Policy Research Working Paper 5710

Crossing Boundaries

Gender, Caste and Schooling in Rural Pakistan

Hanan G. Jacoby
Ghazala Mansuri

The World Bank
Development Research Group
Agriculture and Rural Development Team
&
Poverty Reduction and Economic Management Network
Poverty Reduction and Equity
June 2011

Policy Research Working Paper 5710

Abstract

Can communal heterogeneity explain persistent educational inequities in developing countries? The paper uses a novel data-set from rural Pakistan that explicitly recognizes the geographic structure of villages and the social makeup of constituent hamlets to show that demand for schooling is sensitive to the allocation of schools across ethnically fragmented communities. The analysis focuses on two types of social barriers: stigma based on caste affiliation and female seclusion that is more rigidly enforced outside a girl's own hamlet. Results indicate a substantial decrease in primary school enrollment rates for girls who have to cross hamlet boundaries to attend, irrespective of school distance, an effect not present for boys. However, low-caste children, both boys and girls, are deterred from enrolling when the most convenient school is in a hamlet dominated by high-caste households. In particular, low-caste girls, the most educationally disadvantaged group, benefit from improved school access only when the school is also caste-concordant. A policy experiment indicates that providing schools in low-caste dominant hamlets would increase overall enrollment by almost twice as much as a policy of placing a school in every unserved hamlet, and would do so at one-sixth of the cost.

This paper is a product of the Agriculture and Rural Development Team, Development Research Group; and the Poverty Reduction and Equity, Poverty Reduction and Economic Management Network. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at hjacoby@worldbank.org and gmansuri@worldbank.org.

Crossing Boundaries: Gender, Caste and Schooling in Rural Pakistan

Hanan G. Jacoby Ghazala Mansuri*

JEL codes: I21, J16, O15, Z13

Keywords: Social Fragmentation, Stigma, Demand for Education, Inequality of Opportunity

*The World Bank, 1818 H St. NW, Washington DC 20433. e-mail: hjacoby@worldbank, gmansuri@worldbank.org. The data used in this study were collected by the Pakistan Institute of Development Economics, funded by a grant (no. 84150) from the World Bank Research Support Budget, both of which we gratefully acknowledge. Thanks also to Salma Khalid for excellent research assistance, Naghma Imdad for help with the caste questions in PRHS-II, and to Jishnu Das, Chico Ferreira, and Sylvie Lambert for comments on an earlier draft.

1 Introduction

Large and persistent schooling differentials–between boys and girls and across children from different social groups–in much of the developing world pose a challenge to the canonical human capital investment model with its focus on market returns to education.[1] How much of the male-bias in educational attainment still observed in South Asia (Filmer et al., 1998) or of the disadvantage of children from low-caste families (Drèze and Kingdon, 2001) can be explained by differences in economic returns? Or rather, to turn the question around, how important are differences across groups in the *costs* of attending school? In this paper, we investigate a hitherto neglected dimension of educational choice: social barriers to schooling arising from communal heterogeneity.

An emerging empirical literature highlights the negative impact of social fragmentation on the provision of local public goods (see Alesina and La Ferrara, 2005; Banerjee et al. 2005; Khwaja, 2009). Miguel and Gugerty (2005), for example, show how ethnic diversity can lead to collective action failure, manifesting itself as an underfunding of school facilities in rural Kenya. While social heterogeneity can thus constrain educational attainment through the supply channel, we turn the spotlight for the first time on the demand channel. Since it is usually not economical to provide a school to each settlement, let alone multiple facilities catering to each social group within a settlement, access to schooling will inevitably require many to cross boundaries, be they geographical or hierarchical ones. The question we ask is whether this constitutes a significant constraint on school enrollment.

Our context is rural Pakistan, where boys are not only much more likely than girls to attend primary school, but the enrollment gap between children from "high-caste" households over those from "low-caste" households is nearly as large. Two features of Pakistani society are central to our analysis. The first, the custom of *purdah* or female seclusion, restricts women's and girl's access to public spaces, but is more rigidly enforced outside of the immediate community–specifically, the hamlet or settlement–than within it. Thus, we will argue, sending a daughter to a school located outside of her settlement (even though still within her village) is more costly than sending her to one equally far away but *within* her settlement. The second feature involves caste differentiation in which the low status group is marked as "unclean" and hence subject to discrimination and ostracism by the high

[1] For applications emphasizing human capital return differentials see, e.g., Rosenzweig and Schultz (1982), Foster and Rosenzweig (1996), Behrman et al. (1999), and Jensen (2010).

status group. To the extent that the latter group dominates the community (cf., Anderson, 2011), it may be able (and willing) to exclude the former group from accessing local public services. Thus, for low-caste households, sending a child, male or female, to a school dominated by higher castes is more costly than sending him or her to a caste-concordant school.

We examine these issues using a novel data-set that explicitly recognizes the geographic structure of villages and the social structure of their constituent settlements. Our evidence shows that, controlling for distance to school, the odds of ever enrolling are substantially lower for girls who would need to cross settlement boundaries to attend, an effect not present for boys and one that does not appear to be an artifact of school allocation decisions. Once we distinguish children by their caste, however, the crossing-boundary effect remains relevant only for high-status girls; among low-castes, neither boys nor girls benefit from having a school located within their own settlement. This phenomenon can be attributed to the fact that many low-caste children reside in high-caste dominant settlements, where their families originally settled to provide services to the landowning class. In these settlements, low caste children are stigmatized and face high psychic costs of attending school. But, when the most convenient school happens to be in a low-caste dominant settlement, we find low-caste children are much more likely to enroll.

Experiments conducted in India by Hoff and Pandey (2006) suggest that caste-based stigma plays a powerful and early role in human capital formation. Low-caste children perform poorly on cognitive tests when in the presence of their high-caste counterparts, but only if their caste status is made public; if caste is kept private, both groups do equally well. Munshi and Rosenzweig (2006), meanwhile, argue that low-caste households, who tend to have looser ties to ancestral kin networks, perceive a *greater* return to education (see also Luke and Munshi, 2007). Our findings imply that once the playing field is level–i.e., absent communal barriers and the associated stigma–low-caste children are actually no less likely to enroll in school than high-caste children. Indeed, low-caste girls, the most educationally disadvantaged group, would achieve substantially higher enrollment rates if given access to caste-concordant schools, whether these schools are placed inside or outside their own settlement. We will show that a policy of building village schools designed to serve low-caste children would increase overall enrollment by almost twice as much as a policy of placing a school in every currently unserved settlement, and would do so at one-sixth of the cost.

The next section of the paper describes the data used in our empirical analysis. Section 3 lays out the context, describing the structure of villages, the nature of female seclusion norms, and the salience of caste, following which, in sections 4, 5, and 6 we present, respectively, our empirical framework, results, and policy experiment. We conclude in section 7.

2 Data

Our analysis combines representative household survey data with information from a village and school census undertaken in the two most populous provinces of Pakistan, Punjab and Sindh. The household data come from the second round of the Pakistan Rural Household Survey (PRHS-II) conducted in 2004-05, in which 3519 households were randomly drawn from 165 villages. The school census collects detailed information on all schools inside each of these villages as well as schools lying within a 2 km walk of the perimeter of each settlement/habitation of the village.[2] GPS coordinates are available for households and schools, so the distance between each can be calculated. The hamlet in which households and schools are located is also known. The PRHS-II village census provides land ownership data and the caste/clan (*zaat/biradari)* affiliation of every household in each of the 165 villages.

The school census identifies 1326 schools of which 1112 (84%) have classes at the elementary level (up to grade 5). Of these, 63% are exclusively elementary schools, while the rest had an elementary school attached to a middle or high school. Since three-fourths of elementary schools are public (90% in Sindh province), government school allocation is central in determining access to education (Appendix Table A.1, Panel A). Among elementary schools, 51% are coeducational, 28% are exclusively for boys and 21% exclusively for girls. Public elementary schools are far more likely to be single-sex (62%) compared to private ones (13%).

We consider all children who were age 9-15 in the survey year; practically all of the children who ever enroll in school have already done so by the age of nine. Married girls, which constitute just one percent of females in this age range, are

[2] Given the difficulty in identifying all schools to which children in a village may have access, the school census was conducted in two stages. In the first stage, all schools inside the village, as well as all schools within a 2 km radius of each settlement, were visually identified and a school list was drawn up. This list was provided to all survey teams conducting the household survey. During the household survey, any non-listed schools identified by sample households were added to the initial list. Likewise, any additional schools identified during the community focus group exercise were incorporated into the final school census list.

excluded since their initial enrollment decisions would have depended on conditions in their natal homes rather than in their current residence. These criteria yield a sample of 4,699 children, of which 2,192 (47%) are girls and 2,507 (53%) are boys.

Table 1 provides some descriptive statistics for the sample, highlighting, in particular, the importance of the school entry margin in rural Pakistan. Also evident is the substantial gender gap in schooling, one that varies tremendously by region. In the more developed North Punjab, 92% of boys and 85% girls had enrolled in school by age nine, whereas the poorer southern part of Punjab and the entire province of Sindh have much lower enrollment rates (just above 50% for girls) and huge gender gaps favoring boys (17 and 24 percentage points, respectively). By contrast, dropout rates from primary school are much less differentiated by gender and region, perhaps because many girls who would have dropped out never entered school to begin with. Conditional on having enrolled in school and not completed the primary grades, 88% of 9-15 year-old boys are still in school compared to 81% of girls, with gender gaps of 5, 5, and 9 percentage points in North Punjab, South Punjab, and Sindh, respectively.[3]

3 Context

The term "village" in much of Pakistan, and indeed in large parts of South Asia, has a dual connotation. On the one hand, the administrative or revenue village is the lowest tier of the government's administrative structure, at which public services, such as schools and health workers, are allocated. Most administrative or revenue villages, however, consist of several distinct clusters of habitations or hamlets; though most residents would also refer to these as "villages", we will call them "settlements." Among the 165 PHRS-II villages, for example, our village census records 1031 named settlements. However, villages of North Punjab typically are less geographically dispersed and have much fewer settlements than those of Southern Punjab and Sindh (see Table 2 below).

[3] Beyond primary school, strong gender-region differences in dropout rates re-emerge. Conditional on having completed primary school, 81% of 9-15 year-old boys are still enrolled in school compared to 78% of girls, with region-specific gender gaps of 1, 15, and 22 percentage points, respectively. Since post-primary school enrollment, conditional on reaching that decision point, is driven more by the location of middle schools than of primary schools, we will have little more to say about it in this paper.

3.1 Seclusion practices and mobility constraints

Seclusion of women (*purdah*) is practiced widely and in varying degrees throughout rural Pakistan. As they enter puberty, girls usually experience increased enforcement of *purdah* norms and attendant restrictions on mobility and social interaction. Limits on mobility for young unmarried girls include prohibitions on traveling unaccompanied or the need to obtain permission to do so from a male family member. These mobility constraints stem from a desire to maintain family honor (*izzat*), which, in a conservative society, is intimately bound up with female reputation and hence behavior, whether real or perceived. Unchaperoned travel outside the immediate confines of the village and, especially, the settlement can invite damage to a girl's reputation.

Among households living in close proximity with long standing familial, caste, or patronage ties, as would typically be the case within a settlement, there would tend to be less harassment of each others womenfolk than of those from other settlements. Support for this claim is found in the PRHS-II, where all married women age 15-40 were asked about their perception of safety within and outside their own settlement as well as their use of any type of veil or *purdah* in public.[4] While 80% of women report feeling "safe" alone inside their own settlement, only 27% report feeling safe alone outside of it. Similarly, only 14% of women report practicing full *purdah* (coverage of whole body including face) out in public within their own settlement compared to 31% outside their settlement. Thus, settlement boundaries appear to matter as far as female reputational risk is concerned.

Other data also point to mobility restrictions as a key constraint on schooling for girls. The nationally representative Pakistan Integrated Household Survey (PIHS) of 2001-02 shows that enrollment decisions for girls in rural areas are more responsive than those of boys to proximity to school (see World Bank, 2005), but this analysis does not account for settlement boundary-crossing. More suggestive evidence comes from the 2001-02 PRHS-I, in which parents of children never enrolled were asked the main reason for the child's non-enrollment. For boys, far and away the most important reason is economic ("school too expensive" cited by 43% of parents), but the picture is quite different for girls. While economic motives still dominate (32% reported "school too expensive"), social constraints become an important consideration. Respondents were much more likely to report that they did not "approve" of

[4] Veiling or *purdah* in this context is the use of a large shawl (*chador*) to cover the head and the body, and sometimes also part of the face.

their non-enrolled daughters going to school (30%) than to disapprove of schooling for non-enrolled sons (7.5%).

Field interviews conducted in five villages randomly selected, by region, from the PRHS-II sample also highlight the importance of mobility constraints for girls school enrollment, as the following quotes illustrate:[5]

> "*...I took my daughters out of school because it was too far to walk and I feel that things are not safe there...*" (woman, Southern Punjab)

> "*When sons go to schools that are far away we don't get worried, but for our daughter we get worried.*" (woman, Southern Punjab).

> "*The problem is, when children are so small to attend primary school, we cannot expect them to walk all the way to a school in another settlement or another village. But when they grow a bit more mature, 8-9 yrs, we cannot send girls alone to walk this distance. There are tall sugarcane fields and all the way is deserted except young boys, who use the same route for going to school. This is very dangerous! The boys tease young girls – anyone could push a little child, girl or boy in the fields, one would not even know!*" (woman, Southern Punjab)

3.2 Caste and Stigma

Caste has received scant attention in Pakistan, from any disciplinary perspective. Indeed, the PRHS-II is the first large-scale survey, as far as we are aware, to collect systematic data on caste affiliation in rural Pakistan. While caste groupings in Pakistan cannot be precisely mapped into the much more extensively studied social hierarchy of neighboring India, there are significant similarities between the two due to their long shared history. In particular, caste identity is embedded within occupational differences, which are associated with status and notions of purity and pollution. Various exclusionary norms follow from these hierarchies and are exercised in relations of mutual assistance, in social networks, and in the establishment and maintenance of political power within the village economy and without. Unlike the Indian context, there has been no official acknowledgement of caste-based

[5] Khan (1998) provides similar evidence from North Punjab.

discrimination in Pakistan and thus no affirmative action programs to mitigate its impacts. Caste mobility through land transactions is also extremely limited.

We construct a caste-status identifier that categorizes dozens of distinctly named caste/clan (*zaat/biradari*) groups into "high" and "low" caste. High-caste includes all such groups that self-identify on the basis of traditional access to land (*zamindars).* The low-caste group comprises *zaats* that were historically considered either out-castes (similar to the *dalits* in India) or were in clientalist relationships with *zamindars* as providers of services in the village economy; i.e. barbers, metalworkers, clothes washers, etc. Based on this definition, around 25% of the population from which we draw our sample consists of low-caste households, with the highest proportion (35%) found in Sindh province.[6]

Village census data, summarized in Table 2, provide insight into residential patterns by caste in rural Pakistan. Villages typically contain many more distinctly named *zaat/biradari* groups than do their constituent settlements. Moreover, while more than half of all settlements are comprised of just one caste-type (mostly high), this is true of just 3% of villages. In short, settlements are far more segregated along caste/clan lines than villages.

