deviation increase in the test scores of their peers.

Using our estimates we conduct two counterfactual exercises. Motivated by the literature on the demand for new goods, we first estimate the value of private schools for this population, and to a large extent, the value of school choice. Put bluntly, if, as discussed in Hausman (1996) and Bresnahan and Gordon (1997), the arrival of Apple-Cinnamon Cheerios with a share of 1.6% in the market for cereals nevertheless added $78 million in consumer welfare each year, what is the added welfare from private schools with a market share of 39%? We show that for the set of students choosing private schools, the value of private schooling is USD$3.4 for girls and USD$ 11.0 for boys, which corresponds to 2% and 7% of their total annual per capita expenditure.[3] If we consider full universe of students, which includes those choosing public schools or no school at all, these figures are lower (USD$1.4 and USD$4.8 or 1% and 4% of annual per capita expenditure). Extrapolating our estimates from rural Punjab to the entire country, the total value of private schools in Pakistan is estimated to have been at least $138 million in 2003.[4]

[3]Interestingly, 78% of the value of having private schools comes just from the ability to opt-out from the public option. The benefit from having an expanded choice set of private schools offering many differentiated products is much smaller (Hausman (1996)).

[4]The value of private schooling is likely higher in urban areas where private school fees were 70% higher in 2001 than in rural areas (Andrabi et al. (2002)). Scaling-up valuations with

Second, we examine the potential impact of vouchers, which we simulate as a reduction to zero in the price of attending any private school. Such a voucher would cost $13 for each student who uses it, and would increase private school enrollment for girls from 19% to 40% and for boys from 23% to 31%. Since most children never use the voucher, the implied per capita cost of a voucher in the whole population is $5.2 for girls and $4.0 for boys, relative to a valuation of $2.7 and $2.4 respectively. In addition, there is a further reduction in society's direct costs of schooling of $3.3 for girls and $1.4 for boys, resulting from a shift of children from public schools (where costs per student are higher) to private schools. As the implementation of a voucher system at scale would have to pay for children who would have chosen private schooling even without a subsidy, the difference between the cost and the welfare gain provides one estimate of what the shadow value of market failures (such as credit constraints) must be for such schemes to increase welfare.

This paper therefore contributes to a literature on the IO of education markets. Gallego and Hernando (2009), Neilson (2021), Barrera-Osorio et al. (2020), Bau (2021) and Arcidiacono et al. (2021) all estimate variants of a model of demand for differentiated goods applied to education markets. They then use these estimates to examine the link between voucher prices and test scores (Neilson (2021)), horizontal differentiation in instructional levels among private schools (Bau (2021)) , and input choices among private school owners compared a social planner (Barrera-Osorio et al. (2020)).[5] In a recent paper, Arcidiacono et al. (2021) estimate

---

fees would imply that boys choosing private schools in urban Punjab value that choice at $19.0 and girls at $6.1.

[5]Examples from the U.S. include Bayer et al. (2007), who estimate residential choice models and Hastings et al. (2009), who estimate the impact of providing school-level information on test scores on school choice. Dinerstein and Smith (2018) estimate the impact of increased funding for public schools on private school exit and entry in New York and Pathak and Shi (2017) evaluate the performance of structural demand estimates against a change in school allocation mechanisms in Boston.

a demand model (which also incorporates liquidity constraints) to data from a voucher experiment in India and use it to calculate the welfare gains from a voucher program.

Our first contribution to this literature is methodological. The data we use allows us to better account for the endogeneity of school prices, assess the sensitivity of the model to peer effects, control for unobserved household characteristics, and compare experimental and structural estimates of the price elasticity of demand. Our estimates are robust across multiple validation exercises and thus provide support for the continued use of IO models in education markets.

Our second contribution is to show how demand estimates can affect policy, which is especially interesting in settings where culture and social norms can affect school choice (Ashraf et al. (2020) and Borker (2020)). In Chile (Gallego and Hernando (2009), Neilson (2021)), children from poorer households are highly sensitive to price, but are unwilling to travel far. This allows schools in poorer areas to markdown quality. In our setting, restrictions on female mobility that differ by social status implies that children from poorer households are willing to travel *farther* to go to school relative to children from richer households (Jacoby and Mansuri (2015) and Cheema et al. (2018)). Consequently, distance to school affects girls schooling choices more than that of boys, and that of richer girls more than that of poorer girls, with different implications for the market power of schools in poorer areas. As the effects of policies differ by geography and cultural norms, ex-ante simulations specific to the area where a policy is implemented will have very high value.

Our simulation of the impact of vouchers is one example. Punjab introduced vouchers for private schools in 2008 and by 2018, there were 2.5 million beneficiaries. Yet, household survey data does not show a marked change in the proportion of children enrolled in private schools. While the demand for private schooling may have declined due to other factors, a second possibility supported by our estimates is that price was never the fundamental barrier to private schooling and vouchers were primarily a fully fungible (and regressive) income subsidy for children already enrolled in the private sector. Ex-ante simulations would have provided valuable information towards the design of a better voucher scheme, potentially encouraging the entry of new schools in areas where there were none, rather than subsidies for already existing schools, as discussed by Barrera-Osorio et al. (2020).

Our third contribution takes advantage of the fact that prices in our setting are market-determined and therefore we can value school attributes in dollar terms. This allows us to use our demand estimates to compute welfare metrics, such as the value of private schooling. Although such welfare computations are standard in the literature on products, they have not been used for the education market in low-income countries, where outcome-based measures of welfare are more typical. The outcome-based approach would miss that private schools add value not only through test scores but also the utility benefits of shorter commute times as well as other amenities that are directly valued by parents. Our extension of the literature on new goods to education offers a potential option for evaluating the benefits of government programs in the market for schooling.

In the remainder of this paper, we develop these ideas further. Section 2 presents the Data. Section 3 describes the econometric model used to study the determinants of parents choices among different schools. Section 4 presents the estimates from the model and Section 5 provides the results from the simulations. In Section 6 we contrast price elasticity estimates from a voucher experiment with those produced by the model. Finally, Section 7 concludes.

2 Data

We use the first wave of data from the Learning and Education Achievement in Pakistan Schools (LEAPS) project, collected in 2003/04. The LEAPS data were collected from 112 villages in the Punjab province, randomly chosen from those with at least one private school in 2000; in 2003, the majority of the province's rural population lived in such villages. At the time of the first wave, private schools in these villages faced virtually no de-facto regulation and did not receive subsidies from the government or other bodies. (Andrabi et al. (2017)) Therefore, the prices and attributes that they chose reflect market demand and costs.

The LEAPS project administered surveys to both households and schools, in addition to testing students in Mathematics, English and the vernacular, Urdu. The household survey includes information on household demographics, expenditure data, and school attendance by children in the household. The schools attended are separately identified for each child, allowing us to link household and school attributes. The school survey has information on school characteristics including teacher characteristics (sex, education, experience and performance in Mathematics, English and Urdu tests), basic and extra school facilities, and school costs. These include teacher salaries, the cost of utilities, school materials, and other items. We also construct the characteristics of the student body of each school, namely test scores, parental education, and household assets for the average student in the school. Finally, all households and schools were geo-located allowing us to construct the distance from each household's place of residence to each school in the village.

Table 1 (panel A) reports individual and household characteristics for children between 5 and 15 years old in the sample, distinguishing between boys and girls. Each variable is described in the online Appendix Table A.1. There are 2244 girls and 2317 boys in the sample. On average children in the sample are 9.8 years old, their mothers have 1.3 years of education and the average per capita annual expenditure is $121.2. There are no differences in the characteristics of families of boys and girls. However, girls attend schools closer to their residence and are also less likely to attend school than boys in general (see also Reis (2020)).

Table 1 (panel B) shows means and standard deviations of school-level variables, each described in the online Appendix Table A.1. We present one column for all schools in the sample, one for public, and one for private schools. In addition, because we separate our analyses for boys and girls, and because not all schools are attended by children of both genders we also distinguish schools depending on whether they enter the boys or the girls' analysis (with some schools entering both). There are 511 schools attended by girls and 522 schools attended by boys.

Private schools are more likely to be coeducational and report better infrastructure, with more toilets, and extra facilities such as gyms, libraries or computer labs. More than 80%

Table 1: Summary statistics

Panel A - Individual and Household characteristics	Girls		Boys	
	mean	st. dev.	mean	st. dev.
Age (years)	9.9	(3.1)	9.7	(2.8)
Mothers Education (years)	1.4	(2.7)	1.3	(2.7)
Expenditure per capita	118.9	(127.3)	123.8	(168.7)
Household distance to facilities (Kms)	1.23	(2.96)	1.24	(2.86)
Distance to current school (Kms)	0.51	(0.63)	0.68	(0.88)
Distance to all schools (Kms)	1.09	(1.11)	1.25	(1.34)
Attending school (%)	66.8		79.8	
Attending private school (% of attending school)	28.0		28.7	
Number of Children	2244		2317	
Number of Households	1242		1292	

Panel B- School Characteristics	Total		Public		Private	
	Girls	Boys	Girls	Boys	Girls	Boys
Private School (%)	53.6	50.8	-	-	-	-
School fees	-	-	-	-	13.3	13.1
					(9.4)	(9.0)
School with toilets	0.85	0.74	0.73	0.52	0.95	0.95
	(0.36)	(0.44)	(0.44)	(0.50)	(0.22)	(0.22)
School with permanent classroom	0.87	0.86	0.91	0.88	0.84	0.85
	(0.33)	(0.34)	(0.28)	(0.32)	(0.37)	(0.36)
Number of extra facilities	3.0	2.7	2.1	1.7	3.7	3.7
	(1.6)	(1.7)	(1.4)	(1.5)	(1.2)	(1.2)
Percentage of female teachers	0.82	0.44	0.87	0.09	0.77	0.78
	(0.31)	(0.44)	(0.34)	(0.28)	(0.28)	(0.28)
Perc. of teachers with at least 3 years of exp.	0.61	0.62	0.87	0.84	0.39	0.40
	(0.35)	(0.34)	(0.24)	(0.24)	(0.27)	(0.26)
Perc. of teachers with university degree	0.25	0.31	0.32	0.42	0.20	0.20
	(0.25)	(0.27)	(0.30)	(0.30)	(0.19)	(0.19)
Teacher absenteeism	2.0	1.8	3.0	2.6	1.1	1.2
	(3.7)	(2.9)	(4.7)	(3.4)	(2.0)	(2.1)
Teacher test score (average)	0.86	0.87	0.86	0.88	0.86	0.86
	(0.09)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)
Student test score (average)	0.36	0.35	0.29	0.27	0.42	0.42
	(0.13)	(0.13)	(0.11)	(0.11)	(0.11)	(0.11)
Perc. of Mother with some education (sch. level)	0.27	0.24	0.18	0.12	0.36	0.36
	(0.27)	(0.26)	(0.21)	(0.16)	(0.29)	(0.29)
Asset index (sch. level)	-0.35	-0.59	-0.79	-1.23	0.04	0.03
	(1.05)	(1.14)	(1.02)	(0.99)	(0.92)	(0.91)
Pay and Allowance of Teaching staff (Annual Exp.)	2252.1	2504.9	3432.1	3826.3	1231.4	1223.4
	(2494.5)	(3134.9)	(3010.1)	(3867.2)	(1240.0)	(1242.8)
Number of students	155.1	167.4	163.3	189.2	148.1	146.2
	(120.7)	(139.1)	(139.7)	(166.3)	(101.3)	(102.1)
Number of Schools	511	522	237	257	274	265

Notes: Means and the standard deviations of children and their household attributes (Panel A) and school characteristics (Panel B). In panel B, the standard deviation is in brackets. Each variable is described in Table A.1 in the online Appendix. School fees and Annual Expenditure in US dollars. 1 US dollar = 85.6 Pakistani Rupees.

of the schools have permanent classrooms, and almost all have a blackboard. Public schools do not charge tuition while private schools charge an average annual tuition of \$13 per year, which is 11% of annual per capita expenditure. Student test scores (with a mean of 0.35 and a standard deviation of 0.13 in the sample) are 1 standard deviation higher in private compared to public schools. Teachers in public schools are more educated and experienced than teachers in private schools, but report higher absenteeism. Teacher test scores are about the same in both types of schools. Furthermore, the proportion of mothers who have ever attended any school is much higher for students in private schools, as are their household assets.[6] Finally, the annual expenditure on pay and allowance of teaching and non-teaching staff is higher for public schools, while the costs of utilities and educational materials is higher for private schools.

Tables A.2 and A.3 are analogous to Table 1 (panel B) in the online Appendix, showing characteristics of schools attended by boys and girls, but distinguishing families with different levels of maternal education, household expenditure, and average distance between each household and other important (e.g., health and administrative) facilities, which are often located in the center of the village. Strikingly, there is little variation by family background in the average tuition levels of girls attending private schools, although the proportion of girls attending any school and attending private school vary by maternal education, family expenditure, and household average distance to facilities. These patterns are similar for boys, with the difference that average private school tuition for those attending private school is negatively related to household expenditure. Again, this is counterbalanced by the fact that both the proportion of boys attending any school and the proportion of boys attending private school greatly increases with household expenditure.

There are some, but not substantial, differences between the infrastructure of schools attended by children with different family backgrounds. Some teacher characteristics (such as

[6]We observe family expenditure in the household survey, which we use to construct family background characteristics, but not in the school census. The school census only allows us to construct a simple measure of wealth, which we use as a school attribute.

education and experience) are worse for children in more affluent households, perhaps reflecting the fact that they attend mostly private schools, where teachers are less educated and less experienced on average. Average test scores of peers in the school are not very different in schools attended by rich and poor children. This is true even though the average levels of assets and maternal education in the school differ dramatically across schools attended by children with different family backgrounds. Finally, for both boys and girls, children of richer families attend schools that are closer to their residence than children of poorer families. This is different from Chile, where Neilson (2021) shows that richer households are willing to travel farther.

We also note that there is substantial cross village variation in the proportion of children in school, varying from 49% to 100% for boys (with a mean of 82%), and from 19% to 96% for girls (with a mean of 69%). Similarly, among those in school, the proportion of boys in a private institution can vary from 3% to 72% (with a mean of 29%), while for girls this variation is from 3% to 100% (with a mean of 30%).

3 Empirical model

We model the demand for schools following the literature on the demand for differentiated products and a recent literature on neighborhood choice. We adapt the procedures proposed in Berry et al. (1995), Berry et al. (2004), and Bayer and Timmins (2007) to the particular characteristics of our problem and dataset, defining the village as the relevant education market for each household, and estimating different models for boys and girls. This is consistent with the data in our sample, where students do not attend primary schools outside their village of residence.

In each village there are several schools with different attributes. A household chooses a single school among those in her market, and derives utility from its attributes. The utility household i obtains from its child (of gender g) attending school j in village/market t is given by

$$u_{ijtg} = \sum_{k=1}^{K} x_{jktg}\beta_{ikg} + \gamma_{ig} d_{ijtg} + \xi_{jtg} + \varepsilon_{ijtg} \tag{1}$$

where $j = \{0, ..., J\}$ indexes each school competing in a market defined by t. The outside option, corresponding to no enrollment in any school, is represented by $j = 0$. Therefore, u_{i0tg} is the utility of individual i if he does not attend any of the J schools in the village; k indexes observed school characteristics (x_{jktg}) which are valued differently by each individual and ξ_{jtg} is an unobserved school attribute valued equally by everyone. Here, d_{ijtg} is the distance from the house of household i to school j (and represents the role of geography, as in Bayer and Timmins (2007)). Finally, ε_{ijtg} is an individual-specific preference for school j in market t, which is assumed to be independent and to have an extreme value type I distribution.

Let r indicate a specific observed household characteristic, z_{irtg}, and let v_{itg} be an unobserved characteristic of household i. The value of each school characteristic for each household is allowed to vary with the household's own observed and unobserved characteristics. To minimize the danger of over-fitting in the model, we interact log of household expenditure with a single school characteristic, the school fee. In particular:

$$\beta_{ikg} = \overline{\beta}_{kg} + \sum_{r=1}^{R} z_{irtg}\beta^{o}_{rkg} + \beta^{u}_{kg} v_{itg} \tag{2}$$

and

$$\gamma_{ig} = \overline{\gamma}_{g} + \sum_{r=1}^{R} z_{irtg}\gamma_{rg} + \gamma^{u}_{g} v_{itg} \tag{3}$$

In equations (2) and (3), individual preferences can be divided into three parts: $\overline{\beta}_{kg}$, which is constant within gender; β^{o}_{rkg} and γ_{rg}, which vary with observable student attributes, z_{irtg}; and β^{u}_{kg} and γ^{u}_{g}, which vary with unobservable attributes of the individual, v_{itg}.[7]

[7]We impose that v_{itg} does not vary with the k^{th} characteristic being considered (although its coefficient, β^{u}_{kg}, does vary with k). In other words, the unobserved components of the random coefficients in our model are driven by a single factor: v_{itg}. This assumption simplifies our

Integrating (2) and (3) into (1) we get

$$u_{ijtg} = \sum_{k=1}^{K} x_{jktg}\overline{\beta}_{kg} + \xi_{jtg} + \sum_{k=1}^{K}\sum_{r=1}^{R} x_{jktg}z_{irtg}\beta^{o}_{rkg} +$$
$$+ \sum_{k=1}^{K} x_{jktg}v_{itg}\beta^{u}_{kg} + \overline{\gamma}_g d_{ijtg} + \sum_{r=1}^{R} d_{ijtg}z_{irtg}\gamma_{rg} + \gamma^{u}_{g}d_{ijtg}v_{itg} + \varepsilon_{ijtg} \quad (4)$$

Household i chooses the school for a child of gender g to maximize (4). We can further rewrite this equation as:

$$u_{ijtg} = \delta_{jtg} + \sum_{k=1}^{K}\sum_{r=1}^{R} x_{jktg}z_{irtg}\beta^{o}_{rkg} + \sum_{k=1}^{K} x_{jktg}v_{itg}\beta^{u}_{kg} +$$
$$+\overline{\gamma}_g d_{ijtg} + \sum_{r=1}^{R} d_{ijtg}z_{irtg}\gamma_{rg} + d_{ijtg}v_{itg}\gamma^{u}_{g} + \varepsilon_{ijtg} \quad (5)$$

with

$$\delta_{jtg} = \sum_{k=1}^{K} x_{jktg}\overline{\beta}_{kg} + \xi_{jtg}. \quad (6)$$

The coefficients of this model can be estimated using the algorithms described in Berry et al. (1995) and Berry et al. (2004) (under standard assumptions on v_{itg} and ε_{ijtg}, discussed in online Appendix B) and in Bayer and Timmins (2007), which we adapt to our data. As in these papers, we proceed in two steps.

The first step estimates $\delta_{jtg}, \beta^{o}_{rkg}, \beta^{u}_{kg}, \overline{\gamma}_g, \gamma_{rg}, \gamma^{u}_{g}$ by maximum likelihood, including a contraction mapping to obtain δ_{jtg}. This is a hybrid of the procedures proposed in Berry et al. (1995), and Berry et al. (2004). Although we use micro data, and in principle we should be able to estimate all the parameters of the model by maximum likelihood, we do not observe enough households per school to reliably estimate school fixed effects δ_{jtg} (for most schools we do not observe much more than 10 children in the household survey). However, since we

estimation by reducing the number of unobservables over which we need to integrate. It is also reasonable to think that these random coefficients are driven by a low dimensional set of unobservables, so that considering a single unobservable may not be a poor approximation.

also have a household level census detailing school choices in each village, it is possible to reliably estimate market shares, and recover δ_{jtg} using the contraction mapping procedure proposed in Berry et al. (1995). Apart from this detail, the way we implement these procedures is standard in the literature. See online Appendix B.

The second step estimates $\overline{\beta}_{kg}$, are obtained by running a regression of the school fixed effect (δ_{jtg}) on the observed school characteristics, as in equation (6). δ_{0tg}, which reflects the outside option, is normalized to zero. The household and school variables used to estimate the model are described in online Appendix A - Table A.1. At the school level (x_{jktg}), we use almost every variable available in the dataset, including an indicator variable for whether a school is private. At the individual/household level (z_{irtg}), to minimize the computational burden of our procedure we focus on four variables that are important determinants of educational choices: age of the children, maternal education, (log) of expenditure (which in our setting is a better measure of permanent income), and average household distance to other facilities in the village (capturing the distance to the village center). Finally, we allow for a single household unobservable, v_{itg}, to affect the coefficients on all observable school attributes. Unlike the BLP approach, we do not model the supply-side, a choice that we discuss in in Sections 4 and 5.

As is well understood, prices and other product attributes could be endogenously chosen and observable product attributes could be correlated with unobserved product attributes. In our data, a rich set of school characteristics together with village fixed-effects explain 70% of the total variance of school fixed-effects. Nevertheless, there is still the possibility that school characteristics are missing from the data. One option is to not interpret the coefficients as the households' valuation of the corresponding attributes and consider them instead as coefficients of a projection of all school characteristics on the set of characteristics we observe. This is a standard approach, typically used for all attributes with the exception of price (for which instrumental variables are used), since price plays a particularly important role in most demand models, and it is important to have a credible estimate of the impact of price on demand. In addition to price, in this paper we also consider the potential endogeneity of distance to school and peer quality. We next discuss how prices and distance are addressed in the estimation, postponing a discussion of endogenous peer effects till Section 4.5.4.

3.1 The endogeneity of prices

Our main results instrument price with teachers' costs in the *tehsil*, a group of 100-200 villages, leaving out the own-village in the computations (similar to the instrument used by Bau (2021) and Arcidiacono et al. (2021)). Specifically, we use geographic variation in teacher costs as a cost-shifter. The key assumption is that any one village is too small to change prices in other villages in the tehsil, but villages in the same tehsil are likely to have the same systematic differences in teacher labor supply.

We augment this instrument with total school costs, the number of other schools within 2Km and observed non-price attributes of other competitors as proposed by Berry et al. (1995).[8] The additional "BLP-style" instruments capture how crowded a product is in characteristic space, which should affect the price-cost margin and the substitutability across products. The instruments are justified by assuming that they do not affect the choice of unobserved school attributes, conditional on the observed attributes we include in the model. Our final specification interacts this leave-one out estimate with an indicator variable for whether a school is private, while controlling for cost and a private indicator separately (and the full set of interactions of the private school dummy with other school attributes). We assess the robustness of our estimates using different cost components and Hausman-style instruments. [9]

[8]We exclude rent payments for schools renting their buildings, since there is no available data on user costs for schools that own their buildings.

[9]Our model is a special case of Berry and Haile (2009), who discuss the non-parametric identification of multinomial choice demand models with heterogeneous individuals. Under standard large support and instrumental variables assumptions, they show identifiability of the random utility model.

3.2 Distance to School

Substantial observational and experimental evidence shows that distance to school is a powerful determinant of school attendance, so we devote particular attention to this variable (e.g. Burde and Linden (2013) and Alderman et al. (2001)). The main concern is that households living in the center of the village are generally richer and may also be different in unobserved ways to households living elsewhere. Since private schools tend to locate near the center of villages, these households will also have greater access to private schools, creating a correlation between distance to school and unobserved household characteristics. In order to address this issue we include in the model the average distance between each household and other important facilities in the village, such as for example, hospitals and health clinics, which are also located in the center of the village as well. This follows Andrabi et al. (2020) who demonstrate the validity of this approach in their work on the causal estimates of the impact of private schooling on test-scores, and is justified with recourse to the historical settlement patterns in these villages.

4 Estimates from the model

We consider a mix of household (z_{irtg}) and school variables (x_{jktg}) in the model. The valuation of school characteristics is allowed to vary with both observed and unobserved household characteristics (z_{irtg} and v_{itg}), which means that we can entertain a very rich set of substitution patterns in the data. Our benchmark model does not explicitly consider the endogeneity of peer attributes, which are the average test scores, maternal education, and household assets of other students in the school. We return to this in Section 4.5, where we consider models with endogenous peers, and more generally, discuss the robustness of our estimates to alternate model specifications. In addition, given the large number of parameters in our model one could be worried about overfitting. Therefore, in order to reduce the number of parameters, for our benchmark model we estimated a specification that excluded from the model the interactions of (log) expenditure (which in our setting is a better measure of permanent income) with the school characteristics other than fees, since sensitivity to fees probably varies with income (and perhaps this is not as clear for the other school attributes we consider unless income is capturing unobserved preference heterogeneity). However we show in the appendix that these restrictions do not substantially affect our main results.

4.1 Estimation procedure

We estimate equation (5) using maximum likelihood, with an additional step to estimate the school fixed-effect (as described above and in online Appendix B). The estimated coefficients are shown in tables A.4, and A.5, in online Appendix A. The coefficients in equation (6) can be estimated using instrumental variables, although we also present OLS estimates for comparison. The results for the first stage regressions are displayed in table A.6. Since distance to school is not a fixed school attribute, but depends on each household's location, the coefficients related to distance are estimated in the initial maximum likelihood procedure (see also Bayer and Timmins (2007)).

4.2 Parental willingness to pay for school attributes

Our main results are shown in Tables 2, 3 and 4. Table 2 first shows the estimated coefficients for equation (6) using different specifications for girls and boys; Columns 1 and 4 are OLS results, Columns 2 and 5 our preferred IV estimates and Columns 3 and 6 correspond to IV estimates using an alternative set of instruments, the total cost without rent at the village level using the costs of the other villages in the same sub-district (tehsil). Since we allow the valuation of school attributes to depend on whether the school is public or private, we report the average of the public and private coefficients for each characteristic. Similarly, below we calculate the average willingness to pay for each characteristic, averaged across public and private schools.

Tables 3 (girls) and 4 (boys) then combine the estimated coefficients in equations (5) and (6). Columns 1 to 3 show the impact of each school characteristic on parental utility and

Columns 5 to 7 report the willingness to pay (WTP) for changes in these school characteristics, at the 25th percentile, the mean, and 75th percentile of the joint distribution of maternal education and household assets (labeled 25th, Mean, and 75th).[10] The magnitude of the changes considered in the WTP calculations vary across variables, because each variable has a different scale. The size of the relevant change for each variable is reported in column 4; for example, 0.10 in column 4 for the proportion of female teachers indicates that in columns 5-7 we compute the WTP for a 10 percentage point increase in the proportion of female teachers in the school.

There are three noteworthy patterns. The first is that parents place considerable value on distance and on price for both boys and girls. We discuss these estimates in detail below. Second, parents are willing to pay \$1.7/\$3.3 for an extra facility for girls/boys, and \$1.6/\$4.0 for girls/boys for a one standard deviation increase in test scores, the latter significant at the 90% level of confidence. Third, other school attributes are valued differently for boys and girls. Parents of boys strongly dislike schools with more female teachers and are willing to pay \$2.0 for a 10 percentage point reduction in the proportion of female teachers. In contrast, parents of girls are willing to pay \$0.5 for a 10 percentage point increase in the proportion of female teachers. Finally, girl's parents are willing to pay \$0.6 for a 10 percentage point *reduction* in the proportion of students whose mothers have at least some education. When interpreting this, one should note that the vast majority of mothers in these villages have little or no education. Since the regression already controls for the average test score of peers, one explanation for our results is that conditional on the average test performance of other students, the average mother may prefer to sort into schools with similar mothers, as opposed to schools with very different

[10]We compute WTP for an attribute by dividing the corresponding coefficient by the coefficient on fees, which in this model also measures the marginal utility of income. We then multiply this fraction by the number in the 4th column of the table, generating columns 5, 6, and 7. Coefficients vary across households because of household observed and unobserved variables.

Table 2: OLS vs. IV regressions

	Girls			Boys		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	IV	OLS	IV	IV
School fees	-0.023*	-0.136***	-0.135***	0.022*	-0.043*	-0.042*
	[0.014]	[0.041]	[0.040]	[0.013]	[0.025]	[0.026]
School with toilets	0.031	0.122	0.121	0.220	0.280	0.279
	[0.375]	[0.361]	[0.371]	[0.232]	[0.237]	[0.228]
School with permanent	0.137	0.225	0.224	0.144	0.182	0.182
classroom	[0.274]	[0.266]	[0.264]	[0.201]	[0.204]	[0.198]
Number of extra facilities	0.131*	0.198***	0.197***	0.091	0.122**	0.122**
	[0.070]	[0.072]	[0.071]	[0.056]	[0.056]	[0.055]
Perc. of female teachers	0.831***	0.592*	0.594*	-0.611**	-0.747***	-0.745***
	[0.316]	[0.336]	[0.338]	[0.273]	[0.275]	[0.277]
Perc. of teachers with	0.399	0.251	0.252	0.269	0.186	0.187
at least 3 years of exp.	[0.316]	[0.321]	[0.340]	[0.266]	[0.283]	[0.276]
Perc. of teachers with	0.111	0.530	0.525	-0.148	0.112	0.108
university degree	[0.400]	[0.449]	[0.427]	[0.311]	[0.330]	[0.325]
Student test score	0.571	1.443*	1.434*	0.654	1.146*	1.139*
	[0.708]	[0.777]	[0.755]	[0.625]	[0.640]	[0.643]
Teacher absenteeism	0.039	0.036	0.036	0.004	0.003	0.003
	[0.042]	[0.042]	[0.041]	[0.024]	[0.023]	[0.024]
Teacher test score	1.090	1.600	1.595	0.761	1.051	1.047
	[1.092]	[1.073]	[1.023]	[0.726]	[0.709]	[0.710]
Perc. of mother with	-0.539	-0.711**	-0.710**	-0.300	-0.381	-0.379
some education	[0.353]	[0.363]	[0.339]	[0.305]	[0.307]	[0.309]
Asset index	-0.133	-0.056	-0.057	-0.015	0.030	0.030
	[0.094]	[0.099]	[0.099]	[0.077]	[0.074]	[0.077]
Private	-1.270***	-0.254	-0.264	-1.569***	-0.911**	-0.920**
	[0.321]	[0.494]	[0.475]	[0.348]	[0.396]	[0.411]
F-Test (instruments)						
All schools	-	10.15	10.17	-	15.93	15.76
p-values	-	0.0000	0.0000	-	0.0000	0.0000

Notes: This table shows the estimated coefficients for equation (6) for girls and boys (estimation of $\overline{\beta}_{kg}$ by running a regression of the school fixed effect (δ_{jtg}) on the observed school characteristics (including interactions with private school indicator) using different specifications. The first and fourth columns show the OLS estimates, the second and the fifth columns show our main IV estimates, which includes as instruments, teachers' costs in the tehsil leaving-out the own-village, total school costs excluding rent payments, and BLP-style instruments. The remaining two columns correspond to IV estimates using the total cost excluding rent payments in the tehsil leaving-out the own-village as an alternative leave-out instrument. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Willingness to pay for school characteristics - girls

					Willingness to Pay (in U.S. dollars)		
	25th percentile	mean	75th percentile	Variable variation	25th percentile	mean	75th percentile
School fees	-0.167***	-0.136***	-0.120***				
	[0.041]	[0.041]	[0.041]				
School with toilets	0.164	0.122	0.125	1.00	1.1	1.0	1.2
	[0.379]	[0.361]	[0.370]				
School with permanent classroom	0.300	0.225	0.218	1.00	2.1	1.9	2.1
	[0.277]	[0.266]	[0.270]				
Number of extra facilities	0.178**	0.198***	0.197***	1.00	1.2	1.7	1.9
	[0.071]	[0.072]	[0.073]				
Percentage of female teachers	0.496	0.592*	0.610*	0.10	0.3	0.5	0.6
	[0.341]	[0.336]	[0.342]				
Percentage of teachers with at least 3 years of experience	0.137	0.251	0.304	0.10	0.1	0.2	0.3
	[0.319]	[0.321]	[0.343]				
Percentage of teachers with university degree	0.432	0.530	0.553	0.10	0.3	0.5	0.5
	[0.430]	[0.449]	[0.432]				
Student test score (average)	1.447*	1.443*	1.394*	0.13	1.3	1.6	1.8
	[0.759]	[0.777]	[0.783]				
Teacher absenteeism	0.036	0.036	0.039	1.00	0.3	0.3	0.4
	[0.042]	[0.042]	[0.041]				
Teacher test score (average)	1.620	1.600	1.574	0.08	0.9	1.1	1.2
	[1.028]	[1.073]	[1.096]				
Perc. of Mother with some education (school level)	-0.976***	-0.711**	-0.620*	0,10	-0,7	-0,6	-0,6
	[0.333]	[0.363]	[0.348]				
Asset index (school level)	-0.064	-0.056	-0.046	1.05	-0.5	-0.5	-0.5
	[0.099]	[0.099]	[0.098]				
Distance	-2.223***	-2.233***	-2,252***	0.50	-7.8	-9.6	-11.0
	[0.145]	[0.136]	[0.142]				
Private	-0.357	-0.254	-0.211	1,00	-2.5	-2.2	-2.1
	[0.483]	[0.494]	[0.465]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the willingness to pay for each of them, change with the family background of the girl. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile$\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 7 report the willingness to pay for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Willingness to pay for school characteristics - boys

					Willingness to Pay (in U.S. dollars)		
	25th percentile	mean	75th percentile	Variable variation	25th percentile	mean	75th percentile
School fees	-0.051**	-0.043*	-0.039				
	[0.025]	[0.025]	[0.025]				
School with toilets	0.162	0.280	0.329	1.00	3.7	7.6	9.9
	[0.238]	[0.237]	[0.251]				
School with permanent classroom	0.107	0.182	0.215	1.00	2.5	4.9	6.4
	[0.198]	[0.204]	[0.203]				
Number of extra facilities	0.115**	0.122**	0.127**	1.00	2.6	3.3	3.8
	[0.055]	[0.056]	[0.057]				
Percentage of female teachers	-0.852***	-0.747***	-0.675**	0.10	-2.0	-2.0	-2.0
	[0.269]	[0.275]	[0.271]				
Percentage of teachers with at least 3 years of experience	0.091	0.186	0.240	0.10	0.2	0.5	0.7
	[0.286]	[0.283]	[0.281]				
Percentage of teachers with university degree	0.171	0.112	0.091	0.10	0.4	0.3	0.3
	[0.329]	[0.330]	[0.310]				
Student test score (average)	0.996	1.146*	1.231*	0.13	3.0	4.0	4.8
	[0.644]	[0.640]	[0.642]				
Teacher absenteeism	0.002	0.003	0.004	1.00	0.0	0.1	0.1
	[0.024]	[0.023]	[0.024]				
Teacher test score (average)	1.104	1.051	0.992	0.09	2.3	2.6	2.7
	[0.726]	[0.709]	[0.767]				
Perc. of Mother with some education (school level)	-0.368	-0.381	-0.391	0.10	-0.8	-1.0	-1.2
	[0.298]	[0.307]	[0.303]				
Asset index (school level)	-0.009	0.030	0.052	1.14	-0.2	0.9	1.8
	[0.075]	[0.074]	[0.076]				
Distance	-1,131***	-1,151***	-1,165***	0.50	-13.0	-15.6	-17.4
	[0.089]	[0.079]	[0.084]				
Private	-0.929**	-0.911**	-0.909**	1.00	-21.3	-24.8	-27.2
	[0.405]	[0.396]	[0.387]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the willingness to pay for each of them, change with the family background of the boy. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile$\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 7 report the willingness to pay for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

(and more educated mothers). [11]

We also examine how the household's valuation of a school attribute varies with the family background of the student, restricting our discussion to the school attributes that interact significantly with observable family characteristics (Tables A.4, and A.5). For girls, the statistically significant interactions are between maternal education and school fees, maternal education and the average maternal education of other students in the school, and family expenditure and school fees. For boys, the statistically important interactions are between maternal education and the proportion of female teachers in the school, maternal education and whether schools have toilets, maternal education and the asset index of the other students in the school, and age of the children and number of extra facilities.

Columns 1 to 3 of Table 3, shows that the sensitivity of girls' enrollment to fees, average maternal education of peers, and distance to school, declines with family background. As we would expect, the own-price elasticity is significantly lower for girls from a higher family background; the coefficients in the table correspond to an elasticity of -1.41 for girls at the 25th percentile relative to -0.94 for girls from the 75th percentile. The negative valuation of the maternal education of peers could reflect social stratification in these villages. Most mothers have little or no education. Therefore, conditional on the average test performance of other students, mothers who are less educated may prefer to sort into schools with similar mothers,

[11] Following Barrera-Osorio et al. (2020) we also examined the correlation between parental preferences for different school attributes and compared this to the bundles of attributes that schools actually offer. These correlations, reported in table A.11, are restricted to attributes that were statistically significant in equation (6). For girls, preferences for school attributes are positively and strongly correlated: Parents who value one of these attributes also value all the others. For boys the patterns are irregular and the strength of the correlations is weaker. Interestingly, the correlations among the bundles of these attributes that schools actually offer is much weaker and not necessarily positive (Table A.12). We find similarly weak correlations for private schools, which suggests that some costs may be school specific

as opposed to schools with very different, more educated mothers. Given the decline in price elasticity, the WTP for changes in either distance or the family background of peers is estimated to increase with family background.[12]

Columns 1 to 3 of Table 4 shows that, like for girls, the elasticity with respect to fees for boys declines with family background (-0.45 at 25th percentile relative to -0.33 at the 75th percentile). In addition, the sensitivity of boys' enrollment to whether the school has more facilities rises with background variables, and, with regards to the proportion of female teachers in the school, declines with the family background of the student.

As these attributes are not randomly assigned across schools, these patterns are best regarded as descriptive without a causal interpretation. We also cannot make direct comparisons of the magnitudes of the coefficients across gender groups unless we assume that the variance of ε_{ijtg} in the random utility model does not vary with gender. However, we can still compute demand elasticities, which, in the following sections, we discuss in detail for two attributes, fees and distance to school, where we also argue for a causal interpretation of the estimates based on the IV specification.

4.3 School fees

Our most striking result is that the own-price elasticity of demand is well below 1 for most of the schools. The own-price elasticity is estimated to be -1.12 for girls and -0.37 for boys, which implies that if a single school increases its price by 10%, demand among girls/boys will reduce by 11%/4% . The own-price elasticity increases (in absolute value) with the level of the fee in the school, suggesting that more expensive schools price in a more elastic section of the demand

[12]While it is sensible that the negative coefficient on the maternal education of peers becomes less important as one's education increases, it does not make as much sense that (at the same time) the WTP for uneducated mothers is increasing in one's education. This result may be a consequence of our linearity assumptions, and could potentially disappear in a more flexible model.

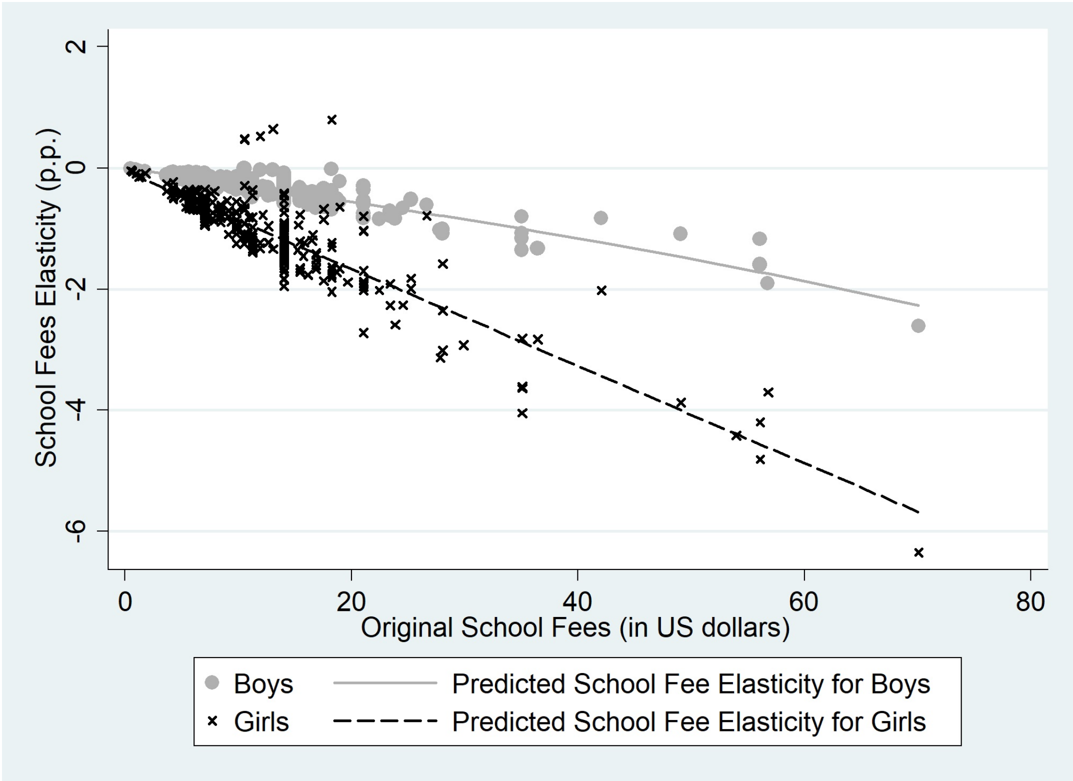
curve (Figure 1). Several additional features of the price elasticity are noteworthy. First, the sectoral price elasticity, which reflects the increase in demand when *all* schools increase prices simultaneously is lower at -0.27 for girls and -0.10 for boys. Second, online Appendix Table A.7 shows that own-price elasticities in the transitional grades (Grade 5) are higher than in non-transitional grades (Grades 3 and 4). Therefore, the elasticities we estimate are averages over different groups. One implication is that the optimal pricing strategy then needs to account for potential non-linearities in market demand as well as switching costs.

The fact that different groups have different elasticities and that elasticity changes across grades implies that schools must solve a difficult dynamic pricing problem in order to price optimally. Perhaps not surprisingly, we indeed find that a static model of profit maximization, which requires that schools never price in the inelastic portion of the demand curve, is insufficient to characterize this market: Appendix Table A.8 computes the price elasticity by school-fee quartiles, and it is only once we are in the top quartile that schools price in the elastic part of the demand curve for girls. For boys, in all parts of the distribution, schools price in the inelastic portion of the demand curve. Our inability to explain school pricing based on per-period profit maximization is an important puzzle for future research. What schools are maximizing and the dynamic nature of price elasticities have not been investigated in this literature thus far, and our assessment is that such an investigation will be necessary in order to estimate a fully specified supply-side model as in BLP (Berry et al. (1995) and Berry et al. (2004)).

4.4 Distance

Our second main result is that distance is a key determinant of school choice for both boys and girls, but more so for girls. Increasing the distance to school by 500 meters decreases the likelihood of choosing that school by 11.1 p.p for girls and 6.0 p.p for boys. Tables 3 and 4 show that parents are willing to pay \$15.6 for a 500 meters reduction in distance to school for boys (from an average distance of 680 meters to the current school, and 1250 meters to all

Figure 1: Elasticity of enrollment with respect to school fees



Notes: This figure represents the elasticity of demand with respect to fees, as a function of the original school fees, for both girls and boys. The school fee elasticity is a measure of how much the enrollment in each school changes (in percentage points) when the price increases by 1 percent. Schools not charging fees (public) are excluded from the sample.

schools in the village) and \$9.6 for girls.[13] The magnitudes of the estimates are substantial, especially when compared to the annual fee in a typical private school. Notice also that the willingness to pay for distance is higher for boys than for girls despite the fact that the elasticity of demand with respect to distance is higher for girls than for boys. This is because the demand for boys' schooling is less price elastic, and therefore parents are willing to pay more for the same reduction in distance.

Another way to highlight the importance of distance relative to other school attributes in the demand for schooling is to express WTP for each school attribute in terms of distance to school, instead of in monetary terms (computed by dividing the coefficient of each attribute of equation (6) by the coefficient on distance in the same equation). The results in online Appendix Tables A.9 and A.10 suggest that that parents are willing to travel very small additional distances in response to relatively large changes in other school attributes. For example, parents of girls are only willing to travel 90 meters more (110 meters for boys) for an additional extra facility, or 810 meters (500 meters for boys) for a \$13.3 reduction in school fees, which would make private schools free on average.

4.5 Robustness to alternate specifications

We now investigate the robustness of our estimates to alternate specifications, different instruments and the potential endogeneity of peer attributes. We report the consolidated results from multiple robustness checks in the online Appendix Tables A.14 and A.15, and include individual estimates from each specification in the online Appendix. Tables A.14 shows that estimates of the school fee elasticity, distance elasticity and the willingness to pay for distance are similar

[13]To assess potential misspecification, we estimated a model with a quadratic distance term. The quadratic term is not statistically significant for boys and significant at the 10% level for girls. With this specification, increasing the distance to a school by 500 meters decreases the likelihood of choosing that school by 8.6 p.p. for girls and by 6.2 p.p. for boys (11.1 and 6.0 p.p. in the linear specification, respectively).

between the main specification and four other specifications where we interact all demographics with all school attributes (all interactions), only allow school fees to interact with income and exclude all interactions involving distance to village health and administrative facilities (exclude some interactions), add number of children to the set of household covariates (number of children), and add a quadratic distance term to the model allowing the impacts of distance on choices to be non-linear in distance (quadratic distance term). Table A.15 shows that the estimates are robust to different sets of instruments for school fees. In the remaining of this section we present more detail on some of these robustness checks.

4.5.1 Alternate Specifications

As discussed in the paper, in order to address the potential for overfitting, we exclude some interactions between household covariates and school attributes from the model. To further investigate whether overfitting continues to pose a problem for our estimates, we estimated an alternate model that drops the interaction of school fees with maternal education, age and household distance to facilities, as well as all interactions with household distance to facilities with similar results to our main specification (Table A.16). We then estimated a second, even more parsimonious specification that reduced the total number of parameters by: (i) including a limited set of school characteristics using principal components to summarize school-level peer and facility variables; (ii) excluding all non-income interactions with school fees and; (iii) excluding all interactions with household distance to facilities. In this exercise, the AIC of our preferred specification is smaller than the one obtained from the parsimonious model for both girls (6904.0 versus 6936.8) and boys (7547.6 and 7724.0), lending further credibility to our estimates. Finally, we estimated our coefficients on a 'training' dataset that excluded 50% of villages from the estimating sample and checked if these estimates were also valid in the hold-out sample. Overall, for both, girls and boys, we achieved a close out-of-sample fit, with the predicted moments –private school enrollment shares in the aggregate and for different subgroups– similar to the moments observed in the data. These exercises are shown in the online Appendix (Table A.17).

4.5.2 Alternate Instruments

Tables A.18 and A.19 first show that our results are robust to using non-teacher costs of the school in the first stage. Since cost data are typically not available to researchers, we also estimated new first stage regressions using Hausman-style instruments. We combined private schools by size into 4 and 10 categories. For each case we calculated (i) the average prices and (ii) the median prices of the same-group schools in other villages for each school (Tables A.20 and A.21).[14] Our main takeaway is that hausman-style instruments such as these are quite weak in our setting. These additional variables are not statistically significant and do not improve the power of the first stage. Consequently, our results remain substantively unchanged.

4.5.3 Incorporating school size

A potentially important school attribute that has been excluded from the model is a measure of school size. Parents may have an intrinsic preference for school size. The inclusion of school size as an attribute is clearly problematic in our model, because schools in high demand will tend to be larger than schools in low demand. The coefficient on school size is therefore likely to be positive, not because parents prefer larger schools, but because high demand is a consequence of good quality. This is precisely what happens in our estimates, shown in tables A.22 and A.23 in the online Appendix A. Furthermore, all our remaining coefficients in equation (6) become very imprecise, in particular for boys. Consequently, without an instrument for school size we cannot include this attribute in our specification.

4.5.4 Considering different specifications of peer attributes

We also examined how our estimates of equation (6) changed when we either allowed peer attributes in schools to be endogenous in the model, or simply omitted these variables from the

[14]For each case, we also use only other villages in the same district with similar results.

model. There are some small changes in our estimates, when peer attributes are omitted, shown in tables A.24 and A.25 in the online Appendix A. For robustness we also consider the potential endogeneity of the measures of peer group "quality" that are likely important determinants of school choice. The endogeneity of peer effects has been extensively discussed in the literature on school (and neighborhood) choice (Bayer et al. (2007)).

In principle, in order to account for endogenous peer attributes in schools one would need to fully specify and solve the equilibrium model governing the sorting of students to schools, taking into account that each household's decision depends on the decision of every other household in the village. Bayer and Timmins (2007) propose a simpler IV procedure to estimate the individuals' valuation of peer attributes in a school, which is consistent with an equilibrium model, but does not require the full solution of a model (even in cases where there are likely to be multiple equilibria). In online Appendix C we present the full IV procedure for addressing peer effects using this method. Incorporating endogenous peer characteristics in our model changes the point estimates for the peer variables. However, overall they are not statistically significant for girls and for boys. School fee elasticities are similar to our main specification.[15] This suggests that the main conclusions of our paper are robust to how peer effects are modeled.

5 Simulations

5.1 The value of private schools and private school vouchers

We now use the demand model to examine the welfare implications of potential policies. Our motivation here is two-fold. First, the structure of the education system in Pakistan, like in many other low and middle-income countries, has changed substantially with a 10-fold expansion in the number of primary schools over the last two decades. How the emergence of private schools and potential policies towards this sector affects consumer welfare is therefore a first order question. Second, we are interested in the tension between using outcomes (such as

[15] These results are robust to a specification that allows changes in the peer composition.

test scores) as a measure of welfare versus the demand-based aggregates more common in the product literature. The fact that private schools charge (market-determined) prices in our setting opens up the possibility of using welfare measures derived from the demand model, which is what we attempt to do here. Any such exercise requires several assumptions. Most notably, we have not specified a supply-side model. The key assumptions therefore are that congestion costs, potential spillovers arising through the peer attributes in each school, as well as public school responses to a change in the private school environment, are all small. We discuss the limitations of our exercise in Section 5.2 below.[16]

We first focus on the welfare gains from private schooling. Using estimates from equations (5) and (6), we simulate the welfare consequences of closing down all private schools or alternatively, leaving one private school open in each village. This exercise is similar to valuing private schooling as a whole and valuing the product differentiation from multiple private schools. We then simulate the welfare impacts of an active schooling policy that provides education vouchers to those attending private schools, implying that effective fees in private schools are reduced to zero.

We use a standard measure of Compensating Variation (CV), which represents the change in a household's income that equates utility across two states: a benchmark state, which is the status quo, and the alternative state, which is the environment without private schools, or the environment with vouchers. Following Nevo (2000), and as shown in McFadden (1980) and Small and Rosen (1981), if the marginal utility of income is fixed for each individual, the compensating variation for individual i is given by

$$CV_i = \frac{\ln\left[\sum_{\widetilde{j}=0}^{\widetilde{J}} \exp(V_{i\widetilde{j}}^{Public})\right] - \ln\left[\sum_{j=0}^{J} \exp(V_{ij}^{\Pr ivate})\right]}{\frac{\partial V_{ij}^{private}}{\partial school fees}} \tag{7}$$

where $V_{ij}^{Private}$ represents the utility in the benchmark economy where both private and public

[16]We also assume that the policy changes do not affect the utility of not enrolling in any school.

schools coexist in the choice set of students, and $V_{i\tilde{j}}^{Public}$ represents the counterfactual scenario where only public schools are available to students.[17] The denominator represents the marginal utility of income.

In order to compute the total change in consumer welfare (TCV), one could average the compensating variation across sample and multiply by the number of students (M):

$$TCV = M \int CV_i dP_v(v) \tag{8}$$

where P is a distribution function. In practice, this average can be driven by extreme values both in the upper and the lower tails of the distribution of CV_i. In our setting these reflect extreme values of $\frac{\partial V_{ij}^{private}}{\partial schoolfees}$, which may be sensitive to modifications in the specification of observed and unobserved heterogeneity in the valuation of school fees. A more robust alternative is to present results based on the median (as opposed to the mean) value of CV_i in the sample. We use this as our main measure in the calculation of the welfare impacts of different policies. To estimate the total welfare of a policy we multiply this figure by the total number of students in the region we are considering.[18]

Table 5 presents estimates of the median compensating variation for a policy that forces private schools to shut down, separately for boys and girls. If we close all private schools, the estimated annual median compensating variation is \$4.8 dollars (37% of the average school fee) for boys, and \$1.4 for girls. If we focus only on those affected by the policy, i.e., those attending private schools in the current regime, then the estimated compensating variation is \$11.0 for boys and \$3.4 for girls. This compares to the average value of the fee of \$13 and is the amount that would have to be given to households to compensate them fully in money metric utility for the closure of private schools. The net benefit of private schools is therefore

[17] $V_{ij} = \delta_{jtg} + \sum_{k=1}^{K}\sum_{r=1}^{R} x_{jktg} z_{irtg} \beta_{rk}^{o} + \sum_{k=1}^{K} x_{jktg} v_{itg} \beta_k^u + \overline{\gamma} d_{ijtg} + \sum_{r=1}^{R} d_{ijtg} z_{irtg} \gamma_r + d_{ijtg} v_{itg} \gamma^u$

[18] An alternative, which we also implement (online Appendix - Table A.26), is to take the average of CV_i after trimming the bottom and top 5% of the distribution of this variable.

Table 5: No private schools - policy that forces private schools to shut down

	Girls	Boys
Median compensating variation (in U.S. dollars)	1.4	4.8
Median compensating variation - affected by the policy	3.4	11.0
Total change in consumer welfare (in thousand U.S. dollars)	51.0	242.5
Changes in total school enrollment rate (in percentage points)	-5.7	-5.4

Notes: This table presents changes in welfare, and school enrollment from the closure of all private schools. We use compensating variation (CV) to measure the changes in a household's income that equates utility across two states: a benchmark state, which is the status quo, and the alternative state, which is the environment without private schools. The first two rows present estimates of the median CV (in USD) for a policy that forces private schools to shut down, separately for boys and girls. The first row shows the results for everyone, while the second one shows the results for those affected by the policy. In this scenario (no private schools), those not affected by the policy intervention have no change in their consumer surplus. In the third row, we compute a measure of the total change in consumer welfare, in thousand USD taking the median CV across the sample and multiplying by the total number of students enrolled in the regions from our sample in rural Punjab, separately for girls and boys. The last row shows how total school enrollment changes (in percentage points) when the "no private schools" policy is implemented, separately for boys and girls.

1 U.S. dollars $\approx$ 85.6 Pakistani Rupees.

26% of the value of fees for girls, and 85% for boys. Another way to think about the value of private schools is that, for households whose children are in such schools, the benefit is equivalent to 7% of annual per-capita expenditure for boys, and 2% for girls. We also consider an alternative and less extreme way to restrict access to choice, where instead of forcing the closure of all private schools, we close all but one private school in each village. The private school that is allowed to remain open has the average characteristics of all the private schools in the village, and is located at the mean distance of private schools to the village (although the latter is clearly artificial since distance to a particular school should depend on where one resides). The amounts required to compensate families for such a change relative to the status quo (where public and private schools coexist), are 21% and 25% as high as those reported in the first row of Table 5 for girls and boys, respectively (see Table A.27 in the online Appendix). Therefore, a substantial part of the value of private schools comes from the fact that they make it possible to opt-out from the available public schools.[19]

Figure 2 plots the average CV estimates per village against the proportion of female and male students in the village in private schools. Not surprisingly, the correlation between these two variables is very strong for both boys and girls, showing that private school enrollment is high in villages where the valuation of the private school market is also high. The cross-village variation in this valuation is again striking. Our estimates of CV_i for the average student in a village ranges from \$0 to \$35 in the case of boys (with a mean of \$5 and a standard deviation of \$5), and from \$ 0 to \$12 in the case of girls (with a mean of \$2 and a standard deviation of \$2).

In the third row of Table 5 we multiply the numbers in the first row by the total number of students enrolled in the regions of our sample.[20] This gives us a measure of the annual welfare benefits of having private schools in these villages, relative to having no private school,

[19]The fact that a single private school adds considerably to consumer welfare captures, in part, that such a school reduces distances and therefore will apply to a *public* school as well.

[20]This assumes that the median CV numbers reported above are similar to the mean we would have obtained if we could perfectly correct for outlier CV values that are caused by

separately for girls and boys. The total value of private schools for parents of children in the regions we are considering is \$293,519 per year. If we extrapolated these values to the whole country, assuming similar valuations in other regions including urban centers (a likely underestimate as school fees are higher in urban areas) the value of private schools rises to \$138 million per year.

The fourth row of Table 5 shows how total school enrollment changes when the 'no private school' policy is implemented. Even though girls value private schools less than boys, the declines in overall school enrollment that we observe as a result of the policy are 5.7 p.p. for girls and 5.4 p.p. for boys. This is a relatively more important decline for girls, who start from a baseline enrollment rate of 67%, than for boys, who have an average enrollment rate of 80% in our sample. This means that the differential private school valuation across gender groups does not come from the fact that individuals are less likely to attend any school when private schools cease to exist, but from the fact that they have to switch from a private to a public school that is less desirable.

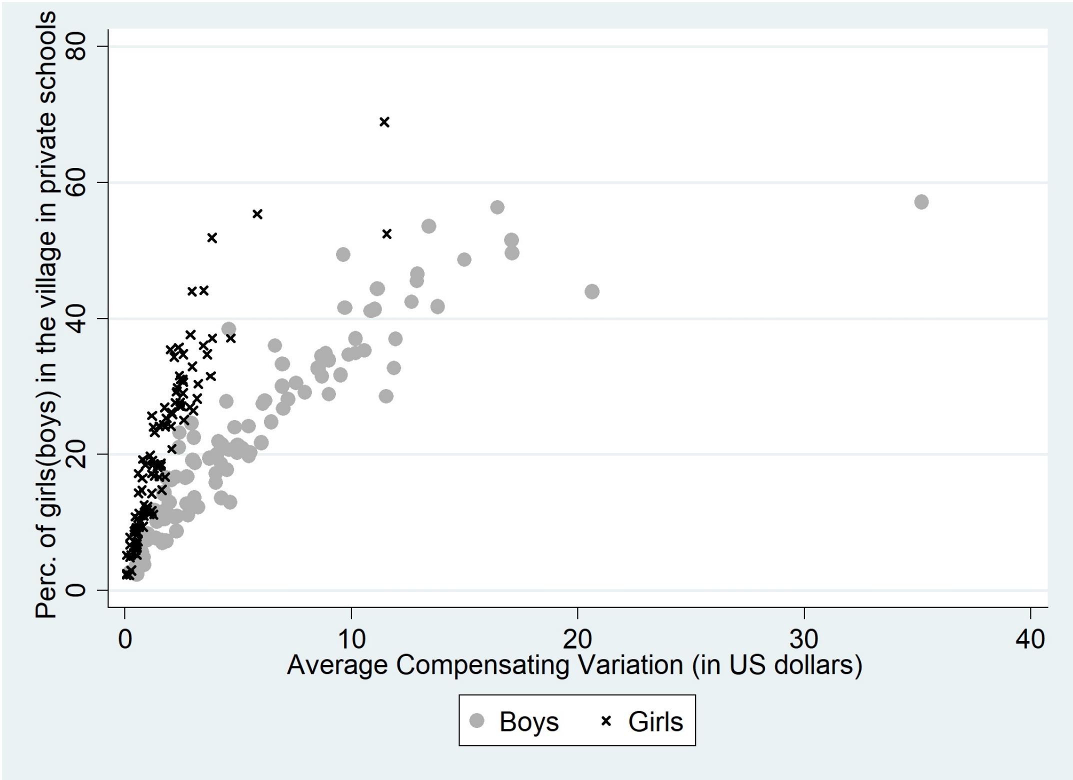
Table 6 considers a second policy, where school fees are equalized to zero across all schools.[21] One way to implement such a policy would be to offer each student a school voucher equal to the fees charged in each private school, which would be \$13 per student if every potential student decided to enrol in private school as a result. Table 6 shows that for the entire population of children in our sample, the median value of such a voucher would be \$2.7 for girls and \$2.4 boys. If we focus only on those attending private schools in the current regime, the estimated compensating variation is \$4.2 for girls and \$4.5 for boys. In the second row of Table 6 we again multiply these figures by the total number of boys and girls in the region we are considering.[22]

model misspecification.

[21] In this simulation, we reduce school fees but retain additional money that parents pay towards textbooks, uniforms and school supplies; in our data these costs are \$12 a year, which is very similar to the cost of private school tuition.

[22] Online Appendix A Table A.28 shows that the value of private schools and the value of

Figure 2: Compensating variation and proportion of girls (boys) in the village in private schools



Notes: This figure represents the average compensating variation per village and the proportion of girls (boys) in the village in a private school.

Our voucher policy (Table 6) increases total school enrollment by 8.4 p.p. for girls and 2.1 p.p. for boys. Private school enrollment rises by 21.1 p.p. for girls (from 19% to 40%) and 7.4 p.p. for boys (from 23% to 31%).[23] Public school enrollments decline by 12.7 and 5.3 p.p. for girls and boys, respectively. This means that the cost of the voucher per student is \$5.2 for girls ($= \$13 * 40\%$) and \$4.0 for boys.[24] Further, Andrabi et al. (2020) estimate that the cost per student in public schools is \$26. Therefore, the 12.7% for girls and the 5.3% for boys who move from public to private schools will save the government \$3.3 and \$1.4, respectively. This reduces the deadweight loss, and it is possible that the shadow value of frictions like credit constraints is higher than the remaining amount. Nevertheless, the increase in private schooling is smaller than what we would have expected if school fees were the only constraint on higher attendance.

5.2 A discussion of the limitations

Our estimates suggest that private schools add considerable value, especially for those who choose to use them, but that the value of vouchers is considerably lower than their costs. This is not a surprising result; absent any market failures, those who value the product at more than its price are already purchasers. What is of interest is the size of this gap as well as the simulated change in enrollment, which suggest that price is not the main barrier to private school attendance. As our emphasis on demand-based measures of welfare is not common in the education literature, we now discuss the limitations of our approach and the robustness of our estimates to alternate specifications.

school vouchers is higher for children with more educated mothers, especially for boys.

[23] Table 1 shows that 66.8% of all girls are enrolled in a school, and 28.0% of these are in a private school so that the proportion of girls attending a private school is 19% . An analogous calculation can be done for boys.

[24] Using the median school fee of \$11 to compute the costs of the policy gives a total cost per student of \$4.4 for girls and \$3.4 for boys.

Table 6: Voucher program simulation - policy where school fees are equalized to zero

	Girls	Boys
Median compensating variation (in U.S. dollars)	-2.7	-2.4
Total change in consumer welfare (in thousand U.S. dollars)	-102.5	-119.3
Changes in total school enrollment rate (in percentage points)	8.4	2.1
Changes in private school enrollment rate (in percentage points)	21.1	7.4
Changes in public school enrollment rate (in percentage points)	-12.7	-5.3

Notes: This table presents changes in welfare, and changes in total school enrollment from the introduction of vouchers. We use compensating variation (CV) to measure the changes in a household's income that equates utility across two states: a benchmark state, which is the status quo, and the alternative state, which is the environment where school fees are equalized to zero across all schools. The first row presents estimates of the median CV (in USD) for a policy where school fees are equalized to zero across all schools, separately for boys and girls. In the second row, we compute a measure of the total change in consumer welfare, in thousand USD, taking the median CV across the sample and multiplying by the total number of students enrolled in the regions from our sample in rural Punjab, separately for girls and boys. The last three rows show how total, private, and public school enrollment changes (in percentage points) when the "voucher program" policy is implemented, separately for boys and girls.
1 U.S. dollars ≈ 85.6 Pakistani Rupees.

5.2.1 Specification of the error term

The first concern with our welfare analysis regarding the value of private schools is that the i.i.d nature of the logit error can overstate true welfare from changes in the number of schools. Following Petrin (2002) we calculated the welfare change and simulated the decomposition into two components; one related to the observed characteristics entering the utility function and the other to the idiosyncratic logit taste term. The decomposition of compensation is the average difference in the value of observed and unobserved characteristics. As highlighted by Petrin (2002), introducing greater flexibility with the observed characteristics is likely to reduce the model's dependence on the error term, and lead to more stable results. Our results show that the total change in welfare change is similar to our counterfactual exercise and therefore not dominated by the logit taste component (Table A.29).

5.2.2 Changes in the peer group

Either the closing of private schools or the provision of school vouchers will likely change the peer groups in each school. Our calculations assume that product (school) attributes do not change as a result of the policy being simulated. When we relax this assumption, allowing re-sorting to take place in response to changes in peer attributes (relying on the point estimates of the valuation of peer attributes, even when they are imprecisely estimated), the estimated welfare impacts change at most by 1 to 3%, suggesting that the simpler specification we have used for our welfare computations are robust to changes in the peer composition.[25]

[25]For each simulation we estimated the welfare impacts updating $\tilde{p}_{jptg}$, the simulated value of peer attribute p in school j, with the new simulated probabilities for each individual (without a re-estimation of the model). The practical obstacles in implementing the full simulation arises from the fact that we use the school census to compute the average peer attributes at each school but we estimate the model in the (smaller) household survey. The correlation between the average peer attributes at each school computed using the census and the household survey is 0.5, which implies that, were we to use survey based school attributes for our simulations,

5.2.3 School Responses

Our approach could be rightly criticized both on the assumption that the voucher is made uniformly available to all children and villages, and there are no behavioral responses among public or private schools, ranging from new entry of schools to changes in prices or congestion effects. Each of these effects or changing the targeting design of the voucher would yield different impacts; if private schools respond by increasing prices, or if there are congestion effects, we are estimating an upper bound to the potential welfare gains. These counterfactuals are not observed in the data, and we have not modelled the supply-side in this paper. Nevertheless, ancillary evidence suggests that congestion effects and behavioral responses among *public* schools may be small.

To begin with, policy towards public schools in our context does not appear to take into account the presence (or responses) of the private sector. Online Appendix Figure A.1 shows that public schools preceded the arrival of private schools (there were a small number prior to 1972, when all schools were nationalized with the exception of some elite private schools) and it is reasonable to assume that their initial location and quality choices were not those of a 'leader' in a Stackelberg game.[26] To date, the government does not have a geographically linked database of public and private schools, and policies towards these two sectors have been undertaken by different bodies within the government with limited data sharing or advance planning.

Nevertheless, it is entirely possible that public schools will respond to changes in their *own* sector–for instance, if many more children enroll because private schools are shut down, schools may see declines in test scores. Our assumption of zero changes among children already choosing public schools is accurate only if congestion effects are small. This is a strong assumption that likely leads to an underestimation of the value of private schooling. Inter-

we would likely introduce substantial measurement error in the procedure.

[26]Although Andrabi et al. (2013) have shown how the construction of public schools itself led to the arrival of private schools by creating the necessary teacher pool in rural areas.

estingly, two recent studies from the LEAPS data both suggest that the assumption may not far-fetched.

First, Andrabi et al. (2020) compute School Value Added in the LEAPS sample and validate SVA measures for public schools using private school closures. This is close to what our simulation does, and they show that the estimates of SVA (computed from existing students, prior to the closure of the private school) is identical to the changes in test scores of children who are forced to move due to a private school closure. This implies that there is very limited response as children from private move to public schools. Second, Leclerc (2020) looks at private school entrance in the LEAPS data and shows that it reduces public school enrollment but again with no effect on test scores up to 4 years post-entry. This might be because the private schools are smaller, and there are more public schools. In our average village, shutting down all private schools would displace 242 students to 5 public schools for an average of 21 additional girls and 28 additional boys per public school, which translates into 9 children per grade (4 girls and 5 boys).

It is also possible that *private* schools will respond to a new voucher policy, either through new entry or through price responses. In our specific case, the vouchers that have been implemented allow for only one school per village (at least according to their rules, although there does seem to be some flexibility in this) so our assumption of no new entry may be plausible. However, it is very likely that private schools will change their pricing (both through 'top-up' pricing of vouchers and prices for regular students), again leading us to over-estimate the value of vouchers in our simulations.

5.2.4 Market Frictions

Finally, market frictions such as credit constraints or imperfect information will lead us to underestimate the valuation of vouchers. In that case, our estimates of the deadweight loss show approximately how large the shadow value of the market frictions must be for the vouchers to be cost-effective.

Our overall assessment is that our ex ante simulations provide valuable information for policy that is robust to alternate technical specifications. For instance, they clarify the key differences between providing a voucher to identify test-score differences between public and private schools and analyzing the welfare consequences of expanding a voucher to an entire schooling system. Nevertheless, the assumptions of limited school responses are very strong and would have to be reevaluated once such policies are actually enacted; the methods proposed here should be straightforward to extend to the actual evaluations of such policies. In the final section, we turn to one such experimental example that we implemented to assess the plausibility of a central parameter in our paper–the price elasticity of demand.

6 Voucher Experiment

In this section we provide suggestive experimental evidence supporting our structural estimates of price elasticities. The experiment is as follows: Between March and April, 2017, in 50% of the villages in our original sample, we offered vouchers of different amounts to cover private school fees to a random set of 812 households. In order to participate in the experiment, a household had to have a child in school, in 5th grade or below, or a child out of school who was between 5 and 15 years of age. Vouchers could be issued in 5 possible experimentally determined amounts: 50, 100, 150, 200, and 250 Pakistani Rupees (PKR) per month (for all school months in a year)[27]. A sixth group of families were assigned no voucher. The average amount of the vouchers (125 PKR $\approx$ 1.5 U.S. dollars) covered 25% of monthly private school tuition in the experimental sample. Online Appendix D provides the details of the experiment along with balance tests at the village and household-level in tables D.1 and D.3 (section D.1.1 in online Appendix D).[28]

[27] In U.S. dollars this corresponds to 0.6, 1.2,1.8, 2.3, and 2.9, respectively.

[28] Table D.4 tests for systematic differences in (a) whether a child is enrolled; (b) whether a child is enrolled in a public school and; (c) whether a child is enrolled in a private school by the voucher amounts. We never find any significant difference in the means, suggesting that the

The experimental and structural estimates are difficult to compare directly; the experiment takes place 14 years after the data used in the rest of the paper was collected and the subsidy was given for one year as opposed to the structural estimates, which are based on a permanent fee reduction.[29]. Nevertheless, the experiment was conducted in the same villages and the same households and there was little change in the schooling environment in terms of overall enrollment or aggregate test scores over this time, although there is some indication of more schooling at younger ages.[30] We therefore use the experiment, not to validate the structural model, but to rule-out elasticities that are much higher than what we have estimated.

We assess the comparability of our structural and experimental estimates in two ways. First, we regress private school attendance on the voucher size (online Appendix table D.8) and estimate sectoral price elasticities of private schools of -0.14 for girls and -0.35 for boys, compared to -0.27 and -0.10 from the structural estimates. Based on standard errors, the probability that the true elasticities are larger than 1 in absolute value is 4% for boys and 4.5%for girls. We cannot reject that the sectoral elasticities in the experiment and the structural estimates are the same; this comes with the substantial caveat that this is equally a problem of imprecision.

Second, we use the structural model to simulate what would happen to private school enrolment when a voucher is introduced. To replicate the experiment, the average value of the voucher in the simulation is set to be 25% of average private school fees, similar to the experiment. As in the experiment, in our simulation we also assigned five voucher amounts at random to our pseudo-population, corresponding to 10%, 20%, 30%, 40%, and 50% of the average tuition fees in our data. In the experimental data, offering the voucher has an average impact of 2.2 percentage points (p.p) and 1.7 p.p. for the private school enrollment of girls and boys,

experimental allocation is balanced across these categories.

[29]The extent to which this leads to lower elasticities in the experiment depends both on switching costs and the depreciation of test scores when modeled as a stock. See Das et al. (2013).

[30]Table D.5 shows that the difference in enrollment between the estimation and the experimental sample is not significant.

respectively. Our demand model implies instead an average impact of the voucher on private school enrollment of 4.9 p.p for girls and 1.6 p.p for boys. Although imprecision in both sets of estimates (as well as the comparability issues just discussed) makes it difficult to use one as a 'validation' for the other, the experiment, like the model shows a surprisingly low demand response to price reductions. These low estimates suggest that even in a poor environments such as the one we study, vouchers for private school attendance are unlikely to substantially change private school attendance. Instead, a voucher program will primarily translate into a cash windfall for those families whose children are already attending a private school.

7 Conclusion

Low cost private schools have expanded school choice to very poor areas, and in many countries more than half of total school enrollment is in private institutions. These are all environments where parents are, on average, poor and relatively less educated, but make active schooling decisions, often choosing to opt out of the free public school system. In order to understand the importance of private school markets for education in poor countries, we need to understand the parameters driving the demand and supply of private schooling in such settings. This is a central issue in the economics of education, where the roles of choice and competition in the provision of education are increasingly discussed. See Bau (2021), Neilson (2021) Burgess et al. (2015), Bayer et al. (2007), and Checchi and Jappelli (2004)).

Our demand estimates and policy simulations from Punjab, Pakistan highlight why such exercises are critical for policy. Parents value private schooling, but not the product differentiation that occurs when there are multiple private schools in the same village. Further, a voucher program in this setting has some effect on private and public enrollments, but not as large as is usually imagined. These exercises relate to fundamental issues in the economics of school choice and help inform important policy choices that governments are currently debating.

We are also aware of the limitations to this approach. For instance, were we to fully model changes in the schooling system from a counterfactual policy, we would also have to model supply side responses. But to do so, we need to first understand more fundamentally what private schools are maximizing. While clearly they are subject to some market discipline–in that they have to shut down if they cannot cover costs–their pricing decisions may reflect multiple objectives in addition to maximizing profits. As one example, we find that schools price in the inelastic portion of the demand curve with markups below those that would be profit maximizing. These pricing decisions could reflect many different considerations ranging from social concerns to dynamic pricing. Understanding why this is so remains at the frontier of this research.

References

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. *Journal of Human Resources*, 36(2):304–326.

Andrabi, T., Bau, N., Das, J., and Khwaja, A. (2020). Private schools, learning and civic values in a low-income country. *Working Paper*.

Andrabi, T., Das, J., and Khwaja, A. (2002). The rise of private schooling in pakistan: Catering to the urban elite or educating the rural poor? *mimeo*, Harvard University.

Andrabi, T., Das, J., and Khwaja, A. I. (2013). Students today, teachers tomorrow: Identifying constraints on the provision of education. *Journal of Public Economics*, 100:1 – 14.

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, 107(6):1535–63.

Andrabi, T., Das, J., Khwaja, A. I., Vishwanath, T., and Zajonc, T. (2007). Learning and educational achievements in punjab schools (leaps): Insights to inform the education policy debate. *mimeo*, World Bank, Washington, DC.

Arcidiacono, P., Muralidharan, K., Shim, E., and Singleton, J. D. (2021). Experimentally Validating Welfare Evaluation of School Vouchers - Part I. working paper, mimeo.

Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2020). Bride price and female education. *Journal of Political Economy*, 128(2):591–641.

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L., Raju, D., and Ryan, S. P. (2020). Delivering Education to the Underserved through a Public-Private Partnership Program in Pakistan. *The Review of Economics and Statistics*, pages 1–47.

Bau, N. (2021). Estimating an equilibrium model of horizontal competition in education. *Forthcoming Journal of Political Economy*.

Baum, D., Lewis, L., Lusk-Stover, O., and Patrinos, H. (2014). What matters most for engaging the private sector in education : A framework paper. systems approach for better education results. *(SABER) working paper*, World Bank, Washington, DC.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Bayer, P. and Timmins, C. (2007). Estimating Equilibrium Models Of Sorting Across Locations. *Economic Journal*, 117(518):353–374.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–90.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Berry, S. T. and Haile, P. A. (2009). Nonparametric identification of multinomial choice demand models with heterogeneous consumers. NBER Working Papers 15276, NBER.

Borker, G. (2020). Safety first: Perceived risk of street harassment and educational choices of women. Technical report, mimeo.

Bresnahan, T. F. and Gordon, R. J. (1997). *The Economics of New Goods*. University of Chicago Press.

Burde, D. and Linden, L. L. (2013). Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics*, 5(3):27–40.

Burgess, S., Greaves, E., Vignoles, A., and Wilson, D. (2015). What parents want: school preferences and school choice. *The Economic Journal*, 125(587).

Checchi, D. and Jappelli, T. (2004). School choice and quality. *Discussion Papers* 4748, CEPR.

Cheema, A., Khwaja, A., Naseer, F., and Shapiro, J. (2018). Glass Walls: Experimental Evidence on Access Constraints Faced by Women. Technical report.

Crawfurd, L. (2018). Contracting out schools at scale: evidence from Pakistan. Working Paper Series, RISE.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., and Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5(2):29–57.

Dinerstein, M. and Smith, T. (2018). Quantifying the Supply Response of Private Schools to Public Policies. Technical report, mimeo.

Dynarski, S., Gruber, J., and Li, D. (2009). Cheaper by the dozen: using sibling discounts at catholic schools to estimate the price elasticity of private school attendance. *NBER Working Papers* 15461, NBER.

Gallego, F. and Hernando, A. (2009). School choice in chile: looking at the demand side. *Documentos de Trabajo* 356, Instituto de Economia. Pontificia Universidad Católica de Chile.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER Working Papers* 12145, NBER.

Hausman, J. A. (1996). *Valuation of New Goods under Perfect and Imperfect Competition*, pages 207–248. University of Chicago Press.

Jacoby, H. G. and Mansuri, G. (2015). Crossing boundaries: How social hierarchy impedes economic mobility. *Journal of Economic Behavior and Organization*, 117(Sep):135–154.

Leclerc, C. (2020). Private school entry, sorting, and performance of public schools: Evidence from pakistan. *Working Paper*.

McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business*, 53(3):S13–29.

Muralidharan, K. and Sundararaman, V. (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in india. *The Quarterly Journal of Economics*.

Neilson, C. (2021). Targeted vouchers, competition among schools, and the academic achievement of poor students. mimeo, Yale University.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *RAND Journal of Economics*, 31(3):395–421.

PakistanToday (2019). Schools can't increase fees without prior approval. *Pakistan Today*.

Pathak, P. A. and Shi, P. (2017). How well do structural demand models work? counterfactual predictions in school choice. Working Paper 24017, National Bureau of Economic Research.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Reis, H. (2020). Girls' schooling choices and home production: Evidence from Pakistan. *International Economic Review*, 61(2):783–819.

Singh, A. (2015). Private school effects in urban and rural india: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113:16 – 32.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica*, 49(1):105–30.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Online appendices to "The Value of Private Schools: Evidence from Pakistan"

A Appendix tables and figures

In this section, we provide additional tables and figures for more details. The household and school variables used to estimate the model are described in Table A.1.

Tables A.2 and A.3 are analogous to Table 1, showing the attributes of schools attended by girls and boys, but distinguishing families with different levels of maternal education, household expenditure, and average distance between each household and other important facilities, which are often located in the center of the village.

We estimate equation (5) using maximum likelihood, with an additional step to estimate the school fixed effect. The first step estimated coefficients are shown in tables A.4, and A.5.

The coefficients in equation (6) can be estimated using instrumental variables, although we also present OLS estimates for comparison. The results for the first stage regressions are displayed in table A.6.

Table A.7 reports the school fee elasticity by grades, more specifically, on transition and non-transition grades, and Table A.8 presents the school fee elasticity by school fee quartiles.

Tables A.9 and A.10 show the willingness to pay for each school attribute in terms of distance to school, instead of in monetary terms. Table A.11 examines the correlation between parental preferences for different school attributes. Table A.12 compares the correlations between the same list of attributes offered by schools. Table A.13 report the IV regression with the coefficients of all interactions of our main specification.

We report the consolidated results from the multiple robustness checks in Tables A.14 and A.15, for alternative specifications and alternative instruments, respectively. In addition, we also include individual estimates from each specification. Table A.16 presents the IV regression of the specification excluding the interactions of the school fee with mother education, age, and household distance to facilities, as well as all the interactions with household distance to facilities. Table A.17 shows the out-of-sample exercise. Tables A.18 and A.19 present the first stage and IV regression, respectively, of the specification excluding teacher costs from the instrument set. Tables A.20 and A.21 show the first stage and IV regression, respectively, of the specification using hausman-style instruments as an extra instrument to our main specification.

Tables A.22 and A.23 show our estimates when the model includes school size as an attribute. Tables A.24 and A.25 show estimates of equation (6) in a specification where there are no peer variables and where peer variables are taken as exogenous.

Table A.26 present the welfare impacts of the different policies using the average of the compensating variation after trimming the bottom and top 5% of the distribution of this variable. Table A.27 presents changes in welfare from an alternative and less extreme way to restrict access to choice, where we close all but one private school in each village. Table A.28 shows the changes in welfare of the different policies by household type (mother education, expenditure, and household distance to facilities).

In the spirit of Petrin (2002) in table A.29 we have simulated the welfare change and the decomposition into two components. One component is related to the observed characteristics entering the utility function. The second component is related to the idiosyncratic logit taste term.

Figure A.1 shows the number of public and private schools by year of construction.

Table A.1: Variables definition

Variables	Description
School Variables	
School fees	Tuition annual fees
Private	Dichotomous variable indicating whether schools are private
School with toilets	Dichotomous variable indicating whether schools have toilets
School with permanent classroom	Dichotomous variable indicating whether schools have permanent classroom
Number of extra facilities	Number of extra facilities provided by the school
Student test score (average)	Student test score - Average of Math, Urdu and English
Percentage of female teachers	Percentage of female teachers
Percentage of teachers with at least 3 years of experience	Percentage of teachers with at least 3 years of experience
Percentage of teachers with university degree	Percentage of teachers with a university degree
Teacher absenteeism	Number of days teacher were absent in a month
Teacher test score (average)	Teacher test score - Average of Math, Urdu and English
Percentage of Mother with some education (school level)	Percentage of mothers with at least 1 year of education
Asset index (school level)	Average of the asset index at school level
Total costs without rent	Monthly expenditure on Utilities, Pay and Allowance of Teaching and Non-Teaching staff, purchase of educational material such as textbooks and other current disbursements.
Individual/Household variables	
Distance	Reports the distance in Kms from the house to any school available in the village
Age	Reports the children's age in years
Mother Education	Reports the students' mother education in years
Expenditure per capita	Total annual expenditure divided by household size
Household distance to facilities	Reports the average distance in Kms. from the house to the main facilities in the village

Table A.2: Summary statistics - school characteristics by type of household (girls)

Variables	Mother Education		Expenditure per capita				Household distance to facilities	
	Illiterate	At least some education	≤perc.25	>perc.25 and ≤perc.50	>perc.50 and ≤perc.75	>perc.75 and ≤perc.100	below median	above median
School fees	10.3	11.9	10.9	12.0	10.6	10.9	10.4	12.3
	(5.5)	(6.2)	(5.4)	(7.8)	(5.1)	(5.2)	(6.1)	(5.4)
School with toilets	0.81	0.84	0.76	0.81	0.85	0.85	0.86	0.78
	(0.39)	(0.37)	(0.43)	(0.39)	(0.35)	(0.36)	(0.35)	(0.42)
School with permanent classroom	0.91	0.91	0.92	0.92	0.93	0.88	0.93	0.89
	(0.28)	(0.28)	(0.27)	(0.28)	(0.25)	(0.32)	(0.26)	(0.31)
Number of extra facilities	2.95	3.16	2.78	3.03	3.03	3.18	3.20	2.78
	(1.46)	(1.43)	(1.64)	(1.42)	(1.46)	(1.30)	(1.34)	(1.56)
Percentage of female teachers	0.93	0.93	0.91	0.94	0.94	0.92	0.95	0.90
	(0.22)	(0.18)	(0.25)	(0.19)	(0.18)	(0.21)	(0.16)	(0.25)
Percentage of teachers with at least 3 years of experience	0.77	0.72	0.82	0.77	0.76	0.67	0.73	0.77
	(0.30)	(0.32)	(0.27)	(0.30)	(0.31)	(0.32)	(0.29)	(0.29)
Percentage of teachers with university degree	0.31	0.27	0.31	0.31	0.31	0.28	0.30	0.30
	(0.25)	(0.24)	(0.25)	(0.26)	(0.26)	(0.24)	(0.24)	(0.26)
Student test score (average)	0.32	0.34	0.32	0.32	0.32	0.34	0.33	0.32
	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)	(0.12)
Teacher absenteeism	2.41	2.05	2.13	2.76	2.36	1.94	1.95	2.76
	(4.07)	(3.82)	(3.09)	(4.93)	(4.36)	(3.24)	(2.85)	(5.14)
Teacher test score (average)	0.86	0.86	0.86	0.86	0.86	0.87	0.87	0.85
	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)
Percentage of mothers with some educ. (sch. level)	0.24	0.32	0.24	0.24	0.27	0.30	0.30	0.23
	(0.21)	(0.24)	(0.24)	(0.21)	(0.21)	(0.25)	(0.24)	(0.21)
Asset index (school level)	-0.48	-0.27	-0.70	-0.46	-0.31	-0.24	-0.38	-0.45
	(0.79)	(0.80)	(0.79)	(0.82)	(0.73)	(0.78)	(0.75)	(0.87)
Distance	0.53	0.46	0.63	0.51	0.44	0.47	0.36	0.71
	(0.65)	(0.57)	(0.69)	(0.63)	(0.50)	(0.67)	(0.48)	(0.75)
Attending school	59.5	87.8	53.7	64.8	72.2	77.6	76.6	56.7
Attending private school	22.2	39.2	16.1	24.6	28.0	39.6	30.7	24.4

Notes: Means and standard deviations of school characteristics by household type. The standard deviation is in brackets.
This table is analogous to Table 1 (panel B), showing characteristics of schools attended by girls, but distinguishing families with different levels of maternal education, household expenditure, and household distance to village facilities. The first column presents results for the children with illiterate mothers (0 years of education) and column (2) for mother with at least some education (1 or more years of education). Columns (3) to (6) distinguish families by household expenditure: below 25th percentile, between 25th and 50th percentile, between 50th and 75th percentile, and above 75th percentile. The last two columns present the school characteristics for households leaving near the village facilities (below median) and those leaving more distant from the village facilities (above median), respectively.

Table A.3: Summary statistics - school characteristics by type of household (boys)

Variables	Mother Education		Expenditure per capita				Household distance to facilities	
	Illiterate	At least some education	≤perc.25	>perc.25 and ≤perc.50	>perc.50 and ≤perc.75	>perc.75 and ≤perc.100	below median	above median
School fees	12.9	13.0	13.1	13.9	12.9	12.1	12.7	13.3
	(8.7)	(7.9)	(4.7)	(9.2)	(9.3)	(8.0)	(9.0)	(7.1)
School with toilets	0.62	0.73	0.69	0.62	0.65	0.65	0.65	0.66
	(0.48)	(0.44)	(0.46)	(0.49)	(0.48)	(0.48)	(0.48)	(0.48)
School with permanent classroom	0.90	0.93	0.93	0.87	0.94	0.88	0.93	0.88
	(0.29)	(0.26)	(0.22)	(0.33)	(0.24)	(0.31)	(0.24)	(0.31)
Number of extra facilities	2.43	2.93	2.56	2.45	2.61	2.66	2.74	2.40
	(1.57)	(1.66)	(1.48)	(1.49)	(1.63)	(1.64)	(1.60)	(1.61)
Percentage of female teachers	0.20	0.42	0.15	0.30	0.25	0.34	0.30	0.21
	(0.37)	(0.45)	(0.32)	(0.42)	(0.40)	(0.44)	(0.43)	(0.37)
Percentage of teachers with at least 3 years of experience	0.76	0.65	0.80	0.71	0.74	0.67	0.70	0.76
	(0.29)	(0.33)	(0.28)	(0.30)	(0.30)	(0.33)	(0.32)	(0.29)
Percentage of teachers with university degree	0.40	0.32	0.44	0.36	0.38	0.33	0.37	0.39
	(0.28)	(0.27)	(0.27)	(0.28)	(0.28)	(0.27)	(0.27)	(0.29)
Student test score (average)	0.30	0.35	0.31	0.31	0.32	0.32	0.32	0.31
	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)	(0.13)	(0.11)
Teacher absenteeism	1.90	1.43	2.04	1.66	1.78	1.58	1.44	2.11
	(2.57)	(1.99)	(2.65)	(2.49)	(2.46)	(2.06)	(1.91)	(2.84)
Teacher test score (average)	0.89	0.87	0.88	0.88	0.88	0.88	0.88	0.89
	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)	(0.08)	(0.10)	(0.08)
Percentage of mothers with some educ. (sch. level)	0.19	0.28	0.14	0.22	0.22	0.27	0.25	0.17
	(0.21)	(0.23)	(0.18)	(0.23)	(0.21)	(0.24)	(0.23)	(0.20)
Asset index (school level)	-0.85	-0.44	-1.19	-0.64	-0.60	-0.48	-0.61	-0.87
	(0.87)	(0.87)	(0.85)	(0.83)	(0.80)	(0.91)	(0.80)	(0.95)
Distance	0.74	0.52	0.98	0.61	0.56	0.55	0.39	0.97
	(0.92)	(0.71)	(1.18)	(0.73)	(0.69)	(0.69)	(0.46)	(1.08)
Attending school	76.6	90.0	73.5	79.1	83.7	84.2	83.2	76.6
Attending private school	21.5	48.0	16.4	32.0	30.0	37.0	34.0	23.1

Notes: Means and standard deviations of school characteristics by household type. The standard deviation is in brackets.
This table is analogous to Table 1 (panel B), showing characteristics of schools attended by boys, but distinguishing families with different levels of maternal education, household expenditure, and household distance to village facilities. The first column presents results for the children with illiterate mothers (0 years of education) and column (2) for mother with at least some education (1 or more years of education). Columns (3) to (6) distinguish families by household expenditure: below 25th percentile, between 25th and 50th percentile, between 50th and 75th percentile, and above 75th percentile. The last two columns present the school characteristics for households leaving near the village facilities (below median) and those leaving farther away from the village facilities (above median), respectively.

Table A.4: Estimates of interaction terms - observables

Individual/household characteristic	School characteristic	Girls	Boys
Age	School fees	-0.004*	-0.002
		(0.003)	(0.002)
	Number of extra facilities	0.012	0.025***
		(0.010)	(0.008)
	Percentage of female teachers	0.040	-0.074*
		(0.058)	(0.050)
	Percentage of teachers with at least 3 years of experience	0.001	0.025
		(0.048)	(0.049)
	Percentage of teachers with university degree	0.057	0.078*
		(0.052)	(0.047)
	Student test score (average)	-0.081	-0.068
		(0.142)	(0.114)
	Teacher absenteeism	-0.002	0.003
		(0.003)	(0.004)
	Teacher test score (average)	0.004	-0.046
		(0.160)	(0.127)
	Perc. of Mother with some education (school level)	0.049	-0.086
		(0.060)	(0.058)
	Asset index (school level)	0.011	0.014
		(0.016)	(0.014)
	Outside option - not enrolled	0.253	0.151
		(0.158)	(0.131)
	School with toilets	0.006	-0.003
		(0.039)	(0.027)
	School with permanent classroom	0.056	0.058
		(0.041)	(0.039)
	Private	0.027	0.057
		(0.043)	(0.060)
	Distance	0.237	0.138
		(0.163)	(0.120)
Mother Education	School fees	0.008***	0.004
		(0.003)	(0.003)
	Number of extra facilities	0.011	0.008
		(0.011)	(0.009)
	Percentage of female teachers	0.067	0.091**
		(0.072)	(0.053)
	Percentage of teachers with at least 3 years of experience	0.063	0.078
		(0.054)	(0.054)
	Percentage of teachers with university degree	0.039	-0.048
		(0.058)	(0.055)
	Student test score (average)	-0.074	0.093
		(0.139)	(0.131)
	Teacher absenteeism	0.004	0.002
		(0.004)	(0.005)
	Teacher test score (average)	-0.038	-0.059
		(0.224)	(0.121)
	Perc. of Mother with some education (school level)	0.180***	-0.003
		(0.068)	(0.058)
	Asset index (school level)	0.008	0.035***
		(0.018)	(0.016)
	Outside option - not enrolled	-0.134	0.034
		(0.264)	(0.131)
	School with toilets	-0.016	0.082***
		(0.049)	(0.032)
	School with permanent classroom	-0.023	0.057
		(0.047)	(0.046)
	Private	0.070	0.023
		(0.050)	(0.065)
	Distance	0.004	-0.001
		(0.020)	(0.016)

Notes: This table reports estimates of the interaction terms (β^o_{rkg}, and γ_{rg}) for students' age and mother education in equation (5) for both girls and boys. This step entails estimating $\delta_{jtg}, \beta^o_{rkg}, \beta^u_{kg}, \overline{\gamma}_g, \gamma_{rg}, \gamma^u_g$ by maximum likelihood, including a contraction mapping to obtain δ_{jtg}. Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.4: Estimates of interaction terms - observables (continued)

Individual/household characteristic	School characteristic	Girls	Boys
log income	School fees	0.029***	0.003
		(0.008)	(0.007)
Household distance to facilities	School fees	0.006	0.001
		(0.034)	(0.016)
	Number of extra facilities	0.101	-0.025
		(0.177)	(0.053)
	Percentage of female teachers	0.059	0.080
		(0.538)	(0.315)
	Percentage of teachers with at least 3 years of experience	0.431	0.146
		(1.023)	(0.365)
	Percentage of teachers with university degree	0.194	-0.175
		(0.728)	(0.311)
	Student test score (average)	0.695	-0.126
		(1.241)	(0.888)
	Teacher absenteeism	0.026	0.017
		(0.033)	(0.028)
	Teacher test score (average)	-0.542	-0.017
		(0.666)	(0.893)
	Perc. of Mother with some education (school level)	0.186	-0.066
		(0.821)	(0.406)
	Asset index (school level)	-0.020	0.011
		(0.160)	(0.091)
	Outside option - not enrolled	0.346	-0.059
		(0.256)	(0.883)
	School with toilets	-0.047	-0.070
		(0.407)	(0.194)
	School with permanent classroom	-0.269	0.044
		(0.575)	(0.245)
	Private	0.189	-0.045
		(0.652)	(0.395)
	Distance	0.048	0.039
		(0.495)	(0.291)

Notes: This table reports estimates of the interaction terms (β^o_{rkg}, and γ_{rg}) for log of expenditure and household distance to facilities in equation (5) for both girls and boys. For log of expenditure our specification includes only the interation with school fees. The first step entails estimating $\delta_{jtg}, \beta^o_{rkg}, \beta^u_{kg}, \overline{\gamma}_g, \gamma_{rg}, \gamma^u_g$ by maximum likelihood, including a contraction mapping to obtain δ_{jtg}. Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.5: Estimates of interaction terms - unobservables

School Characteristics	Girls	Boys
School fees	-0.0013	-0.0008
	(0.0505)	(0.1628)
Number of extra facilities	-0.0024	-0.0012
	(0.0580)	(0.0449)
Percentage of female teachers	-0.0011	0.0000
	(0.2884)	(0.3142)
Percentage of teachers with at least 3 years of experience	-0.0010	0.0000
	(0.2011)	(0.1852)
Percentage of teachers with university degree	-0.0001	0.0000
	(0.2964)	(0.4246)
Student test score (average)	-0.0003	-0.0001
	(0.8211)	(0.5673)
Teacher absenteeism	-0.0014	0.0008
	(0.0122)	(0.0311)
Teacher test score (average)	-0.0009	0.0002
	(0.3637)	(0.4822)
Perc. of Mother with some education (school level)	-0.0001	0.0000
	(0.1279)	(0.3307)
Asset index (school level)	0.0007	-0.0001
	(0.0799)	(0.0442)
Outside option - not enrolled	0.0009	0.0001
	(0.4360)	(0.2427)
School with toilets	-0.0007	-0.0005
	(0.1253)	(0.2240)
School with permanent classroom	-0.0013	-0.0001
	(0.1472)	(0.1275)
Private	-0.0007	0.0001
	(0.0777)	(0.3057)
Distance	0.0004	-0.0002
	(0.0604)	(0.0748)

Notes: This table reports estimates of the interaction terms for the individual unobservable characteristics in equation (5) for both girls and boys (β^u_{rkg}, and γ^u_g). The first step entails estimating $\delta_{jtg}, \beta^o_{rkg}, \beta^u_{kg}, \overline{\gamma}_g, \gamma_{rg}, \gamma^u_g$ by maximum likelihood, including a contraction mapping to obtain δ_{jtg}.
Standard errors in brackets.
* Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.6: First stage - school fee equation

	Girls		Boys	
	(1)	(2)	(3)	(4)
School with toilets x Private	139.120	138.883	144.454	142.729
	[142.904]	[142.883]	[129.484]	[129.625]
School with permanent classroom x Private	119.682	120.816	73.623	73.718
	[82.455]	[82.438]	[77.753]	[77.825]
Number of extra facilities x Private	76.336***	74.915***	55.101**	54.817**
	[26.773]	[26.778]	[25.292]	[25.317]
Percentage of female teachers x Private	-405.513***	-412.976***	-489.651***	-496.874***
	[110.219]	[110.031]	[100.414]	[100.496]
Perc. of teachers with	-302.712***	-299.171***	-330.681***	-329.681***
at least 3 years of exp. x Private	[112.674]	[112.566]	[104.588]	[104.682]
Perc. of teachers with university	497.712***	499.708***	496.554***	503.974***
degree x Private	[167.327]	[167.220]	[152.667]	[152.680]
Student test score x Private	1,276.462***	1,278.378***	1,161.171***	1,158.599***
	[282.099]	[282.066]	[257.407]	[257.650]
Teacher absenteeism x Private	-8.719	-8.597	-7.311	-6.829
	[14.936]	[14.928]	[13.557]	[13.562]
Teacher test score x Private	833.574**	861.047**	928.399**	941.449***
	[340.432]	[340.940]	[360.835]	[361.500]
Perc. of Mother with some	-275.935**	-273.707**	-204.596**	-203.459**
education x Private	[110.505]	[110.536]	[101.229]	[101.359]
Asset index x Private	73.824**	72.969**	58.044*	58.748*
	[36.264]	[36.275]	[33.860]	[33.885]
Private	-97.505	-162.397	98.086	77.048
	[619.764]	[620.762]	[589.218]	[590.157]
Teacher costs of other schools	0.055***		0.065***	
in the same tehsil x Private	[0.012]		[0.009]	
Total costs without rent of other schools		0.045***		0.052***
in the same tehsil x Private		[0.010]		[0.008]
Total cost without rent x Private	1.131***	1.128***	1.101***	1.092***
	[0.286]	[0.285]	[0.258]	[0.258]
Number of schools within 2Km. x Private	9.601	10.094	15.474*	15.469*
	[9.341]	[9.341]	[8.112]	[8.119]
Number of extra facilities of the	-19.732	-20.508	3.442	4.212
competitors x Private	[45.547]	[45.564]	[43.919]	[43.953]
Perc. of teachers with at least 3 years of	-401.302	-377.680	-205.967	-186.963
experience of the competitors x Private	[274.006]	[273.602]	[254.582]	[254.864]
Teacher test score of the	-120.081	-85.204	-662.637	-633.530
competitors x Private	[556.874]	[556.581]	[533.566]	[533.961]
Asset index of the competitors x Private	216.201***	209.251***	90.416	88.580
	[70.027]	[70.252]	[66.356]	[66.539]
F-test (Instruments)	10.15	10.17	15.93	15.76
p-value	0,000	0,000	0,000	0,000
Observations	511	511	522	522
R-squared	0.671	0.671	0.711	0.710

Notes: This table reports estimates of the first stage regression of school fees for both girls and boys. Columns (1) and (3) report our main specification using teacher costs at village using the costs of the other villages in the same sub-district (tehsil), as well as total costs without rent of each school, and BLP-type of instruments. Columns (2) and (4) use total cost without rent at the village using the costs of the other villages in the same sub-district (tehsil), as well as the school costs without rent, and BLP-type of instruments. In this regression, we interact the school attributes with private school indicator, allowing us to use all schools in the first stage. Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.7: School Fee Elasticity by Grades (average of school elasticities)

	Girls	Boys
Grades 3 and 4	-0.79	-0.17
Grade 5	-1.85	-0.73
Whole Sample	-1.12	-0.37

Notes: This table reports for both, girls and boys, the school fee elasticity for transitional (grade 5) and non-transitional (grades 3 and 4) grades.

Table A.8: School Fee Elasticity by School Fee Quartiles

	Girls	Boys
Mean	-1.12	-0.37
By School Fee		
First Quartile (below percentile 25th)	-0.51	-0.16
Second Quartile (between percentile 25th and 50th)	-0.83	-0.26
Third Quartile (between percentile 50th and 75th)	-1.17	-0.37
Fourth Quartile (above percentile 75th)	-2.02	-0.71

Notes: This table reports for both, girls and boys, the school fee elasticity by school fee quartiles.

Table A.9: Willingness to pay for school characteristics in terms of distance - girls

					Willingness to Pay (in distance terms Kms)		
	25th percentile	mean	75th percentile	Variable variation	25th percentile	mean	75th percentile
School fees	-0.167***	-0.136***	-0.120***	13.3	-1.00	-0.81	-0.71
	[0.041]	[0.041]	[0.041]				
School with toilets	0.164	0.122	0.125	1.00	0.07	0.05	0.06
	[0.379]	[0.361]	[0.370]				
School with permanent classroom	0.300	0.225	0.218	1.00	0.13	0.10	0.10
	[0.277]	[0.266]	[0.270]				
Number of extra facilities	0.178**	0.198***	0.197***	1.00	0.08	0.09	0.09
	[0.071]	[0.072]	[0.073]				
Percentage of female teachers	0.496	0.592*	0.610*	0.10	0.02	0.03	0.03
	[0.341]	[0.336]	[0.342]				
Percentage of teachers with 3 years of experience	0.137	0.251	0.304	0.10	0.01	0.01	0.01
	[0.319]	[0.321]	[0.343]				
Percentage of teachers with university degree	0.432	0.530	0.553	0.10	0.02	0.02	0.02
	[0.430]	[0.449]	[0.432]				
Student test score (average)	1.447*	1.443*	1.394*	0.13	0.08	0.08	0.08
	[0.759]	[0.777]	[0.783]				
Teacher absenteeism	0.036	0.036	0.039	1.00	0.02	0.02	0.02
	[0.042]	[0.042]	[0.041]				
Teacher test score (average)	1.620	1.600	1.574	0.08	0.06	0.06	0.06
	[1.028]	[1.073]	[1.096]				
Perc. of Mother with some education (school level)	-0.976***	-0.711**	-0.620*	0.10	-0.04	-0.03	-0.03
	[0.333]	[0.363]	[0.348]				
Asset index (school level)	-0.064	-0.056	-0.046	1.05	-0.03	-0.03	-0.02
	[0.099]	[0.099]	[0.098]				
Private	-0.357	-0.254	-0.211	1.00	-0.16	-0.11	-0.09
	[0.483]	[0.494]	[0.465]				
Distance	-2.223***	-2.233***	-2.252***				
	[0.145]	[0.136]	[0.142]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the WTP for each of them, change with the family background of the girl. In this table we express WTP for each school attribute in terms of distance to school, instead of expressing it in monetary terms. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile $\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 17 report the WTP for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.10: Willingness to pay for school characteristics in terms of distance - boys

					Willingness to Pay (in distance terms Kms)		
	25th percentile	mean	75th percentile	Variable variation	25th percentile	mean	75th percentile
School fees	-0.051**	-0.043*	-0.039	13.3	-0.60	-0.50	-0.45
	[0.025]	[0.025]	[0.025]				
School with toilets	0.162	0.280	0.329	1.00	0.14	0.24	0.28
	[0.238]	[0.237]	[0.251]				
School with permanent classroom	0.107	0.182	0.215	1.00	0.09	0.16	0.18
	[0.198]	[0.204]	[0.203]				
Number of extra facilities	0.115**	0.122**	0.127**	1.00	0.10	0.11	0.11
	[0.055]	[0.056]	[0.057]				
Percentage of female teachers	-0.852***	-0.747***	-0.675**	0.10	-0.08	-0.06	-0.06
	[0.269]	[0.275]	[0.271]				
Percentage of teachers with at least 3 years of experience	0.091	0.186	0.240	0.10	0.01	0.02	0.02
	[0.286]	[0.283]	[0.281]				
Percentage of teachers with university degree	0.171	0.112	0.091	0.10	0.02	0.01	0.01
	[0.329]	[0.330]	[0.310]				
Student test score (average)	0.996	1.146*	1.231*	0.13	0.11	0.13	0.14
	[0.644]	[0.640]	[0.642]				
Teacher absenteeism	0.002	0.003	0.004	1.00	0.00	0.00	0.00
	[0.024]	[0.023]	[0.024]				
Teacher test score (average)	1.104	1.051	0.992	0.09	0.09	0.08	0.08
	[0.726]	[0.709]	[0.767]				
Perc. of Mother with some education (school level)	-0.368	-0.381	-0.391	0.10	-0.03	-0.03	-0.03
	[0.298]	[0.307]	[0.303]				
Asset index (school level)	-0.009	0.030	0.052	1.14	-0.01	0.03	0.05
	[0.075]	[0.074]	[0.076]				
Private	-0.929**	-0.911**	-0.909**	1.00	-0.82	-0.79	-0.78
	[0.405]	[0.396]	[0.387]				
Distance	-1.131***	-1.151***	-1.165***				
	[0.089]	[0.079]	[0.084]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the WTP for each of them, change with the family background of the boy. In this table we express WTP for each school attribute in terms of distance to school, instead of expressing it in monetary terms. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile $\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 7 report the willingness to pay for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.11: Correlation of the individual preferences regarding school characteristics

GIRLS

	School Fee	Distance	Toilets	Permanent classroom	Number of extra facilities	Female teachers	Student Test Scores
School Fee	1						
Distance	0.2615	1					
Schools with toilets	-0.6454	-0.8074	1				
Schools with permanent classroom	-0.5309	-0.7891	0.9728	1			
Number of Extra Facilities	0.4374	0.9425	-0.9483	-0.9485	1		
Percentage of Female Teachers	0.5835	0.7810	-0.7396	-0.5834	0.7323	1	
Student test scores	0.3808	0.8356	-0.9212	-0.9797	0.9575	0.5171	1

BOYS

	School Fee	Distance	Toilets	Permanent classroom	Number of extra facilities	Female teachers	Student Test Scores
School Fee	1						
Distance	0.0079	1					
Schools with toilets	0.4768	-0.6847	1				
Schools with permanent classroom	0.2619	0.6377	0.0832	1			
Number of Extra Facilities	-0.3682	-0.4286	0.5630	0.2303	1		
Percentage of Female Teachers	0.9395	0.2644	0.1730	0.2728	-0.6205	1	
Student test scores	0.4783	-0.8557	0.9157	-0.2959	0.3271	0.2027	1

Notes: This table reports the correlation between individual preferences for different school attributes for both girls and boys. Recall that two of the attributes in this table have negative coefficients in parental preferences: school fees and distance. We do not show all attributes in this table, but only the ones for which the coefficients were statistically significant in equation (6) for girls or boys.

Table A.12: Correlation of the attributes offered by schools

GIRLS

All Schools

	School Fee	Toilets	Permanent classroom	Number extra fac.	Female teacher	Student test scores
School Fee	1.00					
Schools with toilets	0.27	1.00				
Schools with permanent classroom	0.02	0.11	1.00			
Number of extra facilities	0.50	0.45	0.12	1.00		
Percentage female teachers	-0.17	0.02	0.08	-0.01	1.00	
Student Test Scores	0.48	0.22	0.01	0.34	-0.10	1.00

Private schools

	School Fee	Toilets	Permanent classroom	Number extra fac.	Female teacher	Student test scores
School Fee	1.00					
Schools with toilets	0.14	1.00				
Schools with permanent classroom	0.14	0.16	1.00			
Number of extra facilities	0.31	0.29	0.21	1.00		
Percentage female teachers	-0.17	0.10	0.04	-0.001	1.00	
Student Test Scores	0.31	0.11	0.10	0.18	-0.05	1.00

BOYS

All Schools

	School Fee	Toilets	Permanent classroom	Number extra fac.	Female teacher	Student test scores
School Fee	1.00					
Schools with toilets	0.38	1.00				
Schools with permanent classroom	0.05	0.09	1.00			
Number of extra facilities	0.54	0.48	0.14	1.00		
Percentage female teachers	0.47	0.40	-0.01	0.43	1.00	
Student Test Scores	0.53	0.37	0.03	0.43	0.39	1.00

Private schools

	School Fee	Toilets	Permanent classroom	Number extra fac.	Female teacher	Student test scores
School Fee	1.00					
Schools with toilets	0.14	1.00				
Schools with permanent classroom	0.14	0.16	1.00			
Number of extra facilities	0.31	0.29	0.21	1.00		
Percentage female teachers	-0.17	0.10	0.04	0.001	1.00	
Student Test Scores	0.30	0.11	0.10	0.17	-0.04	1.00

Notes: This table reports the correlation between the list of attributes offered by schools (all and private) for both girls and boys. We do not show all attributes in this table, but only the ones for which the coefficients were statistically significant in equation (6).

Table A.13: IV regression

	Girls	Boys
School Fee	-0.136***	-0.043*
	[0.041]	[0.025]
School with toilets x private	0.191	0.637
	[0.677]	[0.435]
School with permanent classroom x private	-0.158	-0.574
	[0.546]	[0.403]
Number of extra facilities x private	0.043	-0.015
	[0.141]	[0.114]
Percentage of female teachers x private	-1.702***	-0.179
	[0.657]	[0.537]
Perc. of teachers with 3 years of exp. x private	0.619	0.722
	[0.663]	[0.547]
Perc. of teachers with univ. degree x private	1.157	-0.732
	[0.811]	[0.637]
Student test score (average) x private	-0.624	-0.982
	[1.493]	[1.263]
Teacher absenteeism x private	0.083	0.015
	[0.071]	[0.041]
Teacher test score (average) x private	-3.108	0.864
	[2.112]	[1.361]
Perc. mother with some educ. x private	0.039	-0.474
	[0.693]	[0.611]
Asset index (school level) x private	0.332*	-0.005
	[0.188]	[0.148]
School with toilets	0.019	-0.043
	[0.304]	[0.180]
School with permanent classroom	0.309	0.474
	[0.482]	[0.329]
Number of extra facilities	0.175*	0.130
	[0.099]	[0.089]
Percentage of female teachers	1.505***	-0.656
	[0.441]	[0.451]
Perc. of teachers with 3 years of exp.	-0.081	-0.180
	[0.543]	[0.458]
Perc. of teachers with university degree	-0.091	0.484
	[0.493]	[0.421]
Student test score	1.778*	1.644*
	[0.975]	[0.947]
Teacher absenteeism	-0.008	-0.004
	[0.018]	[0.022]
Teacher test score	3.267*	0.613
	[1.713]	[0.918]
Perc. of Mother with some educ. (sch. level)	-0.732	-0.140
	[0.576]	[0.521]
Asset index (school level)	-0.234*	0.033
	[0.135]	[0.105]
private	3.149	-1.309
	[2.234]	[1.332]
Constant	-3.282**	-0.029
	[1.655]	[0.923]
Observations	511	522
R-squared	0.251	0.280

Notes: This table shows the IV estimated coefficients for equation (6) for girls and boys (estimation of $\overline{\beta}_{kg}$ by running a regression of the school fixed effect (δ_{jtg}) on the observed school characteristics (including the interactions with private school indicator). Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.14: Robustness - Main Outcomes - Alternate Specifications

Robustness	Specification	School Fee Elasticity		Distance elasticity		WTP distance	
		Boys	Girls	Boys	Girls	Boys	Girls
Main Specification	**Main Specification**	**-0.37**	**-1.12**	**-6.0**	**-11.1**	**-15.6**	**-9.6**
Overfitting	All interactions	-0.48	-1.04	-4.8	-9.6	-10.8	-9.6
Overfitting	Exc. some interactions*	-0.46	-1.49	-6.6	-11.4	-15.5	-8.9
Number of children	With number of children	-0.36	-1.03	-5.8	-9.5	-16.4	-9.9
Quadratic distance term	Quadratic distance term	-0.44	-0.92	-6.2	-8.6	-13.9	-10.6

Notes: This table reports the consolidated results on key output measures from the multiple robustness checks on alternate specifications, namely, overfitting, number of children and distance quadratic term.
* Excluding some interactions: this specification excludes parameters of school fees interacted with mothers education, age, and household distance to facilities, as well as all the parameters interacted with household distance to facilities.

Table A.15: Robustness - Main Outcomes - Alternate Instruments

Robustness	Instruments	F-Test (p-value)		Price Elasticity		distance Elasticity		WTP distance	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Main specification	**Teacher costs of other villages + costs without rent + BLP-type instr.**	**15.93 (0.000)**	**10.15 (0.000)**	**-0.37**	**-1.12**	**-6.0**	**-11.1**	**-15.6**	**-9.6**
Hausman type - group private schools in 4 categories by size	Main + avg prices in the 4 categories	8.87 (0.000)	14.01 (0.000)	-0.30	-1.08	-6.0	-11.1	-18.7	-9.9
Hausman type - group private schools 10 categories by size	Main + avg prices in the 10 categories	8.86 (0.000)	13.91 (0.000)	-0.35	-1.12	-6.0	-11.1	-16.4	-9.7
The use of teacher costs	Exc. teacher costs	15.16 (0.000)	22.16 (0.000)	-0.37	-0.63	-6.0	-11.1	-15.6	-15.2

Notes: This table reports the consolidated results on key output measures from the multiple robustness checks on alternative instruments, namely, Hausman type (grouping private schools in categories by size), and excluding the teacher cost from the instrument set.

Table A.16: IV regression – Comparison of the main specification with the specification excluding some interactions

	GIRLS		BOYS	
	(1)	(2)	(3)	(4)
School fees	-0.136***	-0.151***	-0.043*	-0.045*
	[0.041]	[0.045]	[0.025]	[0.025]
School with toilets	0.122	0.237	0.280	0.227
	[0.361]	[0.364]	[0.237]	[0.234]
School with permanent	0.225	0.474*	0.182	0.096
classroom	[0.266]	[0.272]	[0.204]	[0.200]
Number of extra facilities	0.198***	0.172**	0.122**	0.122**
	[0.072]	[0.072]	[0.056]	[0.056]
Perc. of female teachers	0.592*	0.325	-0.747***	-0.580**
	[0.336]	[0.313]	[0.275]	[0.262]
Perc. of teachers with	0.251	0.021	0.186	0.154
at least 3 years of exp.	[0.321]	[0.342]	[0.283]	[0.272]
Perc. of teachers with	0.530	0.510	0.112	0.173
university degree	[0.449]	[0.431]	[0.330]	[0.329]
Student test score	1.443*	1.145	1.146*	1.218*
	[0.777]	[0.831]	[0.640]	[0.643]
Teacher absenteeism	0.036	0.034	0.003	-0.002
	[0.042]	[0.041]	[0.023]	[0.026]
Teacher test score	1.600	2.203**	1.051	0.754
	[1.073]	[1.062]	[0.709]	[0.753]
Perc. of mother with	-0.711**	-0.528	-0.381	-0.320
some education	[0.363]	[0.373]	[0.307]	[0.302]
Asset index	-0.056	-0.019	0.030	0.060
	[0.099]	[0.098]	[0.074]	[0.074]
Private	-0.254	-0.147	-0.911**	-0.983**
	[0.494]	[0.511]	[0.396]	[0.403]
Observations	511	511	522	522
R-squared	0.251	0.246	0.280	0.264
Specification	Main	Exc. some interactions	Main	Exc. some interactions

Notes: This table reports, for girls and boys, the comparison of the IV regression of the main specification (Columns (1) and (3)) with the specification excluding the interactions of school fees with mother education, age, and household distance to facilities, as well as all the interactions of school variables with household distance to facilities (Columns (2) and (4)). Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.17: Out-of -sample exercise

	Girls			Boys		
	Out-of-sample			Out-of-sample		
	Data	Prediction		Data	Prediction	
	(1)	(2)	(3)	(4)	(5)	(6)
% enrolled	68.8	72.6	68.8	78.9	83.3	78.9
% enrolled - non educated mother	61.2	65.4	61.2	75.6	81.0	75.6
% enrolled in private schools	26.4	27.2	27.2	27.8	28.7	28.8
total	22.3	21.1	21.2	19.0	21.4	21.4
non educated mother						
assets below median	20.0	20.5	20.5	22.7	25.4	25.4
% enrolled in private schools by age						
6 to 10	27.2	28.7	28.7	31.4	31.5	31.5
11 to 13	24.1	27.8	27.8	21.0	22.9	22.9
14 and 15	23.3	20.5	20.4	20.6	24.4	24.4
% enrolled in private schools by distance	21.9	23.7	23.8	20.1	25.5	25.5
distance above the median						
distance above the median and ages 14-15	16.7	17.2	17.2	17.9	23.7	23.7

Notes: Columns (1) and (4) show the moments for the out-of-sample data for girls and boys, respectively. In the out-of -sample exercise, we used the in-sample estimates for all parameters (β^o_{rkg},β^u_{kg}, $\overline{\gamma}_g$, γ_{rg}, γ^u_g, $\delta_{jtg}(\overline{\beta}_{kg})$) except the outside option per village (δ_{0tg}) since the data was splitted by village. Therefore, for the exercise we have used two different assumptions for the outside option per village (δ_{0tg}):
i) Columns (2) and (5) show the moments for the out-of-sample prediction, using all in-sample parameters described before and using the δ_{0tg} level from an estimation of the out-sample data.
ii) Columns (3) and (6) show the moments for the out-of-sample prediction, using all in-sample parameters described before, and using the δ_{0tg} level that matches the enrollment rate in the data.

Table A.18: First Stage - Excluding teacher costs from the instrument set

	Girls		Boys	
	(1)	(2)	(3)	(4)
VARIABLES	Main Spec.	Exc. Teacher Costs	Main Spec.	Exc. Teacher Costs
School with toilets x Private	139.120	138.883	144.454	142.729
	[142.904]	[142.883]	[129.484]	[129.625]
School with permanent classroom x Private	119.682	120.816	73.623	73.718
	[82.455]	[82.438]	[77.753]	[77.825]
Number of extra facilities x Private	76.336***	74.915***	55.101**	54.817**
	[26.773]	[26.778]	[25.292]	[25.317]
Percentage of female teachers x Private	-405.513***	-412.976***	-489.651***	-496.874***
	[110.219]	[110.031]	[100.414]	[100.496]
Perc. of teachers with	-302.712***	-299.171***	-330.681***	-329.681***
at least 3 years of exp. x Private	[112.674]	[112.566]	[104.588]	[104.682]
Perc. of teachers with university	497.712***	499.708***	496.554***	503.974***
degree x Private	[167.327]	[167.220]	[152.667]	[152.680]
Student test score x Private	1,276.462***	1,278.378***	1,161.171***	1,158.599***
	[282.099]	[282.066]	[257.407]	[257.650]
Teacher absenteeism x Private	-8.719	-8.597	-7.311	-6.829
	[14.936]	[14.928]	[13.557]	[13.562]
Teacher test score x Private	833.574**	861.047**	928.399**	941.449***
	[340.432]	[340.940]	[360.835]	[361.500]
Perc. of Mother with some	-275.935**	-273.707**	-204.596**	-203.459**
education x Private	[110.505]	[110.536]	[101.229]	[101.359]
Asset index x Private	73.824**	72.969**	58.044*	58.748*
	[36.264]	[36.275]	[33.860]	[33.885]
Private	-97.505	-162.397	98.086	77.048
	[619.764]	[620.762]	[589.218]	[590.157]
Teacher costs of other schools	0.055***	0.048***	0.065***	0.056***
in the same tehsil x Private	[0.012]	[0.012]	[0.009]	[0.009]
Total cost without rent x Private	1.131***		1.101***	
	[0.286]		[0.258]	
Total cost without rent and		9.424***		9.691***
teacher cost x Private		[1.370]		[1.305]
Number of schools within 2Km. x Private	9.601	9.026	15.474*	15.961**
	[9.341]	[9.059]	[8.112]	[7.829]
Number of extra facilities of the	-19.732	-24.174	3.442	-11.029
competitors x Private	[45.547]	[44.126]	[43.919]	[42.353]
Perc. of teachers with at least 3 years of	-401.302	-436.654	-205.967	-258.339
experience of the competitors x Private	[274.006]	[265.807]	[254.582]	[245.904]
Teacher test score of the	-120.081	387.110	-662.637	-349.785
competitors x Private	[556.874]	[545.917]	[533.566]	[517.422]
Asset index of the competitors x Private	216.201***	173.451**	90.416	55.644
	[70.027]	[68.314]	[66.356]	[64.276]
Observations	511	511	522	522
R-squared	0.671	0.691	0.711	0.730
10.15 15.16 15.93 22.16				
F-test (Instruments)	10.15	15.16	15.93	22.16
p-value	0.000	0.000	0.000	0.000

Notes: This table reports, for girls and boys, the comparison of the First stage regression of the main specification (Columns (1) and (3)) with the specification excluding teacher costs from the instrument set (columns (2) and (4)). Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.19: IV Regression - Excluding Teacher costs from the instrument set

	GIRLS		BOYS	
	(1)	(2)	(3)	(4)
School fees	-0.136***	-0.086**	-0.043*	-0.043*
	[0.041]	[0.040]	[0.025]	[0.024]
School with toilets	0.122	0.080	0.280	0.280
	[0.361]	[0.364]	[0.237]	[0.231]
School with permanent	0.225	0.185	0.182	0.182
classroom	[0.266]	[0.275]	[0.204]	[0.199]
Number of extra facilities	0.198***	0.168**	0.122**	0.122**
	[0.072]	[0.073]	[0.056]	[0.057]
Perc. of female teachers	0.592*	0.700**	-0.747***	-0.747***
	[0.336]	[0.340]	[0.275]	[0.278]
Perc. of teachers with	0.251	0.314	0.186	0.186
at least 3 years of exp.	[0.321]	[0.329]	[0.283]	[0.271]
Perc. of teachers with	0.530	0.337	0.112	0.113
university degree	[0.449]	[0.417]	[0.330]	[0.323]
Student test score	1.443*	1.048	1.146*	1.147*
	[0.777]	[0.782]	[0.640]	[0.614]
Teacher absenteeism	0.036	0.038	0.003	0.003
	[0.042]	[0.043]	[0.023]	[0.024]
Teacher test score	1.600	1.363	1.051	1.052
	[1.073]	[1.042]	[0.709]	[0.683]
Perc. of mother with	-0.711**	-0.633*	-0.381	-0.381
some education	[0.363]	[0.360]	[0.307]	[0.310]
Asset index	-0.056	-0.093	0.030	0.030
	[0.099]	[0.098]	[0.074]	[0.071]
Private	-0.254	-0.708	-0.911**	-0.909**
	[0.494]	[0.468]	[0.396]	[0.408]
Observations	511	511	522	522
R-squared	0.251	0.245	0.280	0.281
Instruments	Main	Excluding Teacher Costs	Main	Excluding Teacher Costs

Notes: This table reports, for girls and boys, the comparison of the IV regression of the main specification (Columns (1) and (3)) with specifications excluding teacher costs from the instrument set (columns (2) and (4)). Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.20: First Stage - Using hausman-style instrument as an additional instrument

	Girls		Boys	
VARIABLES	(1)	(2)	(3)	(4)
School with toilets x Private	130.739	137.534	125.445	141.102
	[145.927]	[144.350]	[131.572]	[131.071]
School with permanent classroom x Private	120.893	119.728	74.411	74.515
	[82.637]	[82.540]	[77.784]	[78.001]
Number of extra facilities x Private	74.530***	76.048***	50.760*	54.060**
	[27.512]	[27.030]	[25.845]	[26.033]
Percentage of female teachers x Private	-406.025***	-405.263***	-488.842***	-490.003***
	[110.336]	[110.373]	[100.451]	[100.532]
Perc. of teachers with	-306.379***	-303.484***	-341.456***	-332.278***
at least 3 years of exp. x Private	[113.486]	[113.181]	[105.439]	[105.102]
Perc. of teachers with university	503.249***	498.877***	507.480***	498.636***
degree x Private	[168.568]	[168.099]	[153.293]	[153.296]
Student test score x Private	1,277.086***	1,274.909***	1,168.447***	1,159.840***
	[282.370]	[283.019]	[257.642]	[257.773]
Teacher absenteeism x Private	-8.688	-8.669	-7.432	-7.284
	[14.951]	[14.964]	[13.563]	[13.571]
Teacher test score x Private	816.160**	829.806**	888.420**	920.226**
	[345.998]	[343.861]	[364.211]	[364.307]
Perc. of Mother with some	-272.526**	-275.207**	-196.177*	-203.636**
education x Private	[111.231]	[110.972]	[101.778]	[101.481]
Asset index x Private	72.267**	73.420**	53.474	57.397*
	[36.693]	[36.633]	[34.324]	[34.102]
Private	-127.265	-103.108	28.530	88.308
	[628.770]	[624.144]	[595.446]	[592.531]
Average price in other schools (4 cat) x Private	5.795		14.047	
	[19.978]		[17.083]	
Average price in other schools (10 cat) x Private		1.172		2.420
		[14.288]		[14.093]
Teacher costs of other schools	0.055***	0.055***	0.065***	0.065***
in the same tehsil x Private	[0.012]	[0.012]	[0.009]	[0.009]
Total cost without rent x Private	1.137***	1.133***	1.115***	1.106***
	[0.286]	[0.288]	[0.259]	[0.260]
Number of schools within 2Km. x Private	9.702	9.580	15.694*	15.506*
	[9.356]	[9.354]	[8.119]	[8.122]
Number of extra facilities of the	-21.512	-20.101	-0.993	2.640
competitors x Private	[46.001]	[45.815]	[44.263]	[44.209]
Perc. of teachers with at least 3 years of	-399.212	-400.371	-187.297	-205.637
experience of the competitors x Private	[274.357]	[274.519]	[255.675]	[254.836]
Teacher test score of the	-122.735	-120.786	-688.786	-663.901
competitors x Private	[557.469]	[557.505]	[534.685]	[534.133]
Asset index of the competitors x Private	220.697***	217.238***	102.379	91.977
	[71.786]	[71.229]	[67.953]	[67.039]
Observations	511	511	522	522
R-squared	0.671	0.671	0.711	0.711
F- test (instrument)	8.87	8.86	14.01	13.91
p-value	0.0000	0.0000	0.0000	0.0000

Notes: This table reports, for girls and boys, the First stage regression using hausman-style instruments as an additional instrument. In addition to our main specification, Columns (1) and (3) uses 4 categories of the average price in other schools, Columns (2) and (4) uses 10 categories of the average price in other schools.
Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.21: IV Regression - Using hausman-style instrument as an additional instrument

	GIRLS			BOYS		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
School fees	-0.136***	-0.132***	-0.135***	-0.043*	-0.036	-0.041*
	[0.041]	[0.042]	[0.042]	[0.025]	[0.025]	[0.024]
School with toilets	0.122	0.118	0.120	0.280	0.274	0.279
	[0.361]	[0.396]	[0.381]	[0.237]	[0.234]	[0.236]
School with permanent	0.225	0.221	0.223	0.182	0.178	0.181
classroom	[0.266]	[0.270]	[0.276]	[0.204]	[0.206]	[0.206]
Number of extra facilities	0.198***	0.195***	0.197***	0.122**	0.119**	0.121**
	[0.072]	[0.071]	[0.070]	[0.056]	[0.057]	[0.057]
Perc. of female teachers	0.592*	0.603*	0.595*	-0.747***	-0.730***	-0.743***
	[0.336]	[0.337]	[0.333]	[0.275]	[0.264]	[0.273]
Perc. of teachers with	0.251	0.257	0.252	0.186	0.196	0.189
at least 3 years of exp.	[0.321]	[0.331]	[0.334]	[0.283]	[0.267]	[0.275]
Perc. of teachers with	0.530	0.511	0.524	0.112	0.084	0.105
university degree	[0.449]	[0.413]	[0.428]	[0.330]	[0.336]	[0.325]
Student test score	1.443*	1.405*	1.431*	1.146*	1.092*	1.133*
	[0.777]	[0.776]	[0.790]	[0.640]	[0.637]	[0.651]
Teacher absenteeism	0.036	0.036	0.036	0.003	0.003	0.003
	[0.042]	[0.042]	[0.041]	[0.023]	[0.023]	[0.023]
Teacher test score	1.600	1.577	1.594	1.051	1.018	1.043
	[1.073]	[1.069]	[1.063]	[0.709]	[0.722]	[0.735]
Perc. of mother with	-0.711**	-0.704**	-0.709**	-0.381	-0.371	-0.378
some education	[0.363]	[0.360]	[0.340]	[0.307]	[0.314]	[0.311]
Asset index	-0.056	-0.060	-0.057	0.030	0.025	0.029
	[0.099]	[0.099]	[0.099]	[0.074]	[0.076]	[0.078]
Private	-0.254	-0.298	-0.267	-0.911**	-0.986**	-0.929**
	[0.494]	[0.498]	[0.478]	[0.396]	[0.402]	[0.401]
Observations	511	511	511	522	522	522
R-squared	0.251	0.250	0.251	0.280	0.280	0.280
Instruments						
Main Specification	yes	yes	yes	yes	yes	yes
Hausman Style						
4 categories	no	yes	no	no	yes	no
10 categories	no	no	yes	no	no	yes

Notes: This table reports, for girls and boys, the comparison of the IV regression of the main specification (Columns (1) and (4)) with specifications including hausman-style instrument as an extra instrument: columns (2) and (5) show the inclusion of 4 categories of the average price in other schools in the same district, and columns (3) and (6) uses 10 categories of the average price in other schools in the same district. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.22: IV Regression - school size as a school characteristic - girls

	25th perc		mean		75th perc	
	IV	IV	IV	IV	IV	IV
School fees	-0.167***	-0.152***	-0.136***	-0.122***	-0.120***	-0.105***
	[0.041]	[0.041]	[0.041]	[0.041]	[0.041]	[0.041]
School Size		0.003***		0.003***		0.003***
		[0.001]		[0.001]		[0.001]
School with toilets	0.164	0.002	0.122	-0.038	0.125	-0.037
	[0.379]	[0.369]	[0.361]	[0.364]	[0.376]	[0.353]
School with permanent classroom	0.300	0.245	0.225	0.170	0.218	0.163
	[0.277]	[0.266]	[0.266]	[0.271]	[0.270]	[0.272]
Number of extra facilities	0.178**	0.104	0.198***	0.125	0.197***	0.123
	[0.071]	[0.070]	[0.072]	[0.070]	[0.073]	[0.070]
Percentage of female teachers	0.496	0.562*	0.592*	0.657*	0.610*	0.676**
	[0.341]	[0.342]	[0.336]	[0.344]	[0.342]	[0.347]
Perc. of teachers with at least 3 years of experience	0.137	-0.000	0.251	0.117	0.304	0.168
	[0.319]	[0.336]	[0.321]	[0.328]	[0.343]	[0.335]
Perc. of teachers with university degree	0.432	0.398	0.530	0.497	0.553	0.520
	[0.430]	[0.421]	[0.449]	[0.431]	[0.432]	[0.421]
Student test score	1.447*	1.215*	1.443*	1.216*	1.394*	1.163
	[0.759]	[0.767]	[0.777]	[0.781]	[0.783]	[0.788]
Teacher absenteeism	0.036	0.022	0.036	0.022	0.039	0.025
	[0.042]	[0.041]	[0.042]	[0.042]	[0.041]	[0.042]
Teacher test score	1.620	1.086	1.600	1.082	1.574	1.044
	[1.028]	[1.031]	[1.073]	[1.106]	[1.096]	[1.117]
Perc. of mother with some education	-0.976***	-0.997***	-0.711**	-0.733**	-0.620*	-0.642*
	[0.333]	[0.365]	[0.363]	[0.335]	[0.348]	[0.349]
Asset index	-0.064	-0.072	-0.056	-0.065	-0.046	-0.054
	[0.099]	[0.094]	[0.099]	[0.099]	[0.098]	[0.101]
Private	-0.357	-0.403	-0.254	-0.298	-0.211	-0.256
	[0.483]	[0.485]	[0.494]	[0.484]	[0.465]	[0.479]

Notes: This table shows estimates of equation (6) in a specification where an indicator of school size is included in the model as another attribute. We compute the 25th, and 75th of maternal education, household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of our preferred specification. Remaining columns (2, 4, and 6) report the results where an indicator of school size is included in the model as another attribute. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.23: IV Regression - school size as a school characteristic - boys

	25th perc		mean		75th perc	
School fees	-0.051**	-0.025	-0.043*	-0.020	-0.039	-0.018
	[0.025]	[0.025]	[0.025]	[0.025]	[0.025]	[0.026]
School Size		0.004***		0.005***		0.005***
		[0.001]		[0.001]		[0.001]
School with toilets	0.162	0.007	0.280	0.123	0.329	0.177
	[0.238]	[0.229]	[0.237]	[0.235]	[0.251]	[0.227]
School with permanent classroom	0.107	0.379**	0.182	0.463***	0.215	0.498***
	[0.198]	[0.199]	[0.204]	[0.209]	[0.203]	[0.215]
Number of extra facilities	0.115**	-0.014	0.122**	-0.010	0.127**	-0.005
	[0.055]	[0.056]	[0.056]	[0.056]	[0.057]	[0.058]
Percentage of female teachers	-0.852***	-0.970***	-0.747***	-0.856***	-0.675**	-0.788***
	[0.269]	[0.255]	[0.275]	[0.275]	[0.271]	[0.278]
Perc. of teachers with at least 3 years of experience	0.091	-0.211	0.186	-0.118	0.240	-0.067
	[0.286]	[0.280]	[0.283]	[0.276]	[0.281]	[0.277]
Perc. of teachers with university degree	0.171	-0.001	0.112	-0.054	0.091	-0.074
	[0.329]	[0.331]	[0.330]	[0.330]	[0.310]	[0.334]
Student test score	0.996	0.375	1.146*	0.537	1.231*	0.626
	[0.644]	[0.655]	[0.640]	[0.655]	[0.642]	[0.649]
Teacher absenteeism	0.002	-0.019	0.003	-0.019	0.004	-0.018
	[0.024]	[0.024]	[0.023]	[0.023]	[0.024]	[0.024]
Teacher test score	1.104	0.126	1.051	0.041	0.992	0.001
	[0.726]	[0.693]	[0.709]	[0.716]	[0.767]	[0.728]
Perc. of mother with some education	-0.368	-0.172	-0.381	-0.157	-0.391	-0.149
	[0.298]	[0.300]	[0.307]	[0.315]	[0.303]	[0.313]
Asset index	-0.009	-0.010	0.030	0.031	0.052	0.055
	[0.075]	[0.075]	[0.074]	[0.078]	[0.076]	[0.075]
Private	-0.929**	-0.455	-0.911**	-0.414	-0.909**	-0.395
	[0.405]	[0.387]	[0.396]	[0.388]	[0.387]	[0.417]

Notes: This table shows estimates of equation (6) in a specification where an indicator of school size is included in the model as another attribute. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of our preferred specification. Remaining columns (2, 4, and 6) report the results where an indicator of school size is included in the model as another attribute. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.24: IV Regression - Comparison of the main specification (without endog. of peers) with the model without peer variables - girls

	25th perc.		mean		75th perc.	
	Main	No peers	Main	No peers	Main	No peers
School fees	-0.167***	-0.163***	-0.136***	-0.131***	-0.120***	-0.114***
	[0.041]	[0.035]	[0.041]	[0.034]	[0.041]	[0.036]
School with toilets	0.164	0.138	0.122	0.116	0.125	0.126
	[0.379]	[0.368]	[0.361]	[0.360]	[0.370]	[0.373]
School with permanent classroom	0.300	0.255	0.225	0.188	0.218	0.184
	[0.277]	[0.277]	[0.266]	[0.284]	[0.270]	[0.272]
Number of extra facilities	0.178**	0.169**	0.198***	0.193***	0.197***	0.195***
	[0.071]	[0.072]	[0.072]	[0.071]	[0.073]	[0.070]
Percentage of female teachers	0.496	0.225	0.592*	0.373	0.610*	0.413
	[0.341]	[0.324]	[0.336]	[0.319]	[0.342]	[0.341]
Perc. of teachers with at least 3 years of experience	0.137	0.165	0.251	0.283	0.304	0.339
	[0.319]	[0.331]	[0.321]	[0.329]	[0.343]	[0.338]
Perc. of teachers with university degree	0.432	0.483	0.530	0.577	0.553	0.601
	[0.430]	[0.430]	[0.449]	[0.430]	[0.432]	[0.438]
Student test score	1.447*		1.443*		1.394*	
	[0.759]		[0.777]		[0.783]	
Teacher absenteeism	0.036	0.046	0.036	0.045	0.039	0.048
	[0.042]	[0.041]	[0.042]	[0.041]	[0.041]	[0.042]
Teacher test score	1.620	1.756	1.600	1.703	1.574	1.662
	[1.028]	[1.140]	[1.073]	[1.059]	[1.096]	[1.098]
Perc. of mother with some education	-0.976***		-0.711**		-0.620*	
	[0.333]		[0.363]		[0.348]	
Asset index	-0.064		-0.056		-0.046	
	[0.099]		[0.099]		[0.098]	
Private	-0.357	-0.444	-0.254	-0.300	-0.211	-0.246
	[0.483]	[0.462]	[0.494]	[0.464]	[0.465]	[0.473]

Notes: This table shows estimates of equation (6) in a specification where there are no peer variables and where peer variables are taken as exogenous. We compute the 25th and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of the specification without endogeneity. Remaining columns (2, 4, 6) report the results where peer school variables were excluded from the model. Bootstrapped standard errors in brackets.
* Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.25: IV Regression - Comparison of the main specification (without endog. of peers) with the model without peer variables - boys

	25th perc.		mean		75th perc.	
	Main	No peers	Main	No peers	Main	No peers
School fees	-0.051**	-0.052**	-0.043*	-0.042*	-0.039	-0.036*
	[0.025]	[0.021]	[0.025]	[0.022]	[0.025]	[0.022]
School with toilets	0.162	0.156	0.280	0.274	0.329	0.326
	[0.238]	[0.226]	[0.237]	[0.229]	[0.251]	[0.226]
School with permanent classroom	0.107	0.120	0.182	0.194	0.215	0.226
	[0.198]	[0.204]	[0.204]	[0.209]	[0.203]	[0.210]
Number of extra facilities	0.115**	0.115**	0.122**	0.128**	0.127**	0.136**
	[0.055]	[0.055]	[0.056]	[0.055]	[0.057]	[0.055]
Percentage of female teachers	-0.852***	-0.922***	-0.747***	-0.792***	-0.675**	-0.707***
	[0.269]	[0.258]	[0.275]	[0.287]	[0.271]	[0.264]
Perc. of teachers with at least 3 years of experience	0.091	0.079	0.186	0.175	0.240	0.231
	[0.286]	[0.269]	[0.283]	[0.272]	[0.281]	[0.279]
Perc. of teachers with university degree	0.171	0.199	0.112	0.145	0.091	0.126
	[0.329]	[0.316]	[0.330]	[0.324]	[0.310]	[0.313]
Student test score	0.996		1.146*		1.231*	
	[0.644]		[0.640]		[0.642]	
Teacher absenteeism	0.002	0.003	0.003	0.005	0.004	0.007
	[0.024]	[0.022]	[0.023]	[0.024]	[0.024]	[0.023]
Teacher test score	1.104	1.353**	1.051	1.309*	0.992	1.249*
	[0.726]	[0.738]	[0.709]	[0.749]	[0.767]	[0.734]
Perc. of mother with some education	-0.368		-0.381		-0.391	
	[0.298]		[0.307]		[0.303]	
Asset index	-0.009		0.030		0.052	
	[0.075]		[0.074]		[0.076]	
Private	-0.929**	-0.807**	-0.911**	-0.777**	-0.909**	-0.774**
	[0.405]	[0.385]	[0.396]	[0.390]	[0.387]	[0.388]

Notes: This table shows estimates of equation (6) in a specification where there are no peer variables and where peer variables are taken as exogenous. We compute the 25th and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of the specification without endogeneity. Remaining columns (2, 4, 6) report the results where peer school variables were excluded from the model. Bootstrapped standard errors in brackets.
* Significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.26: Compensating variation - after trimming the bottom and top 5% of the distribution

Panel A - Average Compensating Variation (in U.S. dollars)

	All		Affected by the Policy	
	GIRLS	BOYS	GIRLS	BOYS
No Private schools	2.3	8.0	4.1	13.6
Voucher Program	-3.5	-3.1	-	-

Panel B - Total Compensating Variation (in thousand U.S. dollars)

	GIRLS	BOYS	TOTAL	DIF
No Private schools	84.1	400.0	484.1	-315.9
Voucher Program	-132.1	-155.9	-288.0	23.8

Notes: In this table we present changes in welfare using the average of the compensating variation after trimming the bottom and top 5% of the distribution of this variable.
Panel A shows the estimates of the median compensating variation (in U.S. dollars) for a policy that forces all private schools to shut down and from the introduction of vouchers. Columns (1) and (2) show the results for everyone, and columns (3) and (4) display the estimates for those affected by the policy intervention. In the "no private schools" scenario those not affected by the policy intervention have no change in their consumer surplus. In Panel B we obtain the total welfare change, in U.S. thousand dollars, taking the median compensating variation across the sample and multiply by the total number of students enrolled in the regions from our sample in rural Punjab. As before, Columns (1) and (2) show the results for everyone. Column (3) presents the sum of welfare change for girls and boys, and column (4) displays the difference between boys and girls.
1 U.S. dollars ≈ 85.6 Pakistani Rupees.

Table A.27: Compensating variation - one private school

Panel A - Median Compensating Variation (in U.S. dollars)

	All		Affected by the Policy	
	GIRLS	BOYS	GIRLS	BOYS
Only one Private school	0.3	1.2	0.9	3.2

Panel B - Total Compensating Variation (in thousand U.S. dollars)

	GIRLS	BOYS	TOTAL	DIF
Only one Private school	11.4	58.4	69.8	-46.9

Notes: In this simulation we present the changes in welfare for a policy where we close all but one private school in each village. The private school that is allowed to be open in this simulation has the average characteristics of all private schools in the village. Panel A shows estimates of the median compensating variation, in U.S. dollars, separately for boys and girls. Columns (1) and (2) show the results for everyone, and columns (3) and (4) display the estimates for those affected by the policy intervention. In Panel B we obtain the total welfare change, in U.S. thousand dollars, taking the median compensating variation across the sample and multiply by the total number of students enrolled in the regions from our sample in rural Punjab. As before, Columns (1) and (2) show the results for everyone, and columns (3) and (4) display the estimates for those affected by the policy intervention.
1 U.S. dollars $\approx$ 85.6 Pakistani Rupees.

Table A.28: Compensating variation - by household type

GIRLS

Median (in U.S. dollars)

	Mother Education		Expenditure		Household distance to facilities	
	No education	At least some education	below mean	above mean	below mean	above mean
No Private schools	0.9	4.8	0.8	2.0	2.0	0.7
Voucher program	-2.4	-4.3	-2.4	-3.0	-3.2	-2.1

BOYS

Median (in U.S. dollars)

	Mother Education		Expenditure		Household distance to facilities	
	No education	At least some education	below mean	above mean	below mean	above mean
No Private schools	3.5	19.1	3.5	6.3	6.9	2.6
Voucher program	-1.9	-4.4	-1.8	-2.9	-3.1	-1.4

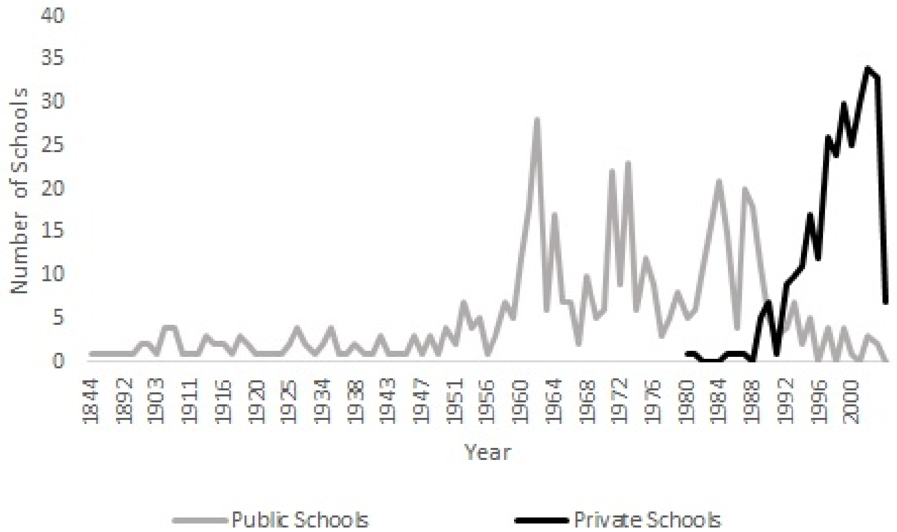
Notes: This table shows the changes in welfare by household type (mother education, expenditure, and household distance to facilities) for a policy that forces all private schools to shut down ("no private schools") and from the introduction of vouchers. For both, girls and boys, the table shows the median compensating variation (in U.S. dollars).
We use compensating variation to measure changes in a household's income that equates utility across two states: a benchmark state, which is the status quo, and the alternative state, which is the environment without private schools and the scenario where school fees are equal to zero. For example, it corresponds to the amount of income required to compensate a given household for the elimination of private schools.
1 U.S. dollars $\approx$ 85.6 Pakistani Rupees.

Table A.29: Welfare Change - Simulation that forces private schools to shut down

	Girls	Boys
Welfare Change	1.3	5.5
Welfare change from difference in:		
Observed attributes	0.9	4.1
Logit error	0.4	1.4

Notes:
In the spirit of Petrin (2002) we have simulated the welfare change and the decomposition into two components. One component is related to the observed characteristics entering the utility function. The second component is related to the idiosyncratic logit taste term. The decomposition of compensation is the average difference in the value of observed and unobserved characteristics. For each simulation we draw a random logit error from the extreme value distribution. It should be highlighted that the total change in welfare is similar to the one calculated in the counterfactual exercise.

Figure A.1: Number of schools by year of construction



Notes: This figure represents the number of public and private schools by year of construction in the LEAPS dataset.

B Appendix - BLP (First step)

In this part of the Appendix, we discuss the estimation procedure of the first step. The coefficients of this model can be estimated using the algorithms described in Berry et al. (1995) and Berry et al. (2004), which we adapt slightly to the type of data we have available.

The first step entails estimating $\delta_{jtg}, \beta^o_{rkg}, \beta^u_{kg}, \overline{\gamma}_g, \gamma_{rg}, \gamma^u_g$ by maximum likelihood, including a contraction mapping to obtain δ_{jtg}.

Under the assumption that ε_{ijtg} has an extreme value Type I distribution, the probability of household i choose school j for children of gender g (i.e. the probability of $u_{ijtg} > u_{iqtg}, \forall j \neq q$) is

$$\begin{aligned} P_{ijtg} &= \Pr(y_i = j | z_{itg}, x_{jtg}, d_{ijtg}, v_{itg}, \beta_g, \gamma_g) \\ &= \frac{\exp(\delta_{jtg} + \sum_{k=1}^{K}\sum_{r=1}^{R} x_{jktg} z_{irtg} \beta^o_{rk} + \overline{\gamma} d_{ijtg} + \sum_{r=1}^{R} d_{ijtg} z_{irtg} \gamma_{rg} + \sum_{k=1}^{K} x_{jktg} v_{itg} \beta^u_{kg} + d_{ijtg} v_{itg} \gamma^u_g)}{\sum_{q=0}^{J} \exp(\delta_{qtg} + \sum_{k=1}^{K}\sum_{r=1}^{R} x_{qktg} z_{irtg} \beta^o_{rkg} + \overline{\gamma}_g d_{iqtg} + \sum_{r=1}^{R} d_{iqtg} z_{irtg} \gamma_{rg} + \sum_{k=1}^{K} x_{qktg} v_{itg} \beta^u_{kg} + d_{iqtg} v_{itg} \gamma^u_g)} \end{aligned} \quad (9)$$

and the likelihood function is given by:

$$L(\beta_g, \gamma_g) = \prod_{j=0}^{J} \prod_{i \in A_j} P_{ijtg}$$

and the log-likelihood by:

$$LL(\beta_g, \gamma_g) = \sum_{j=0}^{J} \sum_{i \in A_j} \ln(P_{ijtg})$$

where, the set of households that choose school j is given by

$$A_{jtg}(x_{jtg}, d_{ijtg}; \delta_{jtg}, \beta^o_{rkg}, \overline{\gamma}_g, \gamma_{rg}) = \{(\varepsilon_{i0tg}, ..., \varepsilon_{iJtg}) | u_{ijtg} > u_{iltg}, \forall j \neq l\}$$

As v_{itg} is unobserved and follows a standard normal distribution, the expected value of the probability unconditional on v_{itg} is given by:

$$\widehat{P}_{ijtg}(z_{itg}, x_{jtg}, d_{ijtg}, \beta_g, \gamma_g) = \int P_{ijtg} f(v) d(v)$$

To calculate the log-likelihood function we approximate this integral using simulation and then sum the log of this probability over students i of gender g.

Let $\widetilde{P}_{iqtg}$ be a simulated approximation to P_{iqtg}. The simulated choice probability is given by

$$\widetilde{P}_{ijtg} = \sum_{n=1}^{ND} \frac{\exp(\delta_{jtg} + \sum_{k=1}^{K} \sum_{r=1}^{R} x_{jktg} z_{irtg} \beta_{rkg}^{o} + \overline{\gamma}_g d_{ijtg} + \sum_{r=1}^{R} d_{ijtg} z_{irtg} \gamma_{rg} + \sum_{k=1}^{K} x_{jktg} v_{itgn} \beta_{kg}^{u} + d_{ijtg} v_{itgn} \gamma_g^{u})}{\sum_{q=0}^{J} \exp(\delta_{qtg} + \sum_{k=1}^{K} \sum_{r=1}^{R} x_{qktg} z_{irtg} \beta_{rkg}^{o} + \overline{\gamma}_g d_{iqtg} + \sum_{r=1}^{R} d_{iqtg} z_{irtg} \gamma_{rg} + \sum_{k=1}^{K} x_{qktg} v_{itgn} \beta_{kg}^{u} + d_{iqtg} \gamma_g^{u})} \tag{10}$$

for random draws $v_{itgn}, n = 1, ..., ND$.

The Simulated log-likelihood function is given by

$SLL(\beta, \gamma) = \sum_{j=0}^{J} \sum_{i \in A_j} \ln(\widetilde{P}_{ijtg})$

This procedure is the same as Maximum Likelihood except that simulated probabilities are used instead of the exact probabilities[31].

Partially differentiating (10) with respect to δ_{qtg} we get

$$\frac{\partial SLL}{\partial \delta_{qtg}} = \sum_{\substack{j=0 \\ j \neq q}}^{J} \sum_{i \in A_j} \frac{1}{\widetilde{P}_{ijtg}} \frac{\partial \widetilde{P}_{ijtg}}{\partial \delta_{qtg}} + \sum_{i \in A_q} \frac{1}{\widetilde{P}_{iqtg}} \frac{\partial \widetilde{P}_{iqtg}}{\partial \delta_{qtg}} \tag{11}$$

Given that

$$\frac{\partial \widetilde{P}_{iqtg}}{\partial \delta_{qtg}} = \widetilde{P}_{iqtg}(1 - \widetilde{P}_{iqtg}) \tag{12}$$

$$\frac{\partial \widetilde{P}_{ijtg}}{\partial \delta_{qtg}} = -\widetilde{P}_{iqtg} \widetilde{P}_{ijtg}, j \neq q \tag{13}$$

the FOC with respect to δ_{qtg} of the MSL problem becomes:

[31] See Train (2009) for further details.

$$\frac{\partial SLL}{\partial \delta_{qtg}} = \sum_{i \in A_q} 1 - \sum_{j=0}^{J} \sum_{i \in A_j} \widetilde{P}_{iqtg}$$
$$= N_q - \sum_{i=1}^{N} \widetilde{P}_{iqtg} = 0$$

Dividing by N we get:

$$sh_{qg} - \frac{1}{N} \sum_{i=1}^{N} \widetilde{P}_{iqtg} = 0 \quad (14)$$

where sh_{qg} is the share of students that attend school q and N is the total number of students[32].

This condition implies that the estimated δ_{jtg} has to guarantee that the empirical share of students attending school j has to be equal to the average probability that a student attends this school.

In order to find estimates for the parameters of interest we need to iterate over

$$\delta_{qtg}^{t+1} = \delta_{qtg}^{t} - \left[\log(sh_{qg}) - \log(\frac{1}{N} \sum_{i=1}^{N} \widetilde{P}_{iqtg}) \right] \quad (15)$$

Each iteration over (15) requires a new calculation of the probabilities in (10)

[32]The procedure is done for each gender. The market is the combination of village t and gender g.

C Appendix - The endogeneity of peer characteristics

In our main set of results we do not consider explicitly the endogeneity of a second set of school attributes: the average test scores, maternal education, and household assets of other students in the school. These are measures of peer group "quality", and therefore they are likely to be important determinants of school choice. They are extensively discussed in the literature on school (and neighborhood) choice (e.g. Bayer et al. (2007)).

In principle, one would need to fully specify and solve the equilibrium model governing the sorting of students to schools, taking into account that each household's decision depends on the decision of every other household in the village. Bayer and Timmins (2007) propose a simpler IV procedure to estimate the individuals' valuation of peer attributes in a school, which is consistent with an equilibrium model, but does not require the full solution of a model (even in cases where there are likely to be multiple equilibria).[33]

Their paper considers models of sorting of individuals across locations, where a central location attribute is the proportion of individuals choosing that location. Their goal is to estimate the individual's valuation of this characteristic. They specify a simple equilibrium sorting model which suggests that, as long as individuals only obtain utility from the characteristics of the location they choose, we can instrument the proportion of individuals choosing a particular

[33] Bayer and Timmins (2007) discuss the circumstances under which this procedure is robust to the possibility of multiple sorting equilibria, which arise naturally in settings with social interactions and local spillovers, such as the one we consider. When the number of individuals in each market is large, the probability that each equilibrium is selected conditional on the distribution of preferences and household characteristics in a given market is orthogonal to a particular individual's preferences and characteristics. Therefore, the choice model can be estimated conditional on the equilibrium selected in each market, regardless of which one was chosen. This simplifies estimation and the assumption on which it is based is reasonable in villages of considerable size, such as the ones studied in this paper.

location using the non-peer (exogenous) attributes of other locations in the same market.

Starting from one particular location, if the attributes of its close competitors are very attractive, the demand for competitor locations will increase, and the demand for this location will fall. This means that competitor attributes will be good predictors of the proportion of individuals at each location. If, in addition, exogenous attributes of competitors do not directly affect the utility of those choosing this particular location, then the exclusion restriction is likely to be satisfied. One could potentially use any function of competitors' attributes as instruments, and following the literature on optimal instruments, Bayer and Timmins (2007) suggest using the predicted probability that one chooses a particular location, after restricting the coefficient on the (endogenous) peer variable to zero.

Our setting is slightly different than the one in Bayer and Timmins (2007). The peer attributes we care about are not the proportion of students attending a specific school, but the average characteristics of these students. We modify the main ideas of Bayer and Timmins (2007) as follows. Starting from a particular school, the school (non-peer) attributes of its competitors in the same market are likely to affect the composition of the student body in this school. In addition, the attributes of competitor schools will be valid instruments unless they directly affect the utility each household derives from a given school. Therefore, as in Bayer and Timmins (2007), we propose to simulate the equilibrium sorting of households to schools when the valuation of peer attributes is restricted to be equal to zero, and use the predicted average peer characteristics in each school, resulting from this simulation, as an instrument for the actual average peer characteristics in the school.

To be precise, we start by estimating the model of equations (5) and (6), ignoring the endogeneity of peer attributes. We then set equal to zero the coefficients ($\overline{\beta}_{kg}$,β^{o}_{rkg},β^{u}_{kg}) on all peer characteristics in each school (average student test scores, average maternal education, average student assets). In addition, we set the school specific unobservable (ξ_{jtg}) also equal to zero. We simulate the proportion of students attending each school once these restrictions are imposed, as well as predict their average test scores, the average education of their mothers, and their average assets.

Let $\tilde{\pi}_{ijtg}$ denote the simulated probability that individual i (of gender g in village t) chooses school j, in the absence of peer variables and school unobservables, and given the household's characteristics and the remaining school attributes. Then, for each peer characteristic p_{iptg}, we compute $\tilde{p}_{jptg} = \frac{\sum_{i=1}^{N_{tg}} p_{iptg}\tilde{\pi}_{ijtg}}{\sum_{i=1}^{N_{tg}} \tilde{\pi}_{ijtg}}$, which is the simulated value of peer attribute p in school j (in village t, and considering only gender g), where N_{tg} is the number of families with children of gender g in village t. Finally we can use these predicted values as instruments for the actual peer variables in equation (6), which includes also as regressors the non-peer attributes of each school.

In addition, to increase the power of our estimates, we compute the predicted values of the peer variables for all other schools in the village, giving us additional functions of the instruments which we can use to predict peer characteristics in each school. Then we estimate a weighted average of these values, using as weights the (relative) distance between a school and each of its competitors. We expect the weighted average of predicted peer attributes in competitor schools to be negatively related to the value of peer variables in a given school. For example, if a village has two schools, as we increase the average value of maternal education in one school, we decrease it in the other school.

Formally, let e_{jlt} be the distance between schools j and l, both in village t. Then, for each peer characteristic p, we compute $\tilde{q}_{jptg} = \frac{\sum_{l=1}^{J_t} e_{jlt}\tilde{p}_{lptg}}{\sum_{l=1}^{J_t} e_{jlt}}$. We use $\tilde{q}_{jptg}$, in addition to $\tilde{p}_{jptg}$, as an instrument for the corresponding peer variable in equation (6).

C.1 First Stage Regressions of peer variables

Table C.1 shows the first stage regressions of peer variables (student test score, mother education, and assets on the predicted value of the peer variables in the school and on the predicted value of the peer variable in competitor schools, weighted by distance to each competitor and other school attributes).

There we also see that non-fee (and non-peer) attributes of other schools predict peer variables, especially after we use the optimal instrument proposed by adapting the procedure in Bayer and Timmins (2007): $\tilde{p}_{jptg}$ (the predicted value of the peer variable p in school j) and $\tilde{q}_{jptg}$ (the predicted value of the peer variable p in competitor schools, weighted by distance from school j to each competitor school).[34] For girls, we can predict average maternal education and average wealth of students using these instruments, but not average test scores of other students. For boys, they are good predictors of all three variables. This means that we may have difficulty estimating the valuation of peer test scores for parents of girls.

[34]Table C.1 shows that the coefficient on $\tilde{p}_{jptg}$ is positive, indicating that the higher the predicted value of the peer variable in the school, based on a model with only exogenous school attributes, the higher the actual value of the peer variable in the school. The coefficient on $\tilde{q}_{jptg}$ is negative, indicating that the value of the peer variables in the school decline with the predicted value of peer variables in competitor schools. Therefore, the signs of the coefficients on these two variables are as expected.

Table C.1: First stage - peer variables equations

		GIRLS			BOYS	
	Student Test	Mother Education	Asset Index	Student Test	Mother Education	Asset Index
Private	0.064	0.422	0.753	0.087	0.241	0.935
	[0.127]	[0.260]	[0.962]	[0.116]	[0.226]	[0.889]
School with toilets x Private	0.016	0.145*	0.350	-0.018	0.198***	0.794***
	[0.036]	[0.082]	[0.302]	[0.034]	[0.072]	[0.286]
School with permanent classroom x Private	0.002	-0.154**	-0.222	0.013	-0.099	-0.214
	[0.032]	[0.071]	[0.261]	[0.029]	[0.060]	[0.241]
Number of extra facilities x Private	0.010	0.001	0.080	0.006	0.010	0.028
	[0.008]	[0.018]	[0.067]	[0.008]	[0.016]	[0.065]
Percentage of female teachers x Private	0.010	0.048	-0.929***	-0.014	0.125*	-0.431
	[0.032]	[0.072]	[0.270]	[0.034]	[0.072]	[0.289]
Perc. of teachers with	-0.002	-0.039	0.319	0.013	-0.072	0.559
at least 3 years of exp. x Private	[0.043]	[0.096]	[0.353]	[0.040]	[0.086]	[0.342]
Perc. of teachers with university	0.019	0.040	0.683*	0.028	-0.061	0.709*
degree x Private	[0.046]	[0.103]	[0.381]	[0.044]	[0.093]	[0.370]
Teacher absenteeism x Private	0.006*	-0.002	0.069**	0.009**	-0.001	0.051
	[0.004]	[0.008]	[0.030]	[0.004]	[0.008]	[0.032]
Teacher test score x Private	0.056	-0.343	-0.234	0.015	-0.203	-1.378
	[0.118]	[0.264]	[0.975]	[0.111]	[0.236]	[0.935]
School with toilets	0.016	0.000	-0.024	0.031**	-0.049*	-0.384***
	[0.018]	[0.040]	[0.148]	[0.014]	[0.030]	[0.118]
School with permanent	0.011	0.124**	0.208	-0.004	0.061	0.095
classroom	[0.026]	[0.057]	[0.211]	[0.021]	[0.046]	[0.183]
Number of extra facilities	0.002	0.014	0.058	0.005	0.006	0.121***
	[0.006]	[0.013]	[0.046]	[0.005]	[0.010]	[0.041]
Perc. of female teachers	-0.027	0.103**	1.062***	-0.004	0.025	0.607***
	[0.022]	[0.049]	[0.180]	[0.024]	[0.053]	[0.211]
Perc. of teachers with	0.005	0.057	-0.312	-0.006	0.101	-0.375
at least 3 years of exp.	[0.035]	[0.078]	[0.285]	[0.031]	[0.067]	[0.268]
Perc. of teachers with	0.034	-0.013	-0.104	0.032	0.082	0.004
university degree	[0.028]	[0.063]	[0.232]	[0.025]	[0.054]	[0.214]
Teacher absenteeism	0.001	0.002	-0.011	-0.002	-0.002	0.000
	[0.002]	[0.003]	[0.013]	[0.002]	[0.004]	[0.017]
Teacher test score	0.037	0.080	0.314	0.085	-0.141	0.808
	[0.091]	[0.204]	[0.752]	[0.071]	[0.154]	[0.604]
Predicted value student test	0.314*			0.693***		
	[0.178]			[0.169]		
Predicted value st-test competitors	0.049			-0.371**		
weighted by distance	[0.174]			[0.166]		
Predicted value student test x Private	-0.087			0.170		
	[0.314]			[0.322]		
Predicted value st-test competitors	-0.045			-0.195		
weighted by distance x Private	[0.309]			[0.322]		
Predicted value mother education		0.283**			0.080	
		[0.114]			[0.108]	
Predicted value mother educ. competitors		-0.195*			0.027	
weighted by distance		[0.109]			[0.102]	
Predicted value mother education x Private		-0.126			0.094	
		[0.167]			[0.159]	
Predicted value mother educ. competitors		0.171			-0.020	
weighted by distance x Private		[0.170]			[0.161]	
Predicted value asset index test			0.305**			0.065
			[0.123]			[0.116]
Predicted value asset index competitors			-0.143			0.115
weighted by distance			[0.122]			[0.116]
Predicted value asset index test x Private			-0.109			0.328*
			[0.183]			[0.192]
Predicted value asset index competitors			0.103			-0.380*
weighted by distance x Private			[0.187]			[0.195]
Constant	0.059	-0.196	-1.838**	0.018	0.058	-1.609***
	[0.098]	[0.197]	[0.731]	[0.078]	[0.148]	[0.570]
F-test (Instruments)	6.48	4.89	7.55	11.50	7.73	8.14
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: This table reports estimates of the first stage regression of peer variables (student test score, mother education and assets on the predicted value of the peer variable in the school and on the predicted value of the peer variable in competitor schools, weighted by distance to each competitor and other school attributes. Columns (1) to (3) report the results for girls and columns (4) to (6) the results for boys. Standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

C.2 Estimates of the Model considering peer variables as endogenous

Tables C.2 and C.3 show estimates of equation (6) in a specification where peer variables are considered as exogenous (columns 1, 3, and 5) and endogenous (columns 2, 4, and 6).

Table C.4 shows the estimated coefficients for equation (6) for girls and boys (estimation of $\overline{\beta}_{kg}$ by running a regression of the school fixed effect (δ_{jtg}) on the observed school characteristics) using different specifications. The first column shows OLS estimates, the second column shows our main IV estimates. Finally, Tables C.5 and C.6 show how the effects of the school characteristics in equation (6) on utility, and the willingness to pay for each of them, change with the family background of the girls and boys, respectively.

Table C.2: Robustness check - specifications of peer variables with and without endogeneity of peer variables - girls

	25th perc.		mean		75th perc.	
	No endogeneity	Peers End.	No endogeneity	Peers End.	No endogeneity	Peers End.
School fees	-0.167***	-0.145***	-0.136***	-0.118***	-0.120***	-0.102**
	[0.041]	[0.034]	[0.041]	[0.034]	[0.041]	[0.040]
School with toilets	0.164	0.353	0.122	0.291	0.125	0.306
	[0.379]	[0.393]	[0.361]	[0.393]	[0.370]	[0.690]
School with permanent classroom	0.300	0.433	0.225	0.382	0.218	0.373
	[0.277]	[0.340]	[0.266]	[0.342]	[0.270]	[0.322]
Number of extra facilities	0.178**	0.328***	0.198***	0.340***	0.197***	0.340**
	[0.071]	[0.089]	[0.072]	[0.089]	[0.073]	[0.128]
Percentage of female teachers	0.496	0.284	0.592*	0.296	0.610*	0.348
	[0.341]	[0.517]	[0.336]	[0.514]	[0.342]	[0.554]
Perc. of teachers with at least 3 years of experience	0.137	0.356	0.251	0.482	0.304	0.521*
	[0.319]	[0.354]	[0.321]	[0.358]	[0.343]	[0.275]
Perc. of teachers with university degree	0.432	1.010**	0.530	1.096**	0.553	1.112**
	[0.430]	[0.440]	[0.449]	[0.428]	[0.432]	[0.453]
Student test score	1.447*	-4.241	1.443*	-4.912	1.394*	-4.940
	[0.759]	[3.163]	[0.777]	[3.192]	[0.783]	[4.029]
Teacher absenteeism	0.036	0.123**	0.036	0.122***	0.039	0.123*
	[0.042]	[0.048]	[0.042]	[0.047]	[0.041]	[0.066]
Teacher test score	1.620	2.512**	1.600	2.563**	1.574	2.516*
	[1.028]	[1.182]	[1.073]	[1.201]	[1.096]	[1.206]
Perc. of mother with some education	-0.976***	0.431	-0.711**	0.632	-0.620*	0.634
	[0.333]	[1.819]	[0.363]	[1.822]	[0.348]	[1.764]
Asset index	-0.064	-0.845**	-0.056	-0.700*	-0.046	-0.685
	[0.099]	[0.403]	[0.099]	[0.400]	[0.098]	[0.565]
Private	-0.357	0.448	-0.254	0.596	-0.211	0.638
	[0.483]	[0.652]	[0.494]	[0.649]	[0.465]	[1.043]

Notes: This table shows estimates of equation (6) in a specification where peer variables are taken as exogenous and endogenous. We compute the 25th and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of the specification without endogeneity. Remaining columns (2, 4, 6) report the results considering endogeneity. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table C.3: Robustness check - specifications of peer variables with and without endogeneity of peer variables - boys

	25th perc.		mean		75th perc.	
	No endogeneity	Peers End.	No endogeneity	Peers End.	No endogeneity	Peers End.
School fees	-0.051**	-0.039	-0.043*	-0.029	-0.039	-0.024
	[0.025]	[0.029]	[0.025]	[0.029]	[0.025]	[0.030]
School with toilets	0.162	0.468*	0.280	0.591**	0.329	0.635**
	[0.238]	[0.281]	[0.237]	[0.285]	[0.251]	[0.287]
School with permanent classroom	0.107	0.176	0.182	0.247	0.215	0.280
	[0.198]	[0.218]	[0.204]	[0.221]	[0.203]	[0.223]
Number of extra facilities	0.115**	0.129*	0.122**	0.134*	0.127**	0.140**
	[0.055]	[0.069]	[0.056]	[0.070]	[0.057]	[0.070]
Percentage of female teachers	-0.852***	-0.519*	-0.747***	-0.387	-0.675**	-0.311
	[0.269]	[0.305]	[0.275]	[0.309]	[0.271]	[0.312]
Perc. of teachers with at least 3 years of experience	0.091	0.527	0.186	0.628*	0.240	0.684**
	[0.286]	[0.336]	[0.283]	[0.341]	[0.281]	[0.343]
Perc. of teachers with university degree	0.171	0.387	0.112	0.308	0.091	0.285
	[0.329]	[0.365]	[0.330]	[0.370]	[0.310]	[0.372]
Student test score	0.996	0.280	1.146*	0.556	1.231*	0.610
	[0.644]	[2.039]	[0.640]	[2.068]	[0.642]	[2.082]
Teacher absenteeism	0.002	-0.031	0.003	-0.032	0.004	-0.030
	[0.024]	[0.031]	[0.023]	[0.031]	[0.024]	[0.032]
Teacher test score	1.104	-0.481	1.051	-0.611	0.992	-0.665
	[0.726]	[0.955]	[0.709]	[0.969]	[0.767]	[0.976]
Perc. of mother with some education	-0.368	-5.721***	-0.381	-5.925***	-0.391	-5.904***
	[0.298]	[1.594]	[0.307]	[1.617]	[0.303]	[1.629]
Asset index	-0.009	0.260	0.030	0.314	0.052	0.325
	[0.075]	[0.305]	[0.074]	[0.310]	[0.076]	[0.312]
Private	-0.929**	-0.152	-0.911**	-0.172	-0.909**	-0.175
	[0.405]	[0.627]	[0.396]	[0.636]	[0.387]	[0.640]

Notes: This table shows estimates of equation (6) in a specification where peer variables are taken as exogenous and endogenous. We compute the 25th and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where m = 25th percentile, mean, 75th percentile. We label these: 25th, Mean, and 75th, respectively. Columns 1, 3, and 5 show the impact of the specification without endogeneity. Remaining columns (2, 4, 6) report the results considering endogeneity. Bootstrapped standard errors in brackets.
* Significant at 10%; ** significant at 5%; *** significant at 1%.

Table C.4: OLS vs. IV regressions - Specification with peer endogeneity

	Girls		Boys	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
School fees	-0.023*	-0.118***	0.022*	-0.029
	[0.014]	[0.034]	[0.013]	[0.029]
School with toilets	0.031	0.291	0.220	0.591**
	[0.375]	[0.393]	[0.232]	[0.285]
School with permanent	0.137	0.382	0.144	0.247
classroom	[0.274]	[0.342]	[0.201]	[0.221]
Number of extra facilities	0.131*	0.340***	0.091	0.134*
	[0.070]	[0.089]	[0.056]	[0.070]
Perc. of female teachers	0.831***	0.296	-0.611**	-0.387
	[0.316]	[0.514]	[0.273]	[0.309]
Perc. of teachers with	0.399	0.482	0.269	0.628*
at least 3 years of exp.	[0.316]	[0.358]	[0.266]	[0.341]
Perc. of teachers with	0.111	1.096**	-0.148	0.308
university degree	[0.400]	[0.428]	[0.311]	[0.370]
Student test score	0.571	-4.912	0.654	0.556
	[0.708]	[3.192]	[0.625]	[2.068]
Teacher absenteeism	0.039	0.122***	0.004	-0.032
	[0.042]	[0.047]	[0.024]	[0.031]
Teacher test score	1.090	2.563**	0.761	-0.611
	[1.092]	[1.201]	[0.726]	[0.969]
Perc. of mother with	-0.539	0.632	-0.300	-5.925***
some education	[0.353]	[1.822]	[0.305]	[1.617]
Asset index	-0.133	-0.700*	-0.015	0.314
	[0.094]	[0.400]	[0.077]	[0.310]
Private	-1.270***	0.596	-1.569***	-0.172
	[0.321]	[0.649]	[0.348]	[0.636]

Notes: This table shows the estimated coefficients for equation (6) for girls (estimation of $\overline{\beta}_{kg}$ by running a regression of the school fixed effect (δ_{jtg}) on the observed school characteristics) using different specifications. The first column shows OLS estimates, the second column shows our IV estimates, which includes instruments for school fees and peer variables. For the peer variables (student test score, percentage of mother with some education and asset index) the instruments are the predicted value of the respective peer variable in competitor schools. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table C.5: Willingness to pay for school characteristics - girls

					Willingness to Pay (in U.S. dollars)		
	25th perc.	mean	75th perc.	Variable variation	25th perc.	mean	75th perc.
School fees	-0.145***	-0.118***	-0.102**				
	[0.034]	[0.034]	[0.040]				
School with toilets	0.353	0.291	0.306	1.00	2.8	2.9	3.5
	[0.393]	[0.393]	[0.690]				
School with permanent classroom	0.433	0.382	0.373	1.00	3.5	3.8	4.3
	[0.340]	[0.342]	[0.322]				
Number of extra facilities	0.328***	0.340***	0.340**	1.00	2.6	3.4	3.9
	[0.089]	[0.089]	[0.128]				
Percentage of female teachers	0.284	0.296	0.348	0.10	0.2	0.3	0.4
	[0.517]	[0.514]	[0.554]				
Perc. teachers with at least 3 years of experience	0.356	0.482	0.521*	0.10	0.3	0.5	0.6
	[0.354]	[0.358]	[0.275]				
Perc. teachers with university degree	1.010**	1.096**	1.112**	0.10	0.8	1.1	1.3
	[0.440]	[0.428]	[0.453]				
Student test score	-4.241	-4.912	-4.940	0.13	-4.4	-6.3	-7.4
	[3.163]	[3.192]	[4.029]				
Teacher absenteeism	0.123**	0.122***	0.123*	1.00	1.0	1.2	1.4
	[0.048]	[0.047]	[0.066]				
Teacher test score	2.512**	2.563**	2.516*	0.08	1.6	2.0	2.3
	[1.182]	[1.201]	[1.206]				
Perc. mother with some education	0.431	0.632	0.634	0.10	0.3	0.6	0.7
	[1.819]	[1.822]	[1.764]				
Asset index	-0.845**	-0.700*	-0.685	1.05	-7.1	-7.3	-8.2
	[0.403]	[0.400]	[0.565]				
Distance	-2.223***	-2.233***	-2.252***	0.50	-9.0	-11.1	-12.9
	[0.145]	[0.136]	[0.142]				
Private	0.448	0.596	0.638	1.00	3.6	5.9	7.3
	[0.652]	[0.649]	[1.043]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the willingness to pay for each of them, change with the family background of the girl. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile$\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 7 report the willingness to pay for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table C.6: Willingness to pay for school characteristics - boys

					Willingness to Pay (in U.S. dollars)		
	25th perc.	mean	75th perc.	Variable Variation	25th perc.	mean	75th perc.
School fees	-0.039	-0.029	-0.024				
	[0.029]	[0.029]	[0.030]				
School with toilets	0.468*	0.591**	0.635**	1.00	14.0	23.8	30.9
	[0.281]	[0.285]	[0.287]				
School with permanent classroom	0.176	0.247	0.280	1.00	5.3	10.0	13.6
	[0.218]	[0.221]	[0.223]				
Number of extra facilities	0.129*	0.134*	0.140**	1.00	3.9	5.4	6.8
	[0.069]	[0.070]	[0.070]				
Percentage of female teachers	-0.519*	-0.387	-0.311	0.10	-1.6	-1.6	-1.5
	[0.305]	[0.309]	[0.312]				
Percentage of teachers with at least 3 years of experience	0.527	0.628*	0.684**	0.10	1.6	2.5	3.3
	[0.336]	[0.341]	[0.343]				
Percentage of teachers with university degree	0.387	0.308	0.285	0.10	1.2	1.2	1.4
	[0.365]	[0.370]	[0.372]				
Student test score (average)	0.280	0.556	0.610	0.13	1.1	2.9	3.9
	[2.039]	[2.068]	[2.082]				
Teacher absenteeism	-0.031	-0.032	-0.030	1.00	-0.9	-1.3	-1.5
	[0.031]	[0.031]	[0.032]				
Teacher test score (average)	-0.481	-0.611	-0.665	0.09	-1.3	-2.2	-2.9
	[0.955]	[0.969]	[0.976]				
Perc. of Mother with some education (school level)	-5.721***	-5.925***	-5.904***	0.10	-17.1	-23.9	-28.7
	[1.594]	[1.617]	[1.629]				
Asset index (school level)	0.260	0.314	0.325	1.14	8.9	14.4	18.0
	[0.305]	[0.310]	[0.312]				
Distance	-1.131***	-1.151***	-1.165***	0.50	-16.9	-23.2	-28.4
	[0.089]	[0.079]	[0.084]				
Private	-0.152	-0.172	-0.175	1.00	-4.6	-6.9	-8.5
	[0.627]	[0.636]	[0.640]				

Notes: This table shows how the effects of the school characteristics in equation (6) on utility, and the willingness to pay for each of them, change with the family background of the boy. We compute the 25th, and 75th of maternal education and household assets (our two family background variables), as well as their mean. Then we evaluate the impacts of the school characteristics at 3 points: (m of the distribution of maternal education, m of the distribution of household assets), where $m = \{25th$ percentile, mean, $75th$ percentile$\}$. We label these: 25th, Mean, and 75th, respectively. Columns 1 to 3 show the impact of each school characteristic on utility at 3 different percentiles of the distribution of family background. Columns 5 to 7 report the willingness to pay for changes in each school characteristic, and the size of the change considered is shown in column 4. Bootstrapped standard errors in brackets. * Significant at 10%; ** significant at 5%; *** significant at 1%.

D Appendix - Voucher Experiment

In this appendix we describe additional details of the voucher experiment.

D.1 Sample

The experiment was conducted, between March and April 2017, in the 50% of the villages in the original LEAPS sample. We randomly offered vouchers of different amounts (for covering private school fees) to a random set of households. In order to participate in the experiment, a household had to include a child in school, enrolled at a grade lower or equal to grade 5, or a child not enrolled in school and between 5 and 15 years old.

D.1.1 Balance Tests

This section presents the balance tests at the village and household level. In Table D.1 we observe that the villages from the original LEAPS survey look like the others.

Table D.2 displays the household characteristics by voucher amount in the experiment (0, 50, 100, 150, 200, adn 250 PKR per month). Table D.3 shows that households and children are balanced across the different (randomly drawn) voucher amounts. Table D.4 shows the balance across voucher amounts for enrollment types and Table D.5 displays the comparison in terms of enrollment in the estimation and in the experiment sample. Tables D.6 and D.7 show the balance across voucher amounts for household and child characteristics ,respectively at individual and household level.

D.2 Description of the Experiment

D.2.1 General Principles

The game consists in offering the possibility to households with eligible children to benefit from a decrease in private primary school fees. The households can either switch their child(ren)

Table D.1: Balance Test - Village

	Non-LEAPS (1)	LEAPS (2)	(2) - (1)	p-value from joint orthogonality test
Number of Households in the village (2003)	16.043	16.197	-0.153	0.534
Average gender - Eligible children (Male)	0.509	0.523	-0.014	0.473
Average age - Eligible children	9.154	8.995	0.159	0.637
Number of eligible Households (2017)	7.370	7.258	0.112	0.846
Number of eligible (2017)	16.761	15.985	0.776	0.804
Number of opened private schools (2003)	2.783	2.773	0.010	0.984
Number of opened public schools (2003)	4.261	4.545	-0.285	0.614
Total number of opened schools (2003)	7.261	7.409	-0.148	0.835
Share of private enrollment (2003)	0.281	0.328	-0.047	0.240
PCA Index computed using 5 years (2003)	-0.218	-0.271	0.053	0.559
Number of Female Headed (2003)	1.565	1.758	-0.192	0.311
Number of Male Headed - (2003)	14.478	14.439	0.039	0.891
Level of education - Female Headed - (2003)	2.071	1.333	0.739	0.475
Level of education - Male Head - (2003)	4.106	3.873	0.233	0.473
N	46	66	112	

Notes: This table shows the balance test at the village level comparing the villages from the original LEAPS survey with the other villages in the experiment. Columns 1 and 2 present the characteristics of the Non-LEAPS and LEAPS villages, respectively. For each variable, Column 3 shows the difference between the two type of villages and column 4 displays the p-value from the joint orthogonality test.

Table D.2: Household Characteristics by voucher amount in the experiment

	Voucher Amount					
	0	50	100	150	200	250
Age of the children	8.220	8.599	8.258	7.907	8.856	8.357
Gender of the children (Male)	0.508	0.544	0.520	0.477	0.568	0.575
Parental Education	2.841	2.235	2.258	3.364	2.333	2.818
Household is female headed	0.068	0.088	0.078	0.109	0.083	0.078
Number of scholarships offered	3.235	3.146	2.977	2.884	3.227	3.006
Household Size	10.136	8.956	9.016	8.922	9.644	9.104
PCA Asset Index (Household)	-0.080	-0.474	-0.397	-0.149	-0.194	-0.265
N	132	137	128	129	132	154

Notes: This table shows the household characteristics by voucher amount (randomnly drawn) in the experiment. Columns 1 to 6 present the characteristics for the 6 possible amounts: 0, 50, 100, 150, 200, an 250 PKR per month, respectively.

Table D.3: Balance Test - Household attributes by voucher amount

	(0)vs.(50)	(0)vs.(100)	(0)vs.(150)	(0)vs.(200)	(0)vs.(250)	(50)vs.(100)
Age of the children	-0.379	-0.038	0.313	-0.636	-0.137	0.341
Gender of the children (Male)	-0.037	-0.012	0.031	-0.061	-0.068	0.024
Parental Education	0.606	0.583	-0.523	0.508	0.023	-0.023
Household is female headed	-0.019	-0.010	-0.040	-0.015	-0.010	0.009
Number of scholarships offered	0.089	0.258	0.351	0.008	0.228	0.169
Household Size (2003)	1.180	1.121	1.214	0.492	1.032	-0.059
PCA Asset Index (Household)	0.394	0.317	0.069	0.114	0.185	-0.077
N	269	260	261	264	286	265

	(50)vs.(150)	(50)vs.(200)	(50)vs.(250)	(100)vs.(150)	(100)vs.(200)	(100)vs.(250)
Age of the children	0.692	-0.258	0.241	0.351	-0.598	-0.099
Gender of the children (Male)	0.068	-0.024	-0.031	0.043	-0.048	-0.055
Parental Education	-1.129	-0.098	-0.583	-1.107	-0.076	-0.560
Household is female headed	-0.021	0.004	0.010	-0.030	-0.005	0.000
Number of scholarships offered	0.262	-0.081	0.139	0.093	-0.251	-0.030
Household Size (2003)	0.034	-0.688	-0.148	0.093	-0.628	-0.088
PCA Asset Index (Household)	-0.325	-0.280	-0.209	-0.248	-0.203	-0.132
N	266	269	291	257	260	282

	(150)vs.(200)	(150)vs.(250)	(200)vs.(250)	p-value from joint orthogonality test
Age of the children	-0.949	-0.450	0.499	0.173
Gender of the children (Male)	-0.092	-0.099	-0.007	0.567
Parental Education	1.031	0.546	-0.485	0.135
Household is female headed	0.025	0.031	0.005	0.903
Number of scholarships offered	-0.344	-0.123	0.221	0.558
Household Size (2003)	-0.721	-0.181	0.540	0.073
PCA Asset Index (Household)	0.045	0.116	0.071	0.636
N	261	283	286	

Notes: This table shows the balance tests at the household level comparing all possible pairs of the the different (randomly drawn) voucher amounts. In each column, in brackets it is presented the voucher amount used in the respective comparison.

Table D.4: Balance across voucher amounts for enrollment types

	(1)		(2)		(3)		(4)		(5)		(6)	
		0		*50*		*100*		*150*		*200*		*250*
	N/[Clusters]	*Mean/SD*	*N/[Clusters]*	*Mean/SD*	*N/[Clusters]*	*Mean/SD*	*N/[Clusters]*	*Mean/SD*	*N/[Clusters]*	*Mean/SD*	*N/[Clusters]*	*Mean/SD*
Enrolled	132	0.81	137	0.8	128	0.82	129	0.82	132	0.8	154	0.77
	[47]	[0.40]	[53]	[0.47]	[50]	[0.48]	[55]	[0.49]	[51]	[0.46]	[53]	[0.41]
Enrolled Public School	132	0.45	137	0.47	128	0.48	129	0.46	132	0.42	154	0.41
	[47]	[0.56]	[53]	[0.60]	[50]	[0.59]	[55]	[0.58]	[51]	[0.63]	[53]	[0.53]
Enrolled Private School	132	0.36	137	0.33	128	0.34	129	0.36	132	0.39	154	0.36
	[47]	[0.47]	[53]	[0.53]	[50]	[0.59]	[55]	[0.50]	[51]	[0.64]	[53]	[0.51]

t-test

Difference	*(1)-(2)*	*(1)-(3)*	*(1)-(4)*	*(1)-(5)*	*(1)-(6)*	*(2)-(3)*	*(2)-(4)*	*(2)-(5)*	*(2)-(6)*	*(3)-(4)*	*(3)-(5)*	*(3)-(6)*	*(4)-(5)*	*(4)-(6)*	*(5)-(6)*
Enrolled	0.01	-0.01	-0.01	0.01	0.04	-0.02	-0.02	0	0.04	0	0.02	0.05	0.02	0.06	0.04
	(0.88)	(0.85)	(0.75)	(0.87)	(0.29)	(0.78)	(0.70)	(1.00)	(0.38)	(0.98)	(0.77)	(0.21)	(0.74)	(0.18)	(0.48)
Enrolled Public School	-0.02	-0.02	0	0.04	0.05	0	0.02	0.06	0.07	0.02	0.06	0.07	0.04	0.05	0.01
	(0.75)	(0.73)	(0.96)	(0.54)	(0.45)	(0.98)	(0.77)	(0.42)	(0.26)	(0.74)	(0.36)	(0.17)	(0.53)	(0.40)	(0.90)
Enrolled Private School	0.03	0.01	-0.01	-0.03	0	-0.02	-0.04	-0.06	-0.03	-0.02	-0.04	-0.01	-0.02	0.01	0.03
	(0.65)	(0.85)	(0.87)	(0.64)	(0.98)	(0.82)	(0.51)	(0.40)	(0.63)	(0.71)	(0.55)	(0.81)	(0.74)	(0.88)	(0.62)

Notes: Columns (1) to (5) display the proportion of children enrolled, the proportion of children enrolled in public school, the proportion of children enrolled in private schools, in the experiment sample, for each voucher amount, ranging from 0 to 250. The experiment sample is constituted of children who lived in the 62 villages randomly drawn to be part of the experiment, and were between 5 and 15 years old, or were already enrolled in primary school at the time of the experiment. The sample is constituted of 812 children, including 42 children who were below 5 years old. Standard deviations in brackets are adjusted for clustering at the village level. The value displayed for t-tests are the differences in the means across the groups. P-values in parentheses clustered at the village level in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table D.5: Enrollment in Estimation vs. Experiment Sample

	Children 5-15			
	Child is enrolled			
	Estimation Sample	**Experiment Sample**	**Pooled Sample**	**Pooled Sample**
Sex (1=Male)	0.12***	0.028	0.10***	0.10***
	(0.017)	(0.027)	(0.015)	(0.015)
Age	0.28***	0.22***	0.26***	
	(0.018)	(0.031)	(0.017)	
Age squared	-0.015***	-0.014***	-0.014***	
	(0.00089)	(0.0017)	(0.00082)	
Age=6				0.17***
				(0.037)
Age=7				0.24***
				(0.031)
Age=8				0.25***
				(0.032)
Age=9				0.26***
				(0.033)
Age=10				0.27***
				(0.033)
Age=11				0.24***
				(0.034)
Age=12				0.16***
				(0.039)
Age=13				0.055
				(0.043)
Age=14				-0.13***
				(0.042)
Age=15				-0.14***
				(0.042)
Experiment Sample			0.028	0.025
			(0.020)	(0.020)
Constant	-0.51***	0.045	-0.34***	0.56***
	(0.093)	(0.13)	(0.084)	(0.032)
Observations	3288	845	4133	4133
Adjusted R-sq.	0.12	0.21	0.12	0.13

Notes: This table displays the relationship between age, gender and enrollment in the experiment, estimation and pulled samples respectively. The experiment sample is constituted of 845 children, aged 5 to 15 years old, who were part of households that took part in the experiment in 2017, whether the child herself was eligible or not. The estimation sample is constituted of 3288 children, aged 5 to 15 years old, who were part of the LEAPS sample in 2003 and lived in the same 62 villages than the children from the experiment. The pulled sample combines the experiment and estimation samples. Standard errors in parentheses are clustered the village level. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table D.6: Balance across voucher amounts for household and child characteristics, individual level

	Age of the child (2016)	Age of the child (2017)	Gender of the children (Male)	Enrolled (2017)	Enrolled Pub. School (2017)	Enrolled Priv. School (2017)	Parental Educ. (2003)	HH is female headed (2003)	No. of scholar. offered (2017)	Household Size (2003)	Household Size (2017)	PCA Asset Index (Household, 2003)
Voucher Amount												
50	0.38	0.38	0.040	-0.61	0.019	-0.089	-1.18*	-0.74	-0.39	-0.0077	0.020	-0.028
	(0.38)	(0.38)	(0.061)	(0.49)	(0.032)	(0.24)	(0.61)	(0.76)	(0.33)	(0.050)	(0.063)	(0.061)
100	0.038	0.033	0.016	-0.58	0.0099	-0.26	-1.12*	-1.26*	-0.32	0.0097	0.022	-0.012
	(0.38)	(0.38)	(0.055)	(0.45)	(0.033)	(0.27)	(0.59)	(0.75)	(0.37)	(0.053)	(0.062)	(0.063)
150	-0.31	-0.33	-0.035	0.52	0.040	-0.35	-1.21**	-1.37**	-0.069	0.011	0.0028	0.0083
	(0.32)	(0.31)	(0.065)	(0.47)	(0.033)	(0.25)	(0.49)	(0.66)	(0.29)	(0.034)	(0.053)	(0.052)
200	0.64	0.56	0.061	-0.51	0.015	-0.0076	-0.49	-0.83	-0.11	-0.0076	-0.038	0.030
	(0.44)	(0.44)	(0.058)	(0.46)	(0.034)	(0.23)	(0.77)	(0.78)	(0.31)	(0.047)	(0.061)	(0.064)
250	0.14	0.055	0.064	-0.023	0.0097	-0.23	-1.03**	-0.79	-0.18	-0.044	-0.045	0.0011
	(0.33)	(0.32)	(0.051)	(0.48)	(0.034)	(0.21)	(0.50)	(0.69)	(0.24)	(0.041)	(0.059)	(0.057)
Constant	8.22***	8.30***	0.51***	2.84***	0.068**	3.23***	10.1***	11.8***	-0.080	0.81***	0.45***	0.36***
	(0.27)	(0.27)	(0.044)	(0.40)	(0.028)	(0.27)	(0.72)	(0.87)	(0.36)	(0.035)	(0.048)	(0.041)
Observations	812	812	812	811	812	812	812	812	807	812	812	812
Adjusted R-squared	0.0095	0.0089	0.0049	0.010	0.0020	0.0049	0.012	0.0068	0.0043	0.0023	0.0029	0.0014

Notes: This table shows how the child and household characteristics vary by voucher amount (randomly drawn) in the experiment, ranging from 0 to 250 PKR per month. We show the OLS estimates for the child's age, the child's gender, whether the child is enrolled at all, enrolled in public school and enrolled in private school, the level of parental education, the number of scholarships offered to the child's household, the household size in 2003 and 2017, and the household PCA indexl. Standard errors clustered at the village level in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table D.7: Balance across voucher amounts for household and child characteristics, household level

	Age of the child (2016)	Age of the child (2017)	Gender of the children (Male)	Enrolled (2017)	Enrolled Pub. School (2017)	Enrolled Priv. School (2017)	Parental Educ. (2003)	HH is female headed (2003)	No. of scholar. offered (2017)	Household Size (2003)	Household Size (2017)	PCA Asset Index (Household, 2003)
Mean of Voucher Amount	0.00069	-0.00020	0.00026	0.0052	0.000039	-0.00094	-0.0011	-0.0017	0.0010	-0.0000094	-0.00044	0.00043
	(0.0020)	(0.0019)	(0.00034)	(0.0034)	(0.00024)	(0.00077)	(0.0023)	(0.0023)	(0.0014)	(0.00027)	(0.00038)	(0.00036)
Constant	8.24***	8.40***	0.51***	2.36***	0.081**	2.23***	8.81***	9.53***	-0.43*	0.82***	0.48***	0.35***
	(0.30)	(0.29)	(0.051)	(0.50)	(0.031)	(0.11)	(0.36)	(0.37)	(0.23)	(0.043)	(0.060)	(0.055)
Observations	385	385	385	384	385	385	385	385	383	385	385	385
Adjusted R-squared	0.00037	0.000033	0.0020	0.0077	0.000099	0.0022	0.00051	0.00088	0.0015	0.0000040	0.0047	0.0044

from government to private schools or keep their child(ren) enrolled in private schools. The game has to be played for each eligible child in the household.

The amount of fees reduction is randomly allocated to each child. The possible amounts are: 0, 50, 100, 150, 200 and 250 PKR per month. The amount drawn is the maximum that the household can get to enroll one child in a private primary school. If the school fees are lower, we will only cover the school fees. No cash remains with the household, or with the schools. The game has to be played even if the household is not interested in the scholarship. In case where the enumerator faces a major refusal, he has to write the reason down.

D.2.2 Experiment Steps

The enumerator says the name of the child for which the game is going to be played. The 6 cards, with the possible amounts, are shown to the respondent and are then mixed and put in the box. The respondent picks one card and the enumerator writes down on the experiment sheet the name of the child and the amount drawn. In the end, the enumerator writes down the amount on the 12 vouchers of the child (the child name will be pre-printed on the vouchers). The process starts over for the next eligible child, if any.

D.3 Regression of the private school attendance on the size of the vouchers

Table D.8 presents the marginal effect from a probit regression of the private school attendance on the size of the vouchers. Finally, Table D.9 shows the private and public school enrollment in our experimental sample.

Table D.8: Regression of the private school attendance on the size of the vouchers - Marginal effects from a Probit Specification

	All	Girls	Boys
Size of the vouchers	0.0002	0.0001	0.0003
	(0.0002)	(0.0003)	(0.0003)

Notes: In this table we display the marginal effect of the probit regression of the private school attendance on the size of the vouchers in our experimental sample for the whole sample, and separately for girls and boys. Standard errors are presented in parenthesis.

Table D.9: Private and Public School Enrollment in our Experimental Sample

Voucher	All		Girls		Boys	
Amount	Public	Private	Public	Private	Public	Private
0	0.43	0.36	0.48	0.31	0.39	0.40
50	0.48	0.31	0.45	0.31	0.51	0.32
100	0.43	0.42	0.48	0.36	0.39	0.47
150	0.45	0.39	0.45	0.39	0.46	0.39
200	0.41	0.40	0.39	0.28	0.43	0.49
250	0.44	0.34	0.49	0.29	0.39	0.39

Notes: In this table we display the percentage of children (all, girls, and boys) in our experimental sample enrolled in a public and in a private school, for different values of the voucher. The voucher amount is expressed in Pakistani Rupees. In U.S dollars the Vouchers Amounts are 0.6, 1.2, 1.8, 2.3, and 2.9 U.S. dollars (1 U.S. dollars $\approx$ 85.6 Pakistani Rupees).