Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital

Javaeria A. Qureshi*
University of Chicago

November 30, 2011

A vast literature on the spillovers from girls' education focuses on the impact of maternal education on child outcomes. This paper is the first to investigate whether externalities from investing in girls' education may be realized earlier – before they have children of their own. In many developing countries, oldest sisters share significant child care responsibilities in the household and potentially play an important role in younger siblings' learning. I propose a model incorporating the effect of the oldest sister that predicts competing effects of increasing oldest sister's schooling on younger siblings' human capital. Using an identification strategy that exploits the gender segregation of schools in Pakistan, I estimate the impact of the oldest sister's schooling on the human capital acquisition of her younger brothers. I find that oldest sister's schooling has significant, beneficial impacts for younger brothers' schooling, enrollment, literacy and numeracy. An additional year of schooling for the oldest sister increases the younger brother's completed years of schooling by 0.42 years and his probability of being enrolled by 9.6 percent. It also increases the probability of a primary school-aged younger brother being literate and numerate by 7-19 percent. I discuss the implications of these results for policies targeting girls' education. These findings indicate that evaluations of such policies that consider only effects on the girls and their children may underestimate their total benefits.

*Contact: javaeria@uchicago.edu. I am very grateful to Dan Black, Robert LaLonde, and Ioana Marinescu for their guidance, support and feedback. I thank Guadalupe Bedoya, Pedro Bernal, Kerwin Charles, Jane Herr, Janna Johnson, Ofer Malamud, Nikolas Mittag, Kyung Park, and seminar participants at the Workshop for Human Potential and the Harris School PhD workshop at the University of Chicago for helpful comments. I thank Jishnu Das for providing access to and guidance on the LEAPS data. I acknowledge the Irving Harris Fellowship, and the Center for Human Potential Pre-dissertation and Dissertation awards for financial support.

1. Introduction

A vast literature establishes the beneficial impact of maternal schooling on a range of child outcomes: birth outcomes (Currie and Moretti, 2003), nutrition (Thomas et al, 1991), education (Haveman and Wolfe, 1995), and test scores (Rosezweig and Wolpin, 1994). This paper investigates whether educating girls may yield significant externalities even before they have children of their own. I extend the literature on social returns to educating girls in a new direction by investigating the spillovers from girls' education on the human capital acquisition of younger siblings.

I use data from rural Pakistan to identify the spillover effects of the oldest sister's schooling on younger brothers' years of schooling completed, enrollment, literacy, and numeracy. I exploit the gender segregation of schools in Pakistan to generate quasi-experimental variation in oldest sister's schooling in order to identify causal impacts on younger brothers. While the impact of oldest sister's schooling on both younger brothers and sisters is of theoretical interest, my preferred identification strategy only allows me to estimate causal impacts for younger brothers.

There are several reasons why the oldest sister's schooling could be important for younger siblings, especially in a developing country context. The oldest sister shares significant child care responsibilities in many developing countries (Levinson and Moe 1998, Ilahi 2001, Edmonds 2007). .As someone who looks after and interacts with younger siblings extensively, the oldest sister has potential for significant influence on younger siblings' learning. In Pakistan, older sisters are also the most important source of help with studies for young children. Only 1 in 5 children get help with studies from a parent while over half of the children report the family member helping them with studies on a regular basis is an older sister.

The oldest sister's schooling may play a more important role in child learning if the oldest sister is relatively highly educated as compared with other child care providers, especially mothers. We might expect diminishing marginal returns to the education of family members such that children gain substantially when the first family member acquires education but relatively less from increases in education in a family that is already well-educated. The extremely low parental – and particularly maternal – education in rural Pakistan makes it a particularly rich context in which to investigate the impact of oldest sister's schooling.

I propose a model that incorporates the effect of oldest sister's schooling on younger brother's human capital acquisition. I build on Becker's classic model of human capital by treating time spent with the oldest sister as a direct input into the younger brother's human capital production. The model yields competing effects of increasing oldest sister's schooling on younger brother human capital: a positive quality effect and a negative quantity effect. The positive quality effect captures the fact that increasing the oldest sister's schooling improves the quality of the time she spends with her younger brother. A key way in which oldest sister's schooling differs from parental schooling is that parental schooling is completed before children are born whereas oldest sister's schooling is likely to still be ongoing when the child is young. This creates a tradeoff for oldest sister's schooling which does not exist for parental schooling because increasing the schooling of the oldest sister requires allocating more time towards schooling, and therefore leaves less time with the younger brother. I term this negative spillover effect the quantity effect. Because increases in oldest sister's schooling are associated with a positive quality effect and a negative quantity effect, the net impact on younger brother human capital is theoretically ambiguous and needs to be determined empirically.

Evaluating the impact of oldest sister's schooling on younger siblings is rife with selection issues because households which educate their oldest daughters more are likely to invest highly in the education of all their children, including their younger sons. OLS specifications with a rich set of controls including household characteristics, parents' education and wealth/asset controls indicate that oldest sister's schooling is associated with improved schooling and learning for younger brothers. I use a falsification test that shows these OLS estimates are biased upwards and highlights the need for quasi-experimental variation in oldest sister's schooling.

My preferred estimation strategy relies on the gender segregation of schools in Pakistan. I use distance to closest girls' school, conditioning on distance to village center, and distance to boys' schools to create exogenous variation in the schooling of the oldest sister[1]. Distance to girls' school imposes a harsh penalty on girls' schooling, similar to that documented by Alderman et. al (1996), Andrabi et. al (2008) in

[1] Data contains GPS co-ordinates for all surveyed households and all schools in the village enabling me to calculate distances between each household and school. I use distance to closest government girls' school while controlling for distance to village center, distance to closest government boys' school, and distance to closest private school.

Pakistan, and Burde and Linden (2010) in Afghanistan. This penalty arises because of safety and chastity concerns about girls outside the home. The fact that somebody needs to accompany the girl on her walk to and back from school creates a large burden on girls' schooling that does not exist for boys.

I provide several arguments for and tests of the validity of the instrument of distance to girls' school in this context. Since boys do not attend girls' schools, distance to girls' school can only affect boys indirectly through how it varies with other factors. First, I demonstrate that the usual concern with the use of distance to school instruments in the U.S. literature does not apply in the context of rural Pakistan. I show that most households in the data are agricultural households tied to their land, and have very restricted mobility and capacity to locate systematically with respect to schools. Second, I control for distance to village center to soak up any variation in distance to girls' schools driven by how centrally or remotely a household is located. Third, I control for distance to boys' schools the younger brothers will attend so that identification comes from comparing households equidistant from the village center and equidistant from the boys' school, and different only in their distance to girls' school. Fourth, I show that distance to boys' schools does not have any impact on oldest sister's schooling. Last, I implement a falsification test in which I analyze the impact of oldest sister's schooling on her *older* brothers' outcomes. The proposed model suggests spillover effects of oldest sister's schooling on her younger siblings due to her role and activities as an *older* sibling. Theoretically, the impact should only flow from older sister to younger siblings, and there should be no detectable impact of oldest sister's schooling on her older brothers. I show that while I find positive, significant impacts of oldest sister's schooling on her older brothers using OLS, I do not find any impact on older brothers when I use the instrumental variables (IV) strategy. This evidence collectively provides strong justification for the validity of the conditional instrument of distance to girls' school in this context.

Using the IV, I find that the oldest sister's schooling has significant, beneficial impacts for younger brothers' human capital. In the context of my model, the quality effect of the oldest sister's increased schooling outweighs the quantity effect to produce a net positive impact. I find that an additional year of schooling completed by the oldest sister increases the younger brothers' completed years of schooling by 0.42 years and increases his probability of being enrolled in school by 9.6 percent. It also increases the probability of a primary school-aged younger brother being able to read, add, and count by 19, 12 and 7 percent, respectively. The impact on writing is positive but not statistically significant. Increasing oldest sister's schooling by one year increases her completed schooling by a third of a standard deviation. Relative to the mean, an additional year increases oldest sister's schooling by 25 percent.