We define a caste-type as dominant if it owns the majority of land in the settlement, based on land shares calculated from village census data. Anderson (2011), citing Dumont's (1970) emphasis on economic power rather than numerical preponderance, uses precisely this definition of caste dominance for her study of Indian villages. Of the 496 settlements from which our sample children are drawn, only 76 (15%) are low-caste dominant, reflecting both the smaller low-caste population and their lower landownership.[7] All but five of these settlements are located in Southern Punjab or Sindh. Among the low-caste children in our sample, 48% reside in high-caste dominant settlements. By contrast, only 6% of high-caste children live in low-caste dominant settlements.

Qualitative evidence gathered from field interviews documents how caste-based stigma plays out in schools either to exclude low-caste children altogether or to make them unwelcome. The following snippets are typical:

> *"No, there is no school for girls here. Our zamindar got a teacher*

[6] Comparable numbers for India put the scheduled castes (*dalits*) at 15-20% and other backward castes at 20-27% of the population.

[7] Table 2 reports that 17% of all settlements in the village census are low-caste dominant as compared to only 4% of villages. This is further indication of the relative segregation of settlements based on caste in rural Pakistan.

appointed, mainly for children of his own zaat [high caste]. There is no building for school. Children, both boys and girls sit in the verandah of [the zamindar's] autaq [meeting hall] to study." (Low-caste woman, Sindh)

"As the school for girls is closed since about two years, the [high-caste] girls are taught primary classes in the boys school. Our daughters cannot attend classes together with the [high-caste] girls. So our girls are not attending any school currently..." (Low-caste woman, Sindh)

"The children of rich [high castes] are taught seriously but our children are paid no attention to. Teachers treat children from poor households really badly. While our daughters have no access to the school at all, our boys receive no attention from the teachers. All boys sit together for classes but our boys sit on the floor." (Low-caste woman, Sindh)

"They let the daughters of [high castes] use the latrines, but tell our daughters to use the fields because you stink." (Low-caste woman, Southern Punjab)

"The teachers make the daughters of [high castes] sit inside the rooms, under the fans. Our poor children sit outside, under the sun and dust." (Low-caste woman, Southern Punjab)

3.3 Enrollment decisions

To avoid excessive notation, we verbally sketch a model of school entry. The decision is, first and foremost, a forward looking one, since completing each level or grade has an option value consisting of the net return to going on to the next level. However, there is also a cost to attending school, which includes: (1) direct school fees and other monetary outlays, (2) foregone labor in the home, (3) disutility of walking to school, and (4) psychic cost of social pressure, stigma, reputational harm, etc.

We have argued that *purdah* restrictions on girls come into force at the onset of puberty, which occurs in the later primary school or early middle school years. Yet, there are a couple of reasons to believe that mobility constraints affect the primary school enrollment decision. First, reputational risk to the family is still a concern for younger girls. Allowing them unfettered mobility may signal a loose attitude about female honor in general, with repercussions for other womenfolk. Second, because the decision to enroll a child is driven in large part by the option value of future schooling, and because schooling is costly, anything that raises the likelihood of dropping out of primary school in the future will reduce the incentive to enroll at the outset.

With these considerations in mind, we state our two main hypotheses:

Hypothesis 1: *Holding constant travel time to school*, a girl who would have to cross settlement boundaries to attend school will be less likely to enroll than a girl who can attend a school in her own settlement. The same crossing-boundary effect should not be observed for boys.

Hypothesis 2: Children will be more likely to be enrolled in school if their caste status coincides with that of the dominant caste of the settlement containing their most convenient school (child-school caste concordance).

4 Salience of Settlement Boundaries

4.1 Empirical strategy

We use household and school GPS coordinates, and information on the year of school establishment, to identify the location of each primary school that was available in or near the child's village by the time the child turned 9. As new schools are established (and old ones closed down), the set of available schools can differ between two children of different ages living in the same settlement. Moreover, since single-sex schools are common in Pakistan, the set of available schools will also vary within settlements by the gender of the child. From this information, we can construct the two key variables for our test of hypothesis 1:

$$\theta_{isv} = \begin{cases} 1 \text{ if school for child } i \text{ lies inside settlement } s \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

where settlement s of village v is that in which child i resides, and

$$d_{isv} = \begin{cases} \text{distance to inside-settlement school} & \text{if } \theta_{isv} = 1 \\ \text{distance to } \textit{nearest} \text{ outside-settlement school} & \text{if } \theta_{isv} = 0 \end{cases} . \tag{2}$$

If there are multiple schools available within the settlement, we chose the nearest one to child i for the calculation of d_{isv}.

The summary statistics in Table 1 show that around three-quarters of children in our sample had an appropriate type of school available in their settlement at the time they were nine years-old. While gender differences in θ_{isv} are negligible, the regional differences are stark: 89% of children had access to inside-settlement schools in North Punjab, a figure which drops to 51% in Southern Punjab with its much greater proliferation of settlements in a typical village. To separate the crossing-boundary effect from the effect of distance it is important that the distribution of d_{isv} conditional on $\theta_{isv} = 1$ share common support with the distribution of d_{isv} conditional on $\theta_{isv} = 0$. Figure 1 shows that there is indeed considerable overlap; 86% of inside settlement-schools lie within half a kilometer of the child's residence compared to 38% of out-of-settlement schools. At the tail of the school distance distribution, of course, outside schools dominate.

Ignoring child gender differences for the moment, consider the following regression equation for school enrollment e_{isv}:

$$e_{isv} = \alpha\theta_{ivs} + \beta(d_{isv}) + \gamma\prime X_{isv} + \eta_v + \mu_{sv} + \varepsilon_{isv} \tag{3}$$

where X_{isv} is a vector of child and household characteristics and η_v, μ_{sv}, and ε_{isv} are village-specific, settlement-specific, and child-specific error components, respectively. The parameter α captures the crossing-boundary effect and $\beta(\cdot)$ is the (possibly) non-linear mapping from school distance to the decision ever to enroll the child. In the empirical implementation, we let α, β,and γ differ by child gender. In particular, this allows us to test the null hypothesis that $\alpha_{boy} = \alpha_{girl}$; as mentioned, we anticipate that $\alpha_{girl} > 0$ and $\alpha_{girl} > \alpha_{boy}$.

Non-random placement of schools is the principal threat to the validity of OLS estimates of equation (3). Those villages, and even those settlements *within* villages, that have a particularly high demand for education (or any other unobserved characteristic positively correlated with school enrollment) may be more likely to receive a school, thus inducing a (positive) correlation between θ_{ivs} and η_v and/or μ_{sv}. To deal with the non-random placement of schools across villages, we use village fixed effects throughout our analysis to purge η_v.

In addressing the remaining issue of μ_{sv}, it will be useful to assume

$$\beta(d_{isv}) \perp \mu_{sv} | \theta_{ivs}, X_{isv}, \eta_v \tag{4}$$

which, if β is linear, is tantamount to $d_{isv} \perp \mu_{sv} | \theta_{ivs}, X_{isv}, \eta_v$. In words, given school settlement location, distance to school (or any function thereof) is uncorrelated with the settlement-specific demand for education. The intuition for this conditional exogeneity of d_{isv} is as follows: When $\theta_{isv} = 1$, an available school lies inside the child's own settlement, which, of course, may signal a high μ_{sv}, but the distance between the child's household and the within-settlement school is likely to be random.[8] When $\theta_{isv} = 0$, so that the child's settlement does not have the right type of school, d_{isv} is the distance between the child's residence and an arbitrary location in a neighboring settlement, the one containing the nearest available school. The distance between child and school settlements is–conditional on average distance to school in the village–for all intents and purposes a matter of pure chance, plausibly unrelated to relative demand for education across settlements.

With conditional exogeneity of d_{isv} we can investigate the potential correlation of θ_{ivs} and μ_{sv} without even having to partial out $\beta(d_{ihsv})$. This allows us to work with the simpler regression model

$$e_{isv} = \widetilde{\alpha}\theta_{ivs} + \widetilde{\gamma}\prime X_{isv} + \eta_v + \mu_{sv} + \varepsilon_{isv} \quad (5)$$

where $\widetilde{\alpha} = \alpha + \mathbf{P}(\beta(d_{isv}); \theta_{ivs}, X_{isv}, \eta_v)$ and $\mathbf{P}$ is the linear projection of $\beta(d_{isv})$ on the remaining regressors and village dummies. So, $\widetilde{\alpha}$ confounds the crossing-boundary effect α with the effect that a within-settlement school has on distance to school. But, given that θ_{isv} and d_{isv} are fairly highly correlated, equation (5) should provide more powerful tests of the exogeneity of θ_{ivs} than equation (3).

This brings us to the two alternative testing strategies:

1. Settlement fixed effects: Recall that because a child's set of available schools differs according to their age (i.e., cohort) and gender, θ_{ivs} varies within settlements. Moreover, even if θ_{ivs} were constant within settlements, we could still identify the gender difference $\alpha_{boy} - \alpha_{girl}$ (or $\widetilde{\alpha}_{boy} - \widetilde{\alpha}_{girl}$). Settlement fixed effects supersede village fixed effects, purging both η_v and μ_{sv}.

2. Instrumental variables with village fixed effects: We exploit the fact that, within villages, public schools are largely allocated to high population settlements. Since we know the number of households in each settlement, we

[8] To anticipate a possible concern that might arise in section 5, there is no significant relationship between the distance to an inside-settlement school and a child's caste status, whether or not we condition on the full set of other covariates ultimately used in our regressions. Thus, there is no evidence that schools are placed *within* settlements so as to be closer to high caste neighborhoods.

can construct a variable p_{sv} = proportion of households in village v residing in settlement s. Given this school allocation mechanism, we expect $Cov(\theta_{ivs}, p_{sv}|\eta_v) > 0$. The validity of excluding p_{sv} from equation (5) depends on the extent of intra-village mobility. If those households with a relatively strong demand for education tend to relocate to settlements already having schools or set up new settlements next to existing schools, then settlement population shares could reflect schooling preferences, rendering them inadmissible as instruments. While the establishment of new settlements is far-fetched in our context given long-standing land ownership patterns, the possibility of intra-village migration in response to school availability cannot be as easily dismissed.[9] However, in our analysis of village census data in the Appendix, we fail to find evidence of increased migration into settlements with newly established primary schools. This result supports our strategy of treating settlement population shares within a village as exogenous with respect to school availability.

To facilitate IV estimation and comparability across estimators we employ the linear probability model with standard errors clustered at the settlement, the lowest-level sampling unit in the household survey.[10] However, for the final results in the paper we present the corresponding probit estimates as well.

4.2 Results

To investigate the endogeneity of school location, as just discussed, the first four specifications in Table 3 exclude distance to school. In the first column, we see that, conditional only on village fixed effects and child age, the presence of a school in the settlement increases girls' enrollment rate by 18.6 percentage points, which is very close to the conditional gender gap in school enrollment. In other words, girls with a school in their settlement have practically the same enrollment rate as boys.

[9] Andrabi, et al. 2007, follow a related strategy of using information on village population in rural Pakistan to instrument the availability of schools at the village level. They assume that households do not migrate *across* villages to take advantage of schooling opportunities.

[10] Specifically, we want to directly compare estimates derived from strategies 1 and 2. However, if we take the discrete choice route, we would have to use a fixed effect logit for strategy 1 (given the large number of settlements relative to children) and a probit IV for strategy 2 (noting that the former estimator does not produce marginal effects). Moreover, as we will see later, probit (or logit) estimation is only possible on a reduced sample, which compromises efficiency relative to a linear probability model.

Next we control for household wealth using a dummy variable for whether the household is landless along with the log of per-capita household expenditures measured at the time of the survey. Neither variable is a perfect proxy for wealth, but together they should do a reasonable job of capturing such variation. Allowing the coefficients on these variables to differ by the gender of the child, we see that girls' enrollment is quite a bit more responsive to wealth than that of boys, but that the coefficients on θ_{isv} barely change.

Columns (3) and (4) of Table 3 provide evidence on the exogeneity of θ_{isv} conditional on the village fixed effect, household wealth proxies, and other regressors. Our first procedure involves simply taking out settlement fixed effects. Although the standard errors on $\widetilde{\alpha}_{boy}$ and $\widetilde{\alpha}_{girl}$ more than double as a consequence, the coefficient estimates barely differ from their counterparts in column (2); the estimate of $\widetilde{\alpha}_{girl}$, moreover, remains statistically significant.

The instrumental variable estimates in column (4) also belie school placement endogeneity.[11] In particular, we would expect upward bias in $\widetilde{\alpha}_{boy}$ and $\widetilde{\alpha}_{girl}$, as settlements with stronger education demand are more likely to attract schools. While there is a weak indication of upward bias for boys, just the opposite is observed for girls. Looking at the IV estimate of the gender difference $\widetilde{\alpha}_{girl} - \widetilde{\alpha}_{boy}$, we remain extremely confident that the enrollment impact of having an inside-settlement school (inclusive of its effect through distance to school) is greater for girls than for boys. Finally, using the settlement rank instead of the share of population as the instrument (e.g., dummies for first and second most populous settlements in the village) makes very little difference for the results.

Accepting now that our estimates of crossing-boundary effects are not likely to be an artifact of school placement decisions, we drop back to the least-squares/village dummy variable estimator and begin to control for distance to school. This is done in linear form in column (5) and with a cubic polynomial in distance in column (6). In both cases, the θ_{isv} coefficients fall only marginally relative to column (2). Figure 2 plots predicted enrollment for boys and girls as a function of distance based on the estimates in column (6). Erecting a settlement boundary between her household and a next-door school would reduce a girl's enrollment by as much as moving that school two and a half kilometers away within the same settlement.

[11] Instrument diagnostics reported in the footnote to Table 3 strongly reject underidentification and weak identification and fail to reject overidentification. Also, adding instruments involving squared population shares improves the efficiency of the 2SLS estimate (i.e., redudancy of these instruments is strongly rejected).

Given the paucity of inside-settlement schools at the greater distances, as seen in figure 1, we also experimented with alternative distance cutoffs to check whether our estimates of α_{girl} and α_{boy} are sensitive to the exclusion of far away schools. The results are, in fact, highly robust. For example, re-estimating the model in column (6) using the 3,454 children whose most convenient school is within just 500 meters, we obtain $\widehat{\alpha}_{girl} = 0.187$ (0.056) and $\widehat{\alpha}_{boy} = -0.011$ (0.037).

Lastly, we investigate the impact of middle school location on primary school enrollment decisions by including variables analogous to θ_{isv} and d_{isv} for middle schools. The likelihood of having a middle school inside one's settlement is far lower than of having a primary school (20% for boys and 29% for girls) and middle schools are considerably more distant on average as well. Indeed, since the degree of overlap between the distribution of distance for inside and outside settlement middle schools is so much lower than was the case for primary schools (cf., figure 1), separately identifying crossing-boundary and distance effects for middle school is likely to be far more difficult. Thus, we find in column (7), that having a middle school inside the settlement significantly raises enrollment rates for girls but not for boys, paralleling the results for primary school. However, including a cubic polynomial in distance to middle school in column (8) largely wipes out this middle school crossing-boundary effect (757 middle school distances are missing because the child had no such school available). The important result, returning to our main argument, is that the primary school crossing-boundary effects are not driven by correlations between primary and middle school placement; the coefficients on the primary school θ_{isv} are quite robust (compare columns (6) and (8)).

5 Salience of Caste-based Stigma

5.1 Empirical strategy

The key variable for testing our second hypothesis is an indicator of child-school caste-concordance. Let

$$CC_{isv} = \begin{cases} \text{child } i\text{'s caste} = \text{settlement } s\text{'s dominant caste} & \text{if } \theta_{isv} = 1 \\ \text{child } i\text{'s caste} = \text{nearest school settlement's dominant caste} & \text{if } \theta_{isv} = 0 \end{cases}. \quad (6)$$

If the nearest school outside the settlement is not contained within any one settlement (5% of cases), we match it to the nearest settlement and assign it the corresponding dominant caste.

Means of CC_{isv} by gender, caste, and region are shown in Table 1. Reflecting the residential patterns already discussed, high-caste children are more than twice as likely overall to have access to a caste-concordant school than low-caste children. Regionally, southern Punjab fares the worst on these terms with only 11% of low-caste children having their most convenient school in a low-caste dominant settlement.

5.2 Results

In the first column of Table 4, we introduce caste-status as a determinant of enrollment behavior in a stripped-down model with only gender-age interactions and village fixed effects. Once we look within villages, there is little difference in school enrollment between high and low-caste boys. By contrast, there is a 9.7 percentage point gap in enrollment rates favoring high over low-caste girls, albeit not a statistically significant one. Controlling for household wealth in column (2) using the landless dummy and the per-capita expenditure variable as before lowers this marginal effect of higher caste status to 7.4 percentage points. So a modest share of caste differences in enrollment can be explained by the greater wealth of high-caste households.

Next we investigate the crossing settlement boundary effect by child caste, conditioning again on distance to school. The results, reported in column (3), are remarkable. Practically the entire positive enrollment effect of having a school in one's settlement that we observed in Table 3 is concentrated among high-caste girls. Low-caste girls appear to be indifferent to the presence of an inside-settlement school. One plausible explanation for this finding is that, as noted earlier, there is a much higher proportion of low-caste children living in high-caste dominant settlements (48%) than of high-caste children living in low-caste dominant settlements (6%). Given caste-based stigma, the psychic cost of attending a local school if one exists is high for these low-caste girls.

Results in column (4) confirm our suspicions, as well as field interviews suggesting that caste-based stigma cuts across gender lines.[12] Low-caste girls with access to

[12] We had to drop 229 observations in this analysis because the most convenient school was not within the village and, therefore, information on the caste-dominance of the school settlement was

a caste-concordant school enroll at a rate 28 percentage points higher than their counterparts without such schools, whereas there is no significant difference for high-caste girls. For low-caste boys, the child-school caste concordance effect is also important but only about half as large, at 14 percentage points, as it is for low-caste girls, and there is no effect at all for high-caste boys. An otherwise identical probit model yields marginal caste-concordance effects of comparable magnitude and significance: 0.350 (0.058) for low-caste girls; 0.132 (0.053) for low-caste boys; essentially zero for high-caste children.[13]

Does the saliency of settlement boundaries reemerge for low-caste girls with access to a caste-concordant school? To answer this question (estimate the four-way interaction between girl, caste, θ_{isv}, and CC_{isv}), we would need significant numbers of low-caste girls with caste-concordant schools both inside and outside their own settlements. However, low-caste children are concentrated in relatively few low-caste dominant settlements; only 20 of the 176 low-caste girls in our sample who have a nearest school in a low-caste dominant settlement ($CC_{isv} = 1$) do not have this school in their own settlement ($\theta_{isv} = 0$). This fact renders credible estimation of the four-way interaction practically impossible.

Our estimates, at any rate, imply that if all girls were given access to a caste-concordant school within their own settlement (i.e., $\theta_{isv} = 1$ and $CC_{isv} = 1$), then the enrollment gap between high and low-caste girls would fall to -6.3 percentage points, which is to say that the enrollment rate of low-caste girls would actually be *higher* than that of high-caste girls. The same calculation shows an even bigger marginal effect for boys, with low-caste enrollment now exceeding that of high-castes by 9.1 percentage points. These results suggest that, if anything, the latent demand for education is higher among low-caste households than among their high-caste neighbors in the village; in the absence of communal barriers to schooling, they would actually be more likely to send their children to school. Low-status *per se* does not appear to dampen parents' aspirations for their children.

not available.

[13] Unlike the linear probability model, the probit estimation drops 33 villages (865 observations) in which all sample children had ever enrolled in primary school. Nevertheless, the crossing settlement boundary effects are also practically identical across estimators, with the probit marginal probability for high-caste girls at 0.184 (0.041) and insignificant for all other children.

6 Policy Implications

Our findings on the salience of geographic boundaries and the caste-composition of settlements have important implications for school allocation decisions in rural Pakistan. We illustrate this point by comparing the following hypothetical policies targeted to the population of 9-12 year-old children (the most recent 3-year cohort) in all PRHS-II villages:

Policy 1: Provide a girl's school to every *settlement* currently lacking one.

Policy 2: Provide a school to a low-caste dominant settlement in every *village* currently lacking one.

Based on our estimates (using only statistically significant coefficients), Policy 1 will induce only more high-caste girls to enroll in school, whereas Policy 2 will draw in only additional low-caste boys and girls. To level the playing field somewhat on the cost-side, we restrict the first policy to settlements with at least a 10% high-caste population and the second policy to villages with at least a 10% low-caste population, thus avoiding the construction of a large number of "superfluous" schools.[14] The question is: Which policy increases enrollment rates most cost-effectively?

Table 5 summarizes the horserace. Our analysis uses estimates from column (4) of Table 4 combined with village census data to obtain population weights for each settlement as well as the proportion of low-caste households within each settlement.[15,16] As mentioned, the enrollment impacts of Policy 1 are zero for boys as well as for low-caste girls. High-caste girls increase their enrollment, but not by much (5.3 percentage points). The reason is that, since higher population settlements typically already have a school, Policy 1 disproportionately serves smaller settlements and thus does not greatly increase school access on a population-weighted basis. Policy 2, by contrast, does not suffer the same defect; more populous settlements tend to be in villages without a low-caste dominant school. Thus, Policy 2

[14] From the village census, we know that only 0.1% of the high-caste population resides in settlements that are more than 90% low-caste and that 3.7% of the low-caste population resides in villages that are more than 90% high-caste.

[15] Throughout our simulations we ignore the fact that building new schools will also reduce distance to school. In this sense, the enrollment gains from the policies that we consider are somewhat understated.

[16] The population proportion of girls among 9-12 year old children is calculated from our sample and, along with household size, is assumed to be constant across settlements and caste-types.

greatly boosts the enrollment rates of low-caste children of both genders and, even though these children are in the distinct minority, increases aggregate enrollment by nearly twice as much as Policy 1 (3.4 versus 1.9 percentage points).

Looking at the cost side, we see that implementing Policy 1 requires building a school in 579 (out of 1,031) settlements. On the other hand, for Policy 2 a mere 93 new schools are needed. The profligacy of Policy 1 relative to Policy 2 is obviously due to its targeting the settlement rather than the village.[17] In sum, Policy 2 delivers twice the overall impact at a sixth of the cost. Moreover, Policy 2 does far better at reaching the educationally disadvantaged low-caste population, especially low-caste girls. These findings highlight the importance of recognizing social fragmentation in remediating lack of access to public services.

7 Conclusion

We have seen that social constraints can interact in surprising ways to limit educational opportunities for girls. Entry into primary school is substantially discouraged when girls have to cross settlement boundaries to attend, irrespective of the distance they would have to travel. This effect is concentrated among high-caste girls, however; low-caste girls appear indifferent to settlement boundaries. But this is because many low-caste children live in settlements dominated by high-caste households and thus face particularly high barriers to attending a local school, if one exists.

A broader lesson of this paper is that differences in the returns to education–or in household wealth, for that matter–do not fully account for school enrollment gaps across gender and caste. There is a significant role for between-group variation in schooling costs due to communal heterogeneity and the social barriers that arise therefrom. Indeed, we have seen that if one could eliminate the influence of stigma, low-caste children would enroll in primary school in greater proportions than their high-caste counterparts. Given the difficulty of designing policies that raise the returns to education on a large scale, our findings thus provide hope that suitably targeted supply-side interventions can mitigate educational inequities.

[17] Implicit here is the assumption that capacity constraints are never binding for a new school; they can accomodate any number of additional students induced to enroll.

References

[1] Anderson, S. 2011, "Caste as an impediment to trade," *American Economic Journal: Applied Economics*, 3(1):239-63.

[2] Andrabi, T., Das, J., Khwaja, A., 2007, "Students today, teachers tomorrow? Identifying constraints on the provision of education," unpublished manuscript, Harvard University.

[3] Alesina, A. and E. La Ferrara, 2005, "Ethnic diversity and economic performance," *Journal of Economic Literature*, 43(3): 762-800.

[4] Banerjee, A., L. Iyer, and R. Somanathan, 2005, "History, social divisions and public goods in rural India," *Journal of the European Economic Association*, 3(2-3), 639-47.

[5] Berhman, J., A. Foster, M. Rosenzweig and P. Vashishtha, 1999, "Women's schooling, home teaching and economic growth," *Journal of Political Economy*, 107(4):682-714.

[6] Drèze, J. and G. Kingdon, 2003, "School participation in rural India," *Review of Development Economics*, 5(1):1-24.

[7] Dumont, L., 1970, *Homos hierarchicus: The caste system and its implications.* Chicago: University of Chicago Press.

[8] Filmer, D., E. M. King and L. Pritchett, 1998, "Gender disparity in South Asia: Comparisons between and within countries," Policy Research Working Paper 1867, World Bank, Washington, D.C.

[9] Foster, A. and M. Rosenzweig, 1996. "Technical change and human capital returns and investments: Evidence from the Green Revolution," *American Economic Review*, 86(4):931-53.

[10] Hoff, K. and P. Pandey, 2006, "Discrimination, social identify, and durable inequalities," *American Economic Review (Papers & Proceedings)*, 96(2): 206-211.

[11] Jensen, R. 2010, "The (perceived) returns to education and the demand for schooling," *Quarterly Journal of Economics*, 125(2):515-48.

[12] Khan, A., 1998, "Female mobility and social barriers to accessing health and family planning services: A qualitative research study in three Punjabi villages," Islamabad: Ministry for Population Welfare, London School of Tropical Hygiene and Medicine and Department for International Development, British Government.

[13] Khwaja, A. 2009, "Can good projects succeed in bad communities?" *Journal of Public Economics*, 93(7-8):899-916.

[14] Luke, N. and K. Munshi, 2007, "Social affiliation and the demand for health services: Caste and child health in South India," *Journal of Development Economics*, 83(2): 256-79.

[15] Miguel, E. and M. K. Gugerty, 2005, "Ethnic diversity, social sanctions, and public goods in Kenya," *Journal of Public Economics*, 89(11-12): 2325-2368.

[16] Munshi, K., M. Rosenzweig, 2006, "Traditional Institutions Meet the Modern World: Caste, Gender,and Schooling Choice in a Globalizing Economy," *American Economic Review,* 96(4):1225-1252.

[17] Rosenzweig, M. and T. P. Schultz, 1982, "Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India," *American Economic Review*, 72(4):803-815.

[18] World Bank, 2005. *Bridging the gender gap: Opportunities and challenges*, Pakistan Country Gender Assessment, Washington, D.C.

Appendix
Settlement In-migration and School Availability

Migration information, while not provided in the PRHS-II village census, is available from a different village census conducted in 2006. This census includes 158 villages (420 settlements) in five districts, two in Punjab (one of which overlaps with a district in PRHS-II), two in Sindh, and one in NWFP. The census was supported by the World Bank and the Pakistan Poverty Alleviation Fund (PPAF). Selected villages comprise the sample for the evaluation of the third phase of the PPAF. The census questionnaire asks about the number of years of residence in the settlement. We also have data on all of the schools within the boundaries of each settlement, including their year of establishment.

We construct a settlement-level panel from 1980-2006 containing the number of households migrating into a settlement in each year and indicators for the presence of a boys and a girls public primary school by settlement-year. Excluding settlements in which the first public primary school was established prior to 1980 as well as those that had no public primary school at all up to 2006 leads to a sample of 115 settlements in 67 villages. The median total number of in-migrants over the whole period is only 6 households, but the median settlement population (as of 2006) is also just 42 households. There are 34 settlements that did not receive a single new household since 1980. Overall, zero migrants are recorded in 79 percent of the settlement-years.

Appendix Table A.2 reports regressions of the number of in-migrants on the presence of a boy's or girl's public primary school that include settlement fixed effects and a cubic polynomial in years since 1980 (year dummies gave similar results in all specifications). Thus, we are asking whether migration into a particular settlement (we do not have data on out-migration) increases after a school is established. OLS estimates in column (1) and (2) provide no evidence that this is the case, either for boys and girls schools separately or together for any primary school. A panel-estimator suitable for count data, the fixed effect Poisson, gives similar results in columns (3) and (4), automatically dropping all observations in the 34 settlements that received no migrants over the entire period.

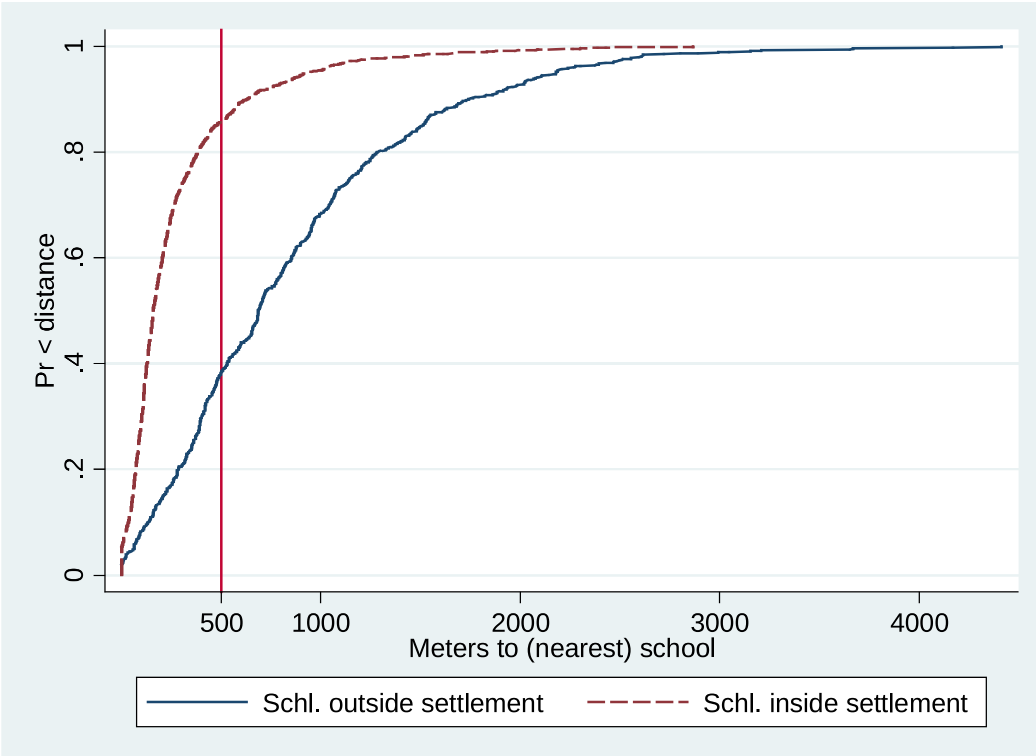

Figure 1: CDF of distance to (nearest) school for inside and outside schools.

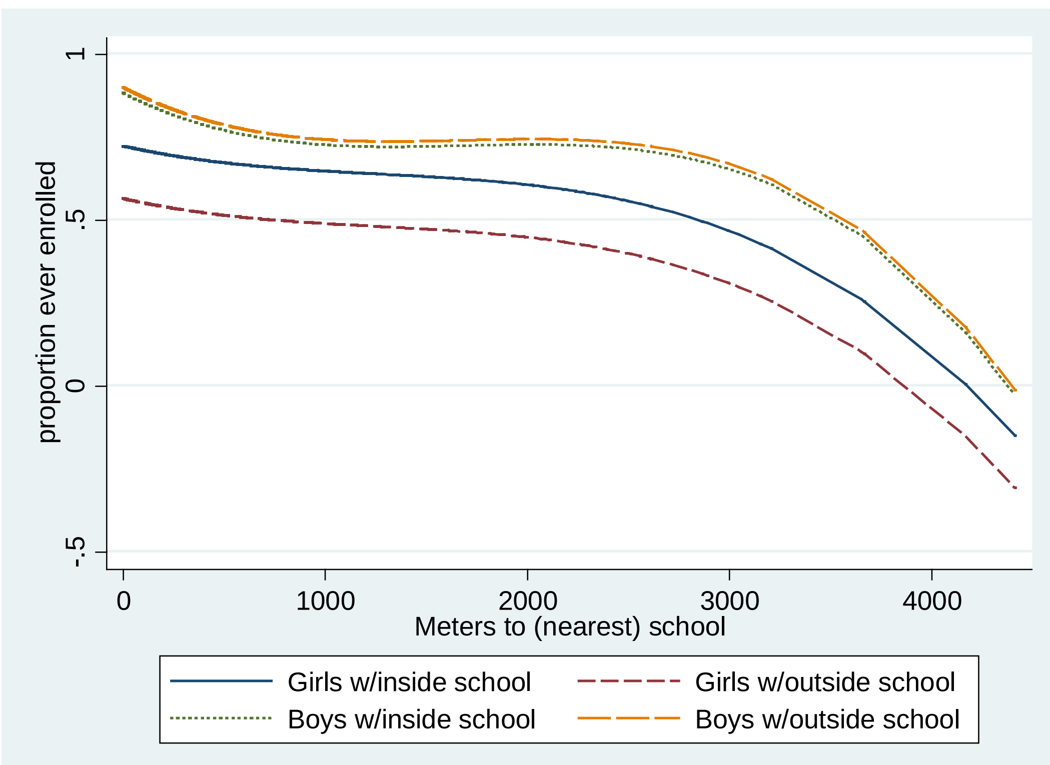

Figure 2: Predicted school enrollment by school distance and location.

Table 1
Descriptive Statistics for Sample

	Boys 9-15			Girls 9-15			All Children 9-15		
Region	Low-caste	High-caste	Total	Low-caste	High-caste	Total	Low-caste	High-caste	Total
North Punjab									
Ever enrolled	0.80	0.94	0.92	0.79	0.86	0.85	0.80	0.90	0.89
$\Theta_{primary\ school}$	0.86	0.91	0.91	0.87	0.87	0.87	0.86	0.90	0.89
CC	0.22	0.94	0.86	0.25	0.95	0.86	0.23	0.94	0.86
N	111	834	945	101	747	848	212	1,581	1,793
South Punjab									
Ever enrolled	0.6	0.70	0.68	0.34	0.55	0.51	0.48	0.63	0.60
$\Theta_{primary\ school}$	0.3	0.57	0.52	0.18	0.57	0.49	0.24	0.57	0.51
CC	0.12	0.92	0.78	0.09	0.92	0.76	0.11	0.92	0.77
N	96	385	481	90	341	431	186	726	912
Sindh									
Ever enrolled	0.70	0.79	0.76	0.44	0.55	0.52	0.58	0.68	0.65
$\Theta_{primary\ school}$	0.76	0.71	0.72	0.74	0.69	0.70	0.75	0.70	0.71
CC	0.50	0.89	0.77	0.58	0.91	0.81	0.53	0.90	0.79
N	335	746	1,081	261	652	913	596	1,398	1,994
All Pakistan									
Ever enrolled	0.70	0.84	0.81	0.50	0.68	0.65	0.61	0.77	0.73
$\Theta_{primary\ school}$	0.70	0.77	0.75	0.66	0.75	0.73	0.68	0.76	0.74
CC	0.38	0.92	0.8	0.41	0.93	0.82	0.40	0.92	0.81
N	542	1,965	2,507	452	1,740	2,192	994	3,705	4,699

Notes: All statistics are means of indicator variables. All Pakistan refers to Punjab and Sindh provinces. $\Theta_{primary\ school}$ and CC are dummies for, respectively, whether a school is available within the child's settlement and whether the most convenient school is in a settlement where the child's own caste is dominant (see text for details).

Table 2
Descriptive Statistics for Village Census

	All		Region					
			Sindh		South Punjab		North Punjab	
	Settlement	Village	Settlement	Village	Settlement	Village	Settlement	Village
Settlements (median no.)	---	10	---	12	---	11	---	4
Zaat/biradari groups (median no.)	2	9	2	10	3	8	4	8
Low-caste households (%)	30	29	36	34	23	23	20	22
Single Caste-type (%)	52	3	60	2	50	0	28	6
Of which, high-caste (%)	75	100	68	100	87	---	98	100
High-caste dominant (%)	83	96	79	97	82	97	96	94
N	1031	165	581	61	268	32	182	72

Note: A *zaat/biradari* group represents a distinctly named caste/clan affiliation, whereas a caste-type is a categorization of the latter into high and low status as described in the text. Caste dominance is based on which type owns the majority of the land in the settlement/village.

Table 3
Crossing-Boundary Effect on School Enrollment: Children Age 9-15

	(1)	(2)	(3)	(4)[a]	(5)	(6)	(7)	(8)
Girl × $\theta_{primary\ school}$	0.186***	0.181***	0.186**	0.216***	0.150***	0.151***	0.118***	0.138***
	(0.035)	(0.034)	(0.076)	(0.063)	(0.038)	(0.039)	(0.040)	(0.044)
Boy × $\theta_{primary\ school}$	0.039	0.035	0.003	-0.086	-0.009	-0.019	-0.009	-0.020
	(0.027)	(0.027)	(0.077)	(0.054)	(0.029)	(0.030)	(0.030)	(0.031)
Girl × landless		-0.089***	-0.073***	-0.090***	-0.085***	-0.086***	-0.085***	-0.103***
		(0.022)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
Boy × landless		-0.066***	-0.062***	-0.069***	-0.063***	-0.065***	-0.067***	-0.077***
		(0.018)	(0.021)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
Girl × ln(pce)		0.104***	0.108***	0.103***	0.106***	0.106***	0.102***	0.090***
		(0.018)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.019)
Boy × ln(pce)		0.048***	0.055***	0.048***	0.048***	0.048***	0.053***	0.043**
		(0.018)	(0.021)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Girl × $\theta_{Middle\ school}$							0.120***	0.051
							(0.036)	(0.042)
Boy × $\theta_{Middle\ school}$							0.002	0.018
							(0.039)	(0.044)
Gender × $d_{primary\ school}$	No	No	No	No	Linear	Cubic	Cubic	Cubic
Gender × $d_{middle\ school}$	No	No	No	No	No	No	No	Cubic
Fixed effects	village	village	settlement	village	village	village	village	Village
N	4,699	4,695	4,695	4,695	4,695	4,695	4,695	3,938
Adj. R^2	0.248	0.265	0.327	---	0.270	0.272	0.275	0.268

Notes: Standard errors in parentheses clustered on settlement (*** p <0.01, ** p <0.05, * p <0.1). There are 496 settlements in 165 villages. Dependent variable is an indicator for whether the child was ever enrolled in school. All regressions include a girl dummy and the age of child interacted with girl and boy dummy variables. Θ_j is a dummy for whether a school of type *j=primary,middle* is available within the child's settlement; d_j is distance to (nearest) school of type *j*.

[a]2SLS estimate using village population share of settlement and its square, both interacted with the Girl and Boy dummies, as instruments. Instrument diagnostics are as follows: Kleibergen-Paap underidentification test Wald stat = 59.0 (3); Kleibergen-Paap weak identification Wald rk F-stat = 33.8; Hansen overidentification J-test stat = 4.2 (2); LM test of redundancy of instruments involving squared population shares = 41.2 (4), where numbers in parentheses are degrees of freedom for chi-square statistics.

Table 4
Caste Effects on School Enrollment: Children Age 9-15

	(1)	(2)	(3)	(4)
Low-caste boys	---	---	---	---
High-caste boys	0.032	0.023	0.022	0.050
	(0.030)	(0.029)	(0.049)	(0.068)
Low-caste girls	-0.103	-0.588**	-0.711***	-0.899***
	(0.080)	(0.228)	(0.237)	(0.243)
High caste girls	-0.006	-0.514**	-0.722***	-0.911***
	(0.081)	(0.229)	(0.237)	(0.245)
Low-caste boys × $\theta_{primary\ sch.}$			-0.011	-0.057
			(0.054)	(0.061)
High-caste boys × $\theta_{primary\ sch.}$			-0.018	0.005
			(0.032)	(0.034)
Low-caste girls × $\theta_{primary\ sch.}$			0.069	0.031
			(0.074)	(0.078)
High caste girls × $\theta_{primary\ sch.}$			0.176***	0.200***
			(0.042)	(0.041)
Low-caste boys × CC				0.138**
				(0.054)
High-caste boys × CC				-0.064
				(0.051)
Low-caste girls × CC				0.277***
				(0.063)
High caste girls × CC				0.057
				(0.077)
Gender × wealth	No	Yes	Yes	Yes
Gender × $d_{primary\ sch}$	No	No	cubic	Cubic
Fixed effects	village	village	village	Village
N	4,699	4,695	4,695	4,466
Adj. R^2	0.240	0.256	0.273	0.287

Notes: Standard errors in parentheses clustered on settlement (*** p <0.01, ** p <0.05, * p <0.1). There are 496 settlements in 165 villages. Dependent variable is an indicator for whether the child was ever enrolled in school. All regressions include the age of child interacted with girl and boy dummy variables. CC is a dummy for whether the most convenient school is in a settlement where the child's own caste is dominant (see notes to Table 3 for definitions of other variables).

Table 5
School Placement Policy Simulations

	Low caste	High caste	All
Boy			
Policy 1	0.000	0.000	0.000
Policy 2	0.092	0.000	0.023
Initial enrollment	0.714	0.855	0.820
Girl			
Policy 1	0.000	0.053	0.040
Policy 2	0.187	0.000	0.047
Initial enrollment	0.507	0.700	0.649
All Children			
Policy 1	0.000	0.025	0.019
Policy 2	0.137	0.000	0.034
Initial enrollment	0.616	0.780	0.739

Notes: All figures, except for the initial enrollment rates of 9-12 year-olds, represent changes in the proportion of children in PRHS-II villages who ever enroll in school. Initial enrollment rates are based on means from the estimation sample by caste-type, which are then weighted using village census data to achieve representativeness across caste-type. See text for a description of the policies.

Table A.1
Descriptive Statistics for Schools

		School Type[a]		Caste Dominance of School Settlement		
	Total (N)	Public (%)	Private (%)	High (%)	Low (%)	No info (%)[b]
Panel A: Elementary-level Schools in the Census						
N Punjab						
All Boys	179	93	6	59	3	38
All Girls	114	94	6	61	4	34
Co-Ed	228	29	68	66	1	33
Total	521	65	33	63	2	35
S Punjab						
All Boys	82	93	4	41	4	55
All Girls	65	95	2	51	5	45
Co-Ed	88	23	69	45	5	50
Total	235	67	28	46	4	50
Sindh						
All Boys	48	100	0	56	15	29
All Girls	48	92	2	63	8	29
Co-Ed	260	89	9	48	13	39
Total	356	91	7	51	12	37
All Schools	1,112	74	24	55	6	39
Panel B: Elementary-level Schools in the Sample						
N Punjab						
All Boys	74	99	1	95	5	0
All Girls	66	97	2	88	8	4
Co-Ed	135	31	65	83	1	16
Total	275	65	33	87	4	9
S Punjab						
All Boys	32	94	3	71	9	20
All Girls	36	97	0	75	6	19
Co-Ed	44	36	52	70	7	23
Total	114	73	21	72	7	21
Sindh						
All Boys	29	100	0	73	17	10
All Girls	29	93	3	73	10	17
Co-Ed	158	95	1	68	16	16
Total	216	95	1	69	16	15
All Schools	605	77	19	78	9	13

Notes: Census data in Panel A reflect an exhaustive list of schools to which children in the 165 sample villages had access to and hence include many peripheral schools. Schools in Panel B are those nearest to the sample children or within the same settlement.

[a] Other types of schools include community, NGO, and religious.

[b] School located outside of village boundaries, hence no caste information from village census.

Table A.2
School Establishment Effect on Settlement In-migration: 1980-2006

	OLS		Poisson	
	(1)	(2)	(3)	(4)
Girl's public primary school in settlement	0.127 (0.199)		0.167 (0.213)	
Boy's public primary school in settlement	0.056 (0.212)		0.135 (0.225)	
Any public primary school in settlement		0.153 (0.137)		0.233 (0.209)
Year	cubic	cubic	cubic	cubic
Fixed effects	settlement	settlement	settlement	settlement
N	3105	3105	2187	2187
Settlements	115	115	81	81

Notes: Standard errors in parentheses clustered on settlement. Dependent variable is the number of households migrating into settlement in year.