This is the first paper I am aware of to estimate the impact of oldest sister's schooling on any younger sibling outcome. Past studies on siblings have focused mostly on the effects of the number and sex composition of siblings on education (Butcher and Case, 1994, Black, Devereaux and Salvanes, 2005). Shrestha (2010) highlights the potential for inter-sibling rivalry in education due to competition for resources by showing that an increase in male education decreases the education of female siblings in the household. My study is the first to conceptualize oldest sister's schooling as an input into younger siblings' learning, and relates closely to the literature on impact of maternal schooling. The effects I find for the impact of oldest sister's schooling are large, exceeding the median impact of maternal schooling on children found in the literature. These effects are also large relative to the impact of maternal schooling I find in my data.

The IV effects I find exceed those found using OLS but the IV effects can only be interpreted as local average treatment effects. I provide evidence that suggests that the reason IV results are larger than OLS is due to the fact that treatment effects for complier households moved by the instrument are much larger than effects for other sub-populations. I show that oldest sister's schooling is most strongly impacted by the instrument in households with uneducated mothers, and treatment impacts appear to be larger in households with uneducated mothers compared to households with educated mothers. The large treatment impacts I find relate closely to the low education of household members. Seventy-five percent of mothers and 40 percent of fathers have no schooling. For many of these households, the oldest sister is one of the first family members to acquire any schooling, and therefore her schooling is expected to generate large spillovers in the family.

The paper is organized as follows. Section 2 provides institutional background and motivation. I propose a simple model for how oldest sister's schooling may impact younger brother human capital in Section 3. Section 4 describes the empirical set-up. Section 5 lays out the empirical identification strategy and results. Section 6 includes robustness checks while Section 7 discusses policy implications and concludes.

2. Institutional Background

This paper estimates the impact of oldest sister's schooling on younger brother human capital using data from rural Pakistan. I use the Learning and Educational Achievement in Punjab Schools (LEAPS) data which is a longitudinal survey of 1800 households in the province of Punjab, which is home to 56% of the country's population[2]. Data was collected for 112 villages from three districts of Attock, Faisalabad, and Rahim Yar Khan[3] from 2003-2006. A random sample of households was selected from each village, and all the schools in these villages were surveyed[4]. The data contains household surveys, school surveys and detailed GPS co-ordinates on all surveyed households and all schools in the 112 villages sampled.

Pakistan is a country with very low educational attainment. The net primary enrolment rate[5] was 56 percent in 2008. Only 67 percent of Pakistani males and 42 percent of females aged 10 and above were literate in 2006-2007[6]. Low educational attainment in Pakistan is further characterized by significant gender disparities. Figure 1 shows the age-specific enrollment rates for girls and boys in my data. Primary school in Pakistan comprises of grades 1-5, middle school grades 6-8 and high school grades 9 and 10. The average age of a child enrolled in grade 1 is 7.3 years while that of a child in grade 5 is 11.7 years. There is an 8 percentage point gender gap in enrollment for children of primary school-age (6-12 years old). For children older than primary school-age (13-18 years old), this gap widens to a 14 percentage point difference. Figure 2 shows the gender gap in years of schooling completed by girls and boys aged 16-20[7]. There is a sizable fraction of children with no completed schooling. While 28.6% of girls have completed zero years of schooling, the

[2] Government of Punjab, 2007. Online at http://pportal.punjab.gov.pk/punjabcms//docimages/12699punjab_2007.pdf.

[3] These districts represent an accepted stratification of Punjab province into North, Middle and South regions

[4] Sample selection was carried out in two stages. In the first stage, 112 villages were selected at random from a list of villages with an existing private school and where the total number of schools did not exceed 20. In the second stage, the selection of households from selected villages was stratified to ensure inclusion of households with at least one child enrolled in grade three as well as households with children 8-10 years old none of whom were enrolled in school.

[5] Net primary enrolment rate is the number of children of official primary school age enrolled in primary education as a percentage of the total children of the official school age population.

[6] Federal Bureau of Statistics, 2008, "Pakistan Social and Living Standards Measurement Survey (PSLM) 2006-07", Islamabad, Pakistan

[7] I show years of schooling completed for boys and girls aged 16-20 because the average 16 year old is old enough to have completed middle school if he/she were to stay enrolled in school till that age.

similar figure for boys is less than half that at 12.6%. The average boy has completed 5.95 years of schooling compared with 4.53 years for the average girl. Since schooling is completed for the overwhelming majority of girls in this age range while it is still in progress for some boys, these figures likely represent lower bounds on the gender gap in educational attainment[8].

These large gender gaps in education fit into a broader context of strong son preference in this society. Many authors – most notably Amartya Sen - have documented the phenomenon of missing girls in many Asian societies including Pakistan. This is also reflected in household composition within my data. The average family has 7.6 members, of whom 4.64 are children, of whom 2.40 are boys and 2.24 are girls. The three adults are typically the two parents and a grandparent. Post-marital residence in Pakistani society is virilocal i.e. the girl moves out of her parents' household and into her husband's household upon marriage. Sons, on the other hand, continue to live with their parents even after marriage and support their parents in old age. Labor force participation rates for women are very low: only 11.85% of the mothers in my data report working. If I exclude agriculture or herding, only 7.15% of mothers report working. Given that girls do not contribute to the earned income of the household through market work, and move away after marriage while sons provide support for parents in their old age, it might likely be efficient for parents to invest less in girls' schooling relative to boys' schooling (Strauss and Thomas, 1995).

Parents have very low educational attainment in my data. Only a quarter of the mothers have ever had any formal schooling compared to 62% of fathers who have had some schooling. The average schooling completed is 1.47 years for mothers, and 4.36 years for fathers (5.25 years versus 6.86 years respectively when excluding those with no schooling). The extremely low parental education in these villages will be an important feature for understanding the size of the impact of oldest sister's schooling.

There are a number of potential mechanisms for how oldest sister's education may be important for younger sibling learning. I give a quick overview of these factors here before describing them more fully, and highlighting why it is oldest sister's education that is instrumental rather than oldest brother's. Older sisters are an important source of child care for younger siblings. In Pakistan, older sisters are the most important

[8] I chose 16-20 as the age range to show the distribution of years of schooling completed to strike a balance between completed schooling while ensuring that I can still capture girls in the household before they move out upon marriage.

source of help with studies and homework. There could also be pure interaction effects from talking and playing with a more educated family member, as well as role model effects wherein older sister's attainment motivates younger siblings. I analyze the relationship between the *oldest* sister's education and outcomes of younger brothers because it is the oldest sister who bears the most responsibility of taking care of younger siblings. Edmond highlights older girls' comparative advantage in home production, especially in caring for younger children, as an explanation for differences in hours worked across siblings. He shows that in Nepal, the oldest girl in a household with four children works 4.2 more hours per week than the second oldest girl.

Oldest sisters share significant child care responsibilities in many developing countries. There is a strict gender division in work performed by girls and boys in Pakistan, with girls responsible for housework and boys engaged in market work. Housework includes cooking, cleaning, and taking care of younger siblings as well as farm and livestock-related chores. While 83.1% of girls aged 15-18 report doing housework on a given day, only 18.5% of boys report doing so. The typical 15-18 year old girl spends an average of 6 hours daily on housework compared with 1.5 hours for similarly-aged boys. If I exclude farm-related activities from housework, the gender gap in housework widens with the average girl working 5.5 hours daily compared to just half an hour for the average boy[9]. Given the strong gender division in work responsibilities, it is the oldest sister and not the oldest brother who are responsible for taking care of younger siblings. As one of the child care providers, the oldest sister's education is likely an important input into younger siblings' learning.

In Pakistan, oldest sisters are the most important source of help for children with their studies. In contrast to developed countries where parents help children with studies, only 1 out of 5 children receiving help from a family member reported getting it from a parent in Pakistan. This is not so surprising given the low level of parental education. When parents are not the ones helping with studies, an older sister is fulfilling that role 70% of the time. She is twice as likely as an older brother to help out, and conditional on helping, older sisters spent an average of 7 hours in the past week helping younger siblings with studies.

[9] Boys are more likely to be engaged in paid work. The average 15-18 year old boy spends 2.5 hours a day on paid work compared with just 14 minutes for a girl of that age.

Even putting aside the role of oldest sister as child care provider and tutor, the oldest sister is more likely to influence younger brother learning because oldest sisters spend significantly more time at home relative to oldest brothers[10]. Since younger children spend the bulk of their time at home, especially in the developmentally significant, pre-school ages, oldest sisters have much greater interaction with younger siblings than oldest brothers. Young children are likely to benefit from their oldest sister's superior verbal abilities. Girls talk earlier, articulate more clearly, use longer sentences, and are more fluent (Maccoby, 1966). Studies have documented that the amount of language exposure is crucial for vocabulary growth in children (Huttenlocher et. al, 1991). Highly educated and high SES mothers talk more, use more complex syntax and a more varied vocabulary when talking with their children, and their children have larger vocabularies (Brooks-Gunn, and Markman, 2005; Hoff, 2003) It is thus plausible that younger siblings can benefit from language exposure from oldest sisters, especially if oldest sisters are highly educated relative to the mothers.

3. Model

This section provides a simple model of how the oldest sister affects the human capital acquisition of a younger brother. I start with the standard model of human capital investment (Becker 1964) where parents make the decision of investment allocations for each of their children by weighing the benefit of the investment against its costs. Just as Becker modeled parents' time investment in children as an input into their human capital production, I model time spent with the oldest sister as an input into the younger brother's human capital. As discussed above, the rationale behind this extension is that oldest sisters are important sources of child care for younger siblings, and they provide significant help with studies.

An increase in the schooling of the oldest sister has two competing effects on her younger brother's human capital. Having a more educated oldest sister increases the quality of the time spent with her. More educated oldest sisters are likely to use a greater number and variety of words in their conversation with younger siblings. A more educated oldest sister is also more likely to provide younger siblings help with

[10] Note the difference in time spent in housework as a rough indicator for time spent at home

studies and that help is likely to be of higher quality. I call this phenomenon whereby increased schooling for the oldest sister improves the quality of each unit of time spent with the younger brother the "quality effect".

A key way in which oldest sister's schooling differs from parental schooling is that parental schooling is completed before children are born whereas oldest sister's schooling is likely to be ongoing when the child is young. This creates a tradeoff for oldest sister's schooling which does not exist for parental schooling because increasing the schooling of the oldest sister requires allocating more time towards schooling, and hence less time with the younger brother. The oldest sister's time spent in school, and traveling to and from school constitutes time that is not spent with the younger brother. This tradeoff is most acute when the younger brother is of pre-school age since he is not yet enrolled in school and spends most of his time at home. I denote this negative impact of increasing oldest sister's schooling on time spent with younger brother, and hence his human capital, the "quantity effect".

Mechanics

I propose a model of human capital acquisition which incorporates the linkages described above. Consider the optimization problem for a household with parents and two children, an older girl and a younger boy. Parents maximize a separable utility function which is concave in period 1 consumption and the utility of their two children which is realized in period 2. In period 1, parents have access to wealth W_p and make the decisions of how much to consume, C and how much of each of their children's time to allocate to schooling, y_i where *i* equals *s* for the older sister and *b* for the younger brother. Children's time can either be allocated to schooling or household production (at home or the family farm). Let y_i denote the share of child's time allocated to schooling, and $x_i = 1 - y_i$ the share allocated to household production. With the length of instruction time in the school day fixed, y_i can best be interpreted as years of schooling for child *i*.

For the oldest sister, we have the usual human capital production function, $H_s = f(y_s)$. Time allocated towards schooling, y_s, is the only input into older sister's learning or human capital, H_s. The function $f(y_s)$ is assumed to be continuous and twice differentiable such that $f_{y_s} > 0$ and $f_{y_s y_s} < 0$ so that years of schooling exhibit positive but diminishing marginal returns in the production of human capital.

My contribution lies in how I modify the human capital production function of the younger brother to allow for an impact of the older sister. I specify the younger brother's human capital production function as $H_b = g(y_b, x_s, H_s)$ so that younger brother's human capital is a function of his own years of schooling, the time the older sister spends at home, and the accumulated human capital of the older sister. Each of the inputs individually have positive and diminishing marginal returns in the production of human capital so that $g_{y_b > 0}$, $g_{x_s} > 0$, $g_{H_S} > 0$, $g_{y_b y_b} < 0$, $g_{x_s x_s} < 0$ and $g_{H_s H_s} < 0$. I further assume complementarities between the inputs with $g_{x_s, y_b} > 0$, and $g_{x_s, H_s} > 0$. Younger brother's years of schooling, y_b, is an input into his human capital, H_b. The older sister's time spent at home, x_s, is also an input into younger brother's human capital. The implicit assumption is that sister's time spent at home doing housework is also spent with the younger brother. This is reasonable because the older sister's responsibilities in the house include caring for younger siblings, and even if the sister is engaged in other household chores, this time is still spent at home and potentially interacting with younger brothers. Following Becker's treatment of parental time investment, I assume a complimentarity between time spent with sister and brother's own schooling investment in his human capital. The oldest sister's human capital also enters the production function directly because higher human capital increases the quality and productivity of the time x_s that she spends with the younger brother.

I allow the productivity of time allocated to household production, x_i, to vary by gender according to the parameters, p_s and p_b. I also model gender-specific cost of schooling with d_s and d_b, and interact these parameters with y_i so as to allow the total cost incurred to vary with years of schooling acquired. Since schools only charge nominal fees for attendance[11], the main cost of schooling captured by d_s and d_b is distance to girls' and boys' school. In period 2, the children who are now adults realize the returns to the

[11] The average annual fee charged by a government school was Rs. 10 in 2003 Pakistani rupees. This is approximately one-sixth of a USD according to exchange rates at the time.

human capital they accumulate in period 1. The parents who are no longer alive derive utility from their children's period 2 utility. Each child's utility is a function V of wealth W_i which equals his/her accumulated human capital H_i times the gender-specific wage rate r_i. Childen's utility translates into parents' utility according to gender-specific altruism parameters, a_b or a_s that are typically assumed to be ≤ 1 because parents may put less weight on children's utility as compared with their own. For reasons discussed in the previous section including strong son preference, reliance on sons' in old age, and daughters moving away after marriage, it is likely that a_s is less than a_b.

The parents' objective function then is to maximize the following utility $U(c)+a_sV(W_s)+a_bV(W_b)$ subject to the budget constraint $c+d_sy_s+d_by_b=W_p+p_s(1-y_s)+p_b(1-y_b)$ and the technologies specified above[12]. I assume the Inada conditions so that that we get interior solutions for consumption, c_p, oldest sister's years of schooling, y_s and younger brother's years of schooling, y_b. Parents choose y_b^* so as to satisfy the following first-order condition:

$$a_bV_{w_b}r_bg_{y_b}=(p_b+d_b)U'(c) \quad (1)$$

Equation (1) yields the standard result that parents will set younger brother's years of schooling y_b^* so that the utility gain due to the brother's increased wealth associated with a marginal increase in y_b (the left-hand side of the equation) equals the increase in the disutility associated with increased schooling costs including foregone consumption in period 1 and distance (the right-hand side of the equation).

Parents choose y_s^* so as to satisfy the following first-order condition:

$$a_sV_{w_s}r_sf_{y_s}+a_bV_{w_b}r_bg_{H_s}f_{y_s}=(p_s+d_s)U'(c)+a_bV_{w_b}r_bg_{x_s} \quad (2)$$

---

[12] U', V_{w_s}, V_{w_b}, f_{y_s}, g_{y_b}, g_{x_s}, g_{H_s} are positive and U'', $V_{w_sw_s}$, $V_{w_bw_b}$, $f_{y_sy_s}$, $g_{y_by_b}$, $g_{x_sx_s}$ and $g_{H_sH_s}$ are negative. The parameters W_p, a_s, r_s, r_b, p_s, p_b, d_s, and d_b are exogenously determined.

Equation (2) shows that as with the younger brother, parents will invest in older sister's schooling till it reaches y_s^* such that the marginal benefit from that investment equals its marginal costs. This comparison for the older sister, however, also involves the spillover effects on the younger brother. An increase in the older sister's schooling y_s increases not only the utility gain due to her own increased wealth, but also that due to the increased wealth of the younger brother through the "quality effect". Increasing the schooling and thus human capital of the older sister increases the quality of her time spent with the younger brother, and hence increases younger brother's human capital. The increased costs associated with increasing older sister's schooling include the foregone consumption and distance but also the "quantity effect" trade-off which increases the disutility from reduced wealth for the younger brother. Increasing the older sister's schooling entails less time spent with the younger brother and thus reduces his human capital and wealth.

The first-order condition in (2) demonstrates that an increase in older sister's schooling y_s is associated with two competing effects on younger brother's human capital acquisition, the positive quality effect and the negative quantity effect. Without knowledge of the relative magnitudes of these effects, the net impact of older sister's schooling is theoretically ambiguous and needs to be empirically determined.

The following expression shows condition 2 re-arranged so that the second term on the left hand side represents the net externality of the older sister's schooling on younger brother.

$$a_s V_{w_s} r_s f_{y_s} + a_b V_{w_b} r_b [g_{H_S} f_{y_s} - g_{x_s}] = U'(c)[p_s + d_s] \quad (3)$$

If $g_{H_s} f_{y_s} - g_{x_s} > 0$, the quality effect overrides the quantity effect, and older sister's schooling has a net positive spillover on younger brother human capital. This expression shows clearly that even if $a_s = 0$ and parents derive no utility from the sister's utility, it might still be optimal for them to invest in her schooling when her schooling generates positive spillovers for the younger brother.

It can be shown that when the net externality of older sister's schooling on younger brother is positive but parents fail to internalize the spillover in their schooling allocation decision, this leads to underinvestment in the older sister's schooling. In this case, parents would only be taking the private benefits and costs of the sister's schooling into account while ignoring the social benefits and costs for the younger brother. It would be optimal for a social planner to intervene in this case to increase investment in older sister's schooling in order to equate social returns with social costs. This idea is expressed more formally below.

Proposition. *Let parents' investment in older sister's schooling be* y_s^{**} *when net externality of the older sister's schooling on younger brother is positive and they take this externality into account when making the schooling decision. Let* y_s^{***} *denote parents' investment in older sister's schooling if they fail to internalize the spillover effect on younger brother. It can be shown that parents will under-invest in the older sister's schooling if they fail to internalize the net positive externality i.e.* y_s^{***} *is lower than* y_s^{**}.

Proof. See appendix A1. □

It is quite plausible that parents in rural Pakistan may not be aware of the spillover effects when making schooling decisions for their children. In this case, parents will fail to internalize the spillover in their choice of schooling allocations but the younger sibling will still be impacted by the spillover because it is present in the human capital production function. I will thus still be able to empirically estimate any spillover effects from older sister's education on younger brothers that exist, as long as the quality effect and quantity effect do not wipe each other out completely.

I assume that the net externality on younger brother human capital is positive for the purpose of calculating comparative statics. I discuss the comparative statics with respect to W_p, a_b, a_s, r_s, r_b, p_s, p_b, d_s, and d_b, and provide detailed proofs in the theoretical appendices A2 and A3[13]. Here I discuss the comparative static of brother's schooling with respect to cost of girl's schooling because that is the source of variation I use to identify the causal impact of oldest sister's schooling.

In a model without any spillovers, an increase in the cost of girl's schooling has an ambiguous effect on brother's schooling because it includes a positive substitution effect (brothers and sisters are competing

[13] Detailed proofs are provided in the theoretical appendix but I present a quick overview here. Children's schooling is increasing in parental wealth. Increasing the relative altruism towards girls results in higher schooling for the sister and lower schooling for the brother. Higher market return to girls' human capital increases sister's schooling while reducing brother's schooling, whereas higher market return to boys' human capital increases the brother's schooling and has an ambiguous impact on sister's schooling (since sister's schooling is also an input into brother's human capital).

for schooling investment from a common pool of resources so decreasing sister's schooling frees up resources for the brother), and a negative income effect (if the sister is getting schooling and incurring the cost, the higher cost squeezes household resources thereby leaving less for the brother). Since schools only charge nominal fees[14],, we expect the income effect of this cost which is a distance burden to be small so that the substitution effect is expected to dominate in the model without spillovers. Without spillovers, brother and sister schooling are hence likely to be in direct competition with each other so that an increase in girl's schooling cost may help increase brother schooling. Shrestha found evidence for such inter-sibling rivalry in Nepal where increases in boys' schooling reduced the schooling of female siblings.

In a model with spillovers, an increase in the cost of girl's schooling has three effects: the positive substitution effect, the negative income effect as well as a negative spillover effect (assuming the net externality of sister schooling is positive). This last new term captures the fact that distance-induced reductions in older sister's schooling lead to a reduction in the net positive spillover for the younger brother. Although the overall sign of the comparative static is ambiguous, relative to the model without spillovers, an increase in the cost of girl's schooling in a world with positive spillovers is associated with a more negative impact on brother's schooling.

4. Empirical Set-up

Conceptualizing Treatment Impact

The treatment of interest is years of schooling completed by the oldest sister, y_s. I consider three formulations of the treatment variable. I use years of schooling completed so that I can compare the marginal effect of an additional year of schooling to effects found in the literature for maternal years of schooling. I also formulate dummy variables for treatment using an indicator for whether the sister has any schooling, and an indicator for whether she has completed primary schooling i.e. 5 years of schooling. Defining treatment as

[14] The average annual fee charged by a government school was Rs. 10 in 2003 Pakistani rupees, approximately one-sixth of a USD according to exchange rates at the time.

an indicator for whether the oldest sister has had any positive schooling is important because one out of every four oldest sisters in my sample have had no schooling. This formulation tells us whether any positive amount of schooling of the oldest sister yields any benefits. The indicator for whether the oldest sister has completed primary schooling is an important threshold as 44.7 percent of oldest sisters have completed at least primary schooling.

Almost 40 percent of the oldest sisters in my sample are currently enrolled in school so the years of schooling observed is a right-censored variable for them. The average years of schooling completed by the oldest sisters in my sample during the panel is 3.97 years, with a standard deviation of 3.22. If I restrict attention to oldest sisters aged 15-20 inclusive, this average increases to 4.54 years. This right-censoring is not a cause for concern because I control for the age of the oldest sister in all my specifications. This means that the interpretation of the treatment effect is "how does the schooling acquired by the older sister, adjusting for her age, affect the outcomes of the younger sibling".

Even though the LEAPS data is a panel dataset, I deviate from the studies (e.g. Rosenzweig and Wolpin, and Moore and Schmidt) that employ fixed effects estimation, and identify the impact of mother's schooling/enrollment by looking only at mothers who go back to school between the births of two of their children. I do not use fixed effects because the identification would come solely from variation in oldest sister's education over time. That would give us the incremental, value added impact of an oldest sister *continuing* her education which does not answer my question of interest. The experiment associated with my research question is to think of a child growing up in a household with an oldest sister who was uneducated, and an otherwise comparable household in which the child's oldest sister had some education, and looking at whether the two children fare significantly differently in their accumulated learning and schooling.

One of my treatment formulations uses an indicator variable for whether the oldest sister has any education. For this treatment definition, a fixed effects approach would yield meaningless estimates because a switch in treatment status observed during the panel for the average oldest sister who is 16 years old is either a really weird person or more likely due to measurement error. When I consider the oldest sister's years of schooling as the treatment variable, only 46 percent of the oldest sisters have time-varying years of schooling and would contribute to the estimates of a fixed effects regression. When using the indicator for whether the oldest sister has primary schooling, the treatment variable is time-varying for only 30 percent of the oldest sisters in my sample. This means that identification of treatment effects in fixed effects specifications would come from a small, non-random sub-sample of the data. Throughout this paper, I have made the decision to not use fixed effects because the results obtained from fixed effects do not answer my research question, and they would be identified from a severely limited, non-random subset of the population.

Outcomes of interest

I am interested in analyzing the impact of oldest sister's schooling on literacy, numeracy, enrollment and schooling of younger brothers. Literacy and numeracy capabilities are important outcomes because these are the basic skills that primary schools are supposed to impart. These are also skills that are likely to be most important in rural, agrarian societies where the return to higher education is often limited. When I consider schooling of younger sibling as an outcome, I use years of schooling completed adjusting for sibling's age because most of the siblings are still continuing their schooling. The last outcome used is whether the younger sibling is currently enrolled in school, also adjusting for his/her age.

Reported literacy and numeracy

Literacy and numeracy indicator outcomes for the younger siblings are derived from responses to questions of the following format that were asked of the female respondent to the household survey:

C*an {name} read a postal letter or newspaper in any language?*

Can {name} write a postal letter or newspaper in any language?

Can {name} add or subtract?

Can {name} count?

Figure 3 shows the gender gap in the mean of these outcomes for younger brothers and younger sisters who are just under 10 years old on average,. While three-fourths of the younger siblings can count, about 65 percent can add and subtract. About 45 percent can read and only two in five can write. The progression of these skills by grade is documented in figure 4, highlighting the fact that while these capabilities may seem very primitive in developed country settings, the learning of children in rural Pakistan is quite poor and has significant room for improvement. A typical 9 year old child would be in class 3 if enrolled and, less than half of the children in class 3 are able to read and only 1 in 3 able to write. Andrabi et. al 2008 report that if a child were to leave school after class 3, he/she would most likely be unable to write a simple sentence in Urdu. Only 31 percent of the children in class 3 could use the word "school" in a sentence. Only 65 percent of the students in class 3 are able to subtract single-digit numbers, and only 19 percent are able to divide a 3-digit number by a single-digit number.

In the vast majority of cases, the female respondent answering these questions about capabilities is the children's mother. Since only a quarter of the mothers have received any education themselves, one may wonder whether uneducated mothers respond differently to questions about capabilities that they themselves may lack! If so, the reported literacy and numeracy responses may measure different things across children with mothers of different education. In order to test if there were any such systematic differences in capability reporting across mothers of varying education, I studied the relationship between mother's education and her arguably subjective response for ability of child to read/write/add/count after controlling for child's test score. The LEAPS data contains test scores for a sub-sample of children who were enrolled in school. I am unable to analyze the impact of oldest sister's schooling on younger brothers' test scores because these are available only for a select sample of children enrolled in school. This raises both sample selection concerns and yields samples that are too small to use with my IV.

The problem in simply looking at capability reports by education of mother is that the mother being educated will have an independent impact on the child's capabilities. A significant difference in reported child abilities by mother's education does not necessarily mean that educated mothers are using different standards to define whether they think the child can read. It could simply be that educated mothers teach their children to read. Since test score is an objective – albeit noisy – measure of the accumulated learning of the child, it should include all the learning impacts mother's education has had on the child. If I find that the mother's education significantly predicts child's reported abilities even after controlling for the test score, I take this as evidence of the mothers' responses capturing different capabilities for mothers of different education levels. I regressed each of the mother-reported capability measures on the most closely related test score[15] and a measure of mother's education[16]. Appendix tables A1 and A2 show that there is no statistical or economically significant relationship between mother's education and reported capability of child after controlling for child's test score. I conclude that child capability reports obtained from mothers are reliable and are not capturing significantly different concepts across mothers of varying education backgrounds.

Sample of Interest

Since I'm interested in the impact of oldest sister's education on younger brothers, the effective sample of interest is households that have at least one daughter living in the household who also has a brother younger than her. Fortunately, the large family sizes with an average of 5 children mean this is a condition that is easily satisfied. I limit analysis to households where the oldest sister is between 8 and 30 years old in round 1. It is important to have an upper bound on age of oldest sister because girls get married early in rural Pakistan and they move out of the parents' household upon marriage. Since I only know oldest sister's education if she is still living with her parents, the use of much older oldest sisters who have not been married yet could bias the results. To further avoid selection from potential marriage of oldest sisters, I use the oldest sister among the siblings still living in the household. In 73% of the cases, the oldest sister I use is the absolute oldest sister by birth rank. I discuss these sample restrictions and robustness checks I implement in more detail in the appendix.

I limit attention to the impact on younger brothers aged 5-12 years old inclusive for literacy and numeracy outcomes. Children aged 5-6 may be enrolled in kindergarten (*kacchi* class) and will have started to learn how to count and add. The average child in class 1 is 7.3 years old and that in class 5 is 11.7 years old. Figure 4 showed the progression of reading, writing, adding and counting abilities of children by class. By the time children are in class 5 i.e. roughly 12 years old, 92 percent can write, 97 percent can read, and essentially all of them can add and count. Since these literacy and numeracy skills are largely acquired in the primary school years of 5-12, I examine impact of oldest sister's education only in this age range. Of the households

---

[15] I regressed reported reading and writing capabilities on both language test scores, and reported addition and counting capabilities on math test score

[16] I used an indicator for any schooling as well as years of schooling of the mother. Additionally, I tried a specification where I interacted the mother's education with the child's test score to see if there is a gradient to the mother's report.

with an oldest sister of the right age range, 94 percent have a younger brother in this age range. For the outcomes of years of schooling completed and enrollment, there is no theoretical reason to limit attention to 5-12 year olds and I use a sample of 5-18 year old younger brothers. Indeed, schooling and enrollment are more elastic for children after the completion of primary schooling. For boys and girls of primary school-age, enrollment rates average above 80 percent while they are considerably lower for middle and high school grades which indicates more room for improvement and potential impact of oldest sister.

Summary Statistics

Table 1 shows the descriptive statistics for the 1215 households used in the analysis[17]. The average rural Pakistani household has 8 household members, of whom 5 are children, and the three adults are the two parents and a grandparent. These households have slightly more girls than boys but this is to be expected given that I selected these households conditional on having an older sister.. Almost all (94 percent) of the households own the house they reside in. Average food expenditures per capita were about 1 USD a day[18].

5. Empirical Strategy and Results

This paper employs two empirical strategies to identify the impact of oldest sister's education. First, I show OLS results under the assumption of selection on observables. Controlling for a rich set of characteristics, I find significant positive impacts of oldest sister's schooling on younger brother learning and schooling. I also present the only estimates of impact for younger sisters in this paper. I show evidence that suggests that OLS estimates are endogenous, and that there is a need for employing a strategy that is robust to selection on unobservables. I then use my preferred identification strategy which exploits the gender

[17] While the LEAPS data consists of four rounds, I do not include the second round of data in the analyses. Due to funding issues, only an abridged version of the survey was administered in round 2 in October-December of 2003, just months after the first round in February-April of 2003, leading to data inconsistencies across rounds. Literacy and numeracy questions were not asked in round 2. While some schooling information was collected, its interpretation is problematic because the academic year for Pakistani schools starts in March/April and I can not determine if the first and second round interviews for a household took place in the same or different academic year. Children who were interviewed twice in the same academic year will have their grade completion appear artificially depressed relative to that of children who were interviewed in separate academic years. Round 2 data also has seasonal differences because unlike all other round interviews which were conducted in the Spring, round 2 was administered in Fall/Winter.

[18] The market exchange rate at the time of the survey was Pakistani rupees 60 to 1 USD.

segregation of schools in Pakistan to analyze causal impacts of oldest sister's schooling on younger brothers. The IV results I find are larger than OLS results which is closely tied to the fact that IV is uncovering a local average treatment effect that is applicable to only a special subset of the population which has higher returns from oldest sister's schooling.

5a. OLS Estimation

OLS results controlling for a rich set of characteristics show that increased schooling for the oldest sister is significantly associated with improved learning outcomes and higher schooling for her younger brothers. Tables 2a and 2b show positive significant impacts of the oldest sister's years of completed schooling on younger brothers' literacy, numeracy, education and enrollment. For the outcomes of literacy and numeracy capabilities, the regressions shown use a sample of 5-12 years old or primary school-aged younger brothers whereas for enrollment and schooling, I use a sample of 5-18 year old younger brothers. Specifications 1-6 include a progressively richer set of controls as we move from the left to the right side of the tables. The set of characteristics I control for in the fullest specification, in column 6, include household size and composition, parents' education, wealth controls and district and year fixed effects[19]. We notice a significant attenuation of the effects in going from specification 1 to specification 2 which adds the controls for parents' education. This is to be expected because parents' education is an important determinant of both oldest sister's education and learning of their younger children.

In specification 3, I additionally control for household wealth and expenditure which matters for the educational attainment and learning of children. The next specification adds district fixed effects so it controls for any time-invariant quality of the district that affects learning and education outcomes differentially. Then I

[19] The household controls include a quadratic in: age of oldest sister, age of younger sibling, number of daughters, number of sons, and total family size, as well as a dummy for language spoken at home, and dummies for month of interview. Since age gap between oldest sister and younger sibling is a linear combination of oldest sister's age and age of the younger sibling, the inclusion of these two variables non-parametrically controls for the age gap as well. Parents' education controls include a set of indicators for whether the mother and father have any education, variables including their years of schooling completed and indicators for whether any of these are missing. As is the case with most developing country data, the survey does not contain information on income since it is so hard to measure. Instead I try to capture wealth and resources by controlling for whether the family owns any land, dummy for the type of house they live in (permanent i.e. made of kiln bricks, temporary i.e. made of mud bricks, or semi-permanent i.e. made of kiln and mud bricks), dummy for the type of water connection the house has, and a quadratic in food expenditure per capita.

add year fixed effects in order to soak up any factors (e.g. economic and weather conditions which affect the opportunity cost of schooling) that make certain years in the 2003-2006 panel better or worse for learning and schooling. The final specification includes all the controls mentioned but is more flexible than specification 5 in that it interacts the district and year fixed effects, allowing these effects to vary by each other. This is my preferred specification because it includes the most stringent controls. Note that once I have controlled for parental education, the results in columns 2-6 display a remarkable robustness to the addition of the controls just described. Given that the coefficients barely change across the successively fuller control specifications, and since this is a pattern that is upheld for all the analyses in this paper, I choose to display only the results for the most stringent specification i.e. equivalent to column 6, in all tables that follow.

Table 3 shows the impact of oldest sister's schooling on younger brother learning and schooling using the preferred control specification of column 6 with all three formulations of the treatment variable. The first panel shows the spillover impact estimated using years of schooling of oldest sister. Each additional year of schooling completed by the oldest sister is associated with 0.151 additional years of schooling and a 1.9 percent increase in probability of younger brother enrolled. It is also associated with a significantly increased probability of younger brother being able to read and write by 2.7 percentage and 1.9 percentage points, and add and count bu 1.8 and 1.3 percentage points respectively. The second and third panels of the table show impact estimates using the indicator variables for oldest sister having any education, and oldest sister having completed primary schooling. Both indicator variables are found to be significantly positively linked to higher literacy and numeracy, increased completed years of schooling and higher enrollment for younger brothers. The positive impact estimates found on younger brother learning and schooling indicate that the positive quality effect from increasing oldest sister's schooling appears to outweigh the negative quantity effect to yield an overall significant positive impact.

Gender of younger sibling

Here I estimate the impact of the oldest sister's schooling on younger brothers and younger sisters separately. This is the only place in the paper where I consider impacts for younger sisters since unfortunately my preferred identification strategy does not allow estimation for them. Younger sisters probably have more frequent interaction with the oldest sister, and it is likely that any role model effects from oldest sister's schooling would be stronger for younger sisters[20]. Given that there is a strong gender division in work, it is possible for oldest sister's schooling to have a negative impact on younger sisters that does not exist for younger brothers because increasing schooling for the oldest sister could mean increasing housework responsibilities get shifted to the younger sister. It is therefore not clear how spillover effects for younger sisters compare to spillover effects for younger brothers,

In table 4, the main treatment effects represent effects of oldest sister's education on younger brothers and the interaction term gives us the incremental effect on younger sisters' outcomes relative to that for younger brothers. I find all the main effects to be positive and statistically significant. All of the 18 interaction terms are positive, and all except two are statistically significant at the 5 percent level indicating that younger sisters benefit significantly more from the education of their oldest sister than younger brothers. Using oldest sister's years of completed schooling, the size of the interaction effect is roughly a fourth to a third of the coefficient on the main effect for reading and adding whereas it is over two-thirds the size of the main coefficient for the outcomes of writing and counting. The impacts for enrollment for younger sisters are double those for younger brothers.

Impact on Older Brothers

While the theoretical framework models important spillover effects from oldest sister's education on her younger brothers, it does not allow for any effect to flow from younger to older siblings. The mechanisms behind the impact of oldest sister's education on her younger brothers are inextricably tied to the role of the oldest sister as an *older* sibling. The oldest sister's education matters because she provides child care, she helps the younger siblings with their studies and is a role model. Therefore, the model predicts that oldest sister's education should have no discernable impact on her *older* brothers[21]. I can test whether the assumption of

[20] When the oldest sister acquires significant schooling in the rural Pakistani society where education levels for girls are very low, this acts as a powerful signal to the younger sisters that such accomplishments are possible, and that their parents will support their daughters in seeking that much education.

[21] While there may be some small spillovers for an older sibling from interacting with more educated younger siblings, I expect such an effect to be drastically smaller than the impact of an oldest sister on younger sibling for the reasons listed above. I also expect this pure interaction effect on older brothers to be negligible because there is very limited

selection on observables is reasonable by checking whether I find an impact of oldest sister's education on her older brothers using my OLS specification with the fullest controls. If my regression controls adequately deal with selection in education of the oldest sister, I expect the treatment variable to have no impact on older brother schooling and learning.

Table 5 shows that I find statistically significant impacts of oldest sister's education on older brother schooling and learning using OLS. The point estimates are qualitatively similar (sometimes smaller, sometimes bigger) to those found for impact of oldest sister's schooling on younger brothers. In percentage terms, these represent smaller effects because mean outcomes are higher for older brothers but these are still economically significant. The finding of significant positive impacts of oldest sister's education on her older brothers when we have no theoretical reason for expecting any such impact suggests that OLS estimation suffers from selection bias. There is likely something unobserved about the households which educate their oldest daughters more which also leads them to educate all their children – younger and older – more. Examples of such unobserved factors could include how much the parents value education or how much of their time they invest in their children's learning. In the presence of such unobserved factors, I expect the OLS estimates to be upward biased, and the true effects of oldest sister's education to be smaller than the OLS results obtained in tables 2-4. This highlights the need for an identification strategy that yields consistent results under the assumption of selection on unobservables.

5b Instrumental Variables Strategy

In my preferred identification strategy, I exploit the gender segregation of government schools in Pakistan to create quasi-experimental variation in oldest sister's schooling and use that to analyze the impact on younger brothers. Government schools are gender segregated at all levels of instruction. Private schools exist at the primary level but are typically co-ed and quite expensive as compared with government schools

---

interaction between older brothers and their younger sisters compared to older sisters' interaction with younger brothers whom they take care of.

which charge only a nominal attendance fee[22]. Government schools enroll more than 70 percent of the children in my sample, and more than 75 percent of the girls. Table 6 shows the summary statistics on schools in the data. The average village has about 7 schools, with 2 government boys' schools, slightly less than 2 government girls' schools and roughly 3 private schools. The average household is located 0.59 km from the closest government girls' school, and 0.61 km from the closest government boys' school.

I use a conditional instrument of distance to girls' school to vary oldest sister's schooling. Since younger sisters also attend girls' schools, distance to girls' school is not a good instrument for analyzing younger sister outcomes. As I show later, distance imposes a strong penalty on girls' schooling. Younger sisters' learning could be impacted by distance through its effect on their schooling. Additionally, distance could directly influence younger sister learning through affecting the number of absences or because time spent walking to a distant school is simply unproductive in terms of learning. For this reason, this paper restricts attention to the impact of oldest sister's schooling on younger brother outcomes.

I use distance to closest government girls' school, conditioning on distance to village center and distance to all possible boys' schools to create exogenous variation in the schooling of the oldest sister. There are a number of reasons why the conditional instrument of distance to girls' school satisfies the exclusion restriction, and does not have a direct impact in the outcome equations for younger brothers. I also present several checks which confirm the validity of this instrument. Since boys do not attend government girls' schools, distance to government girls' schools can only affect boys indirectly, through how it varies with other factors. First, I demonstrate that the usual selection concern with using distance to school in the US literature does not apply to my context of rural Pakistan. Second, I control for distance to village center to soak up any variation in distance to girls' schools which is driven by how centrally or remotely a household is located. Third, I control for distance to boys' schools in all my specifications. Fourth, I show that distance to boys' schools does not have any impact on oldest sister's schooling. Fifth, I implement a falsification test to show that using this conditional instrument, I do not find any impact of oldest sister's schooling on older brother

[22] The average annual fee charged by a government school was Rs. 10 in 2003 Pakistani rupees, approximately one-sixth of a USD according to the exchange rate at the time. The average annual fee charged by a private school was Rs. 870, or USD 14.

outcomes. This provides strong justification for distance to government girls' school being a valid instrument that is excluded from the outcome equation for younger brothers.

A valid concern with the use of distance to school as an instrument as used in the US literature is that household location is not random with respect to children's schooling options. Access to good quality schools is a very important factor in an American family's decision of where to live. In these rural Pakistani villages, the biggest factor which determines where a family lives is land. These villages are highly agrarian societies where people reside on and make a living from land that has been passed down in their family for generations. Half of the households in my data report owning land while 94% own the house they live in[23]. Fifty-eight percent of fathers work in agriculture, and this is likely a lower bound because I treat those who report working in salaried occupations as not working in agriculture. While many in salaried occupations will hold non-farm jobs, a sizable proportion whom I can not identify will include skilled people who operate farm equipment such as harvesters, threshers, tube wells, and tractors, for example.

Given the strong dependence of families on agriculture and land, these households have very little choice in where they locate. The only margin of mobility for these households is for an adult male member to migrate to the city. Even for the positively selected households where somebody manages to migrate to an urban area for work, it is only the adult male members who migrate and send back remittances to the wife and children. Restricted mobility aside, these households place a much smaller value on education as compared with US households making selection of household location based on schooling significantly less likely. The fact that parents value girls' education even less than that of boys works in favor of my specification. I'm controlling for distance to boys' school and only relying on variation in distance to girls' school and I expect very little endogenous location based on access to girls' school.

I use distance to closest government girls' school as an instrument while conditioning on distance to village center because households located far from girls' schools may also be households that are remotely located in general. Distance to girls' school would then be correlated with income and economic opportunities of the household which directly impact child schooling and learning. For this reason, I follow

[23] It is a common arrangement for landlords to give tenants a residence in addition to the share-crop payment. This leads to many tenant families owning a house but not owning any farmland of their own.

Andrabi et al in using distance to school, conditioning on the distance of the household from the population-weighted center of the village. Controlling for distance to center soaks up the variation in distance to school that is simply due to the household being in a central location versus being located out in the middle of nowhere. Most government schools were built on common land which is typically located on the periphery of the village because it was cheaper. Government schools locate largely on the periphery of these villages rather than at the center of the villages where most households are concentrated, particularly richer households with high demand for schooling. First, this indicates that location of government schools, unlike that of private schools that locate in the center, is not likely driven by demand considerations. Alderman et. al (1997) also concluded that the availability of schools in the community can be considered beyond the control of individual households in rural Pakistan. Second, this location of government schools allows there to be sufficient variation in distance to schools even after controlling for distance to center.

The gender segregation of government schools in Pakistan allows me to control for distance to all possible schools that the younger brothers could attend. I control for distance to the closest government boys' school and the closest private school in all IV specifications. Identification therefore comes from comparing households that are equidistant from the village center, equidistant to boys' schools, share all the characteristics of the other controls and differ only in their distance to closest government girls' school.

Finally, recall the falsification test which analyzes impact on older brothers. According to my model, the oldest sister's education matters because as the *older* sibling, she provides child care and helps the younger siblings with their studies. Oldest sister's education should therefore have no discernable impact on her older brothers' outcomes. Using OLS, I found that oldest sister's schooling had significant positive impacts on older brother learning and schooling raising concerns of endogeneity. Table 7 shows that using the conditional instrument of distance to girls' school, oldest sister's schooling does not have a significant impact on any of the older brother outcomes. This falsification test gives us confidence in the validity of the instrument[24].

[24] Far from finding a positive impact, I can actually never reject the one-tailed hypothesis that these estimates are non-positive.

Next I establish that distance to girls' school is a strong instrument, and describe why distance imposes such a strong penalty for girls before turning to results. Distance to school has previously been established as a strong determinant of girls' enrollment in this region – in Pakistan by Alderman, Orazem, Paterno (1996) and Andrabi et. al (2008), and more recently in Afghanistan by Burde and Linden (2010). Table 8 shows that an increase of 1 km in the distance to the closest government girls' school is accompanied by a reduction of 0.18 years in the completed schooling of 5-18 year old girls on average[25]. While the distance penalty attenuates after adding a control for distance to village center, as expected, distance to girls' school continues to have a statistically and economically significant impact on girls' schooling. In contrast, distance to school has a much smaller impact on boys' schooling and adding the control for distance to center causes distance to school to lose statistical significance in two of the three education specifications. Distance to the village center, however, is a significant negative determinant of boys' schooling as it is of girls' schooling. The difference in the penalty for distance to school and distance to village for girls and boys points to the underlying determinants of schooling. Both boys' and girls' schooling is negatively impacted by distance to center which captures the fact that well-off households which are located more centrally are more likely to educate their children. That distance to school impacts girls more strongly than boys captures the fact that girls' mobility outside the house is restricted by safety and chastity concerns, and imposes an additional cost on the schooling of girls in these villages.

The vast majority of children walk to school and the younger ones are accompanied by a relative or friend as they walk to and back from school. Ninety-four percent of children younger than 10 have somebody walk with them to school, and these figures are the same for boys and girls. As children age and become teenagers, however, a gender gap emerges. The fraction of boys accompanied by somebody to school – be it a relative, neighbor or friend – falls as they grow older since they are increasingly capable of taking care of themselves. The fraction of girls accompanied, however, does not decline with girls' age and leads to a statistically significant, 14 percentage point difference in the rates at which girls and boys aged 13 and older are accompanied on their commute to school. This gap emerges because girls' mobility is restricted by safety

[25] This is the distance penalty that exists even after controlling for household characteristics, parents' education, wealth and asset controls and distance times year fixed effects.

concerns and these concerns increase after girls reach the age of puberty. The need for households to ensure that somebody accompanies the girls to school is particularly burdensome as girls get older because the significant drop-off in girls' enrollment after primary school means there are fewer options for girls to walk to school with other girls who live in the neighborhood. The gender segregation of schools also implies that less than 15% of girls are walked to school by a brother. The additional cost imposed by the need for somebody to accompany a girl to school creates a distance penalty for girls which does not exist for boys.

Since the gender difference in the need for somebody to accompany girls and boys to school only arises for older children, one should expect that the gender difference in the distance penalty should arise only for older children. Table 9 shows the distance penalty for schooling and enrollment of boys and girls aged younger than 13, and aged 13 and older separately. While there is no statistically distinguishable distance penalty for schooling or enrollment of young or old boys, the distance penalty for girls is driven primarily by teenage girls. The distance penalty for schooling for girls younger than 13 years old is very small and not statistically distinguishable from zero while distance to girls' school is a strong determinant of schooling for girls 13 years and older. The distance penalty for enrollment of girls younger than 13 is negative and significant at the 10% level but the penalty for enrollment of girls aged 13 and older is 50% larger and highly statistically significant.

IV results

Table 10a shows the IV results for impact of oldest sister's years of completed schooling on younger brother literacy, numeracy, schooling and enrollment. The instrument used is distance to the closest government girls' school from the household, conditioning on distance to center and distance to the closest government boys' school and closest private school. The F statistics on the excluded variable of distance to girls' school in the first stage regressions show that distance to girls' school is a strong instrument. An increase in distance to closest girls' school of 1 km, holding constant distance to village center and a host of other characteristics[26], is associated with the oldest sister completing 0.40 fewer years of schooling. The IV results show statistically significant, big, positive impacts of oldest sister's years of schooling on younger

[26] As mentioned before, I use the fullest control specification which includes household controls, parents education, wealth and asset controls, as well as district*year fixed effects

brother's reading, adding and counting capabilities, years of schooling completed and enrollment. The point estimate for impact on ability to write is positive but lacks statistical significance.

An increase in the oldest sister's years of schooling of one year increases younger brothers' completed schooling by 0.42 years, and his probability of being enrolled by 7.5 percentage points. .It makes the younger brother 7.7 percentage points more likely to be able to read, 7.6 percentage points and 5.5 percentage points more likely to be able to add and count, respectively. Increasing oldest sister's schooling by one year increases her completed schooling by a third of a standard deviation. Since the average oldest sister has completed 4 years of schooling, an increase of one year represents a 25% increase in her schooling. Compared to the mean of these reported capabilities for younger brothers, the effect sizes represent a 14 percent and 9.6 percent increase in younger brothers' schooling and enrollment probability, and a 19, 12 and 7 percent increase in reading, adding and counting capabilities respectively.

Tables 10b and 10c show the IV results for younger brothers using the indicator for whether the oldest sister has had any education, and whether she has completed primary schooling. As with years of oldest sister's schooling, I find important beneficial impacts of oldest sister's education on younger brothers using these two indicator measures. The impact on younger brother's writing ability still lacks statistical significance. It is not clear whether this indicates that oldest sister's schooling does not have an impact on younger brother writing or whether it is too small to detect precisely. I consistently find significant impacts on reading, adding and counting which are all relatively basic skills compared to writing. It is possible that increases in oldest sister's schooling at the margins of schooling relevant for oldest sisters in this data are simply not capable of impacting higher abilities such as writing.

I find that having an educated sister has big, significant returns in terms of whether her younger brother is able to read, add, and count, and for his schooling and enrollment status. The impact of oldest sister having a primary education is qualitatively similar with significant positive impacts for the same outcomes. The positive spillover effects from having an educated sister and having a primary-educated sister are not statistically distinguishable from each other although it appears that having a primary-educated sister tends to generate somewhat bigger returns.

Note that distance to boys' schools does not matter for oldest sister's schooling in any of the first stages of the IV regressions. The coefficients are neither statistically nor economically significant in any of the 18 specifications shown in tables 10a-10c. If households that really valued both girls' and boys' schooling located significantly closer to schools, we would expect to see that picked up as a meaningful correlation between oldest sister's schooling and distance to boys' school. This evidence suggests that households that are located closer to schools do not appear to systematically value education differently.

Comparing IV results to OLS

The story of selection I described and the finding of significant positive impacts of oldest sister's schooling on older brother outcomes indicate that OLS results suffer from positive selection bias. Since IV estimates are free of selection bias, I expected the IV results to be smaller than the OLS results which are upwards biased. Instead, I find that the IV results are considerably larger than the effects found using OLS. While the vast majority of IV estimates are not statistically significantly different from the OLS estimates, the finding of larger IV estimates than OLS estimates in all 18 specifications suggests that this is a systematic trend[27]. This finding of IV coefficients exceeding OLS even though OLS results were expected to be upwards biased is common in the returns to schooling literature, as well as the literature on impact of parental education (for example, Currie and Moretti 2003 and Carneiro, Meghir and Parey forthcoming).

I show that my finding of IV estimates which exceed OLS estimates is closely tied to the fact that IV estimates represent treatment effects for a specific sub-population only, i.e. a local average treatment effect (LATE). Angrist, Imbens and Rubin showed that the IV estimates the LATE which is the average treatment effect only for the compliers i.e. the subpopulation of individuals who are induced into receiving treatment by the instrument. In a world of heterogeneous treatment effects, the LATE can be very different from the average treatment effect (ATE). I show below that households which have stronger compliance with the instrument also appear to have larger treatment effects. Another possible explanation for the IV treatment effects exceeding the effects found using OLS is the presence of non-classical measurement error in the

[27] I conduct the Hausman test to compare OLS and IV estimates. For treatment defined as years of schooling and any primary schooling, OLS and IV estimates are significantly different in only one out of 12 specifications. For the indicator of any schooling, OLS and IV estimates are significantly different in four out of six specifications.

treatment variable. However, I rule out measurement error as a significant factor in explaining why IV effects exceed OLS in the next section on robustness checks.

I provide evidence of heterogeneous treatment effects for different subgroups which indicates that treatment effects are larger for the subset of population which is impacted most by the instrument. I study the differences in distance penalty imposed, and the heterogeneity in treatment effects by level of maternal education. I show that the households which are most strongly moved by the instrument (and hence make a disproportionately bigger contribution to the LATE) are households with uneducated mothers, and that oldest sister's education has larger impacts in these households relative to households with educated mothers.

I consider two categories of households: i) households with uneducated mothers i.e. mothers with zero years of schooling and ii) households with educated mothers, i.e. mothers with any schooling. When a mother is educated (as measured by the indicator for any schooling), the father is almost certain to be educated as well. Only 30 of the 1200 households in my sample has an educated mother but uneducated father so that having an educated mother essentially means having both parents educated[28]. From a theoretical perspective, mother's education is a very important factor to vary when considering heterogeneity in impact of oldest sister's education because oldest sister's education is likely to be a close substitute for mother's education. Ninety-six percent of fathers report work as their primary occupation compared with only 11.5% of mothers[29]. Given the negligible labor force participation for mothers[30], increased market work and earnings are not important channels through which maternal education impacts her children. The main mechanism for impact of maternal education on her children is through her role as the primary child care provider, and the time she spends with her children. Fathers spend significantly less time than mothers at home with their children. Given that I modeled younger siblings' time spent with oldest sister as the main mechanism through which oldest sister's education impacts younger siblings, oldest sister's education is a much closer substitute for mother's education as compared with father's education.

[28] Since very few households have an educated mother but uneducated father, I drop them from the subgroup analysis.
[29] This figure drops to 7.15% if I exclude work in agriculture or herding that is potentially performed on the family farm.
[30] In time use data, only 10.5% of mothers report spending any positive time on paid work in a given day.

Table 11a shows how the distance penalty for oldest sister's schooling differs across households with uneducated and educated mothers. Distance imposes a strong, highly statistically significant penalty on oldest sister's schooling in households with uneducated mothers whereas the impact of distance is considerably attenuated and lacks statistical significance for households with educated mothers. While small sample size for households with educated mothers contributes to the lack of statistical significance by inflating standard errors, the first stage coefficient estimates for the educated mother sample are 25-60% smaller than the coefficients for households with uneducated mothers. The instrument has a stronger impact on oldest sister's education in households with uneducated mothers as compared with households with educated mothers[31]. Recall that both parents are educated when the mother is educated in this sample. One would expect households with both parents educated to place high value on education of their children. While counterfactual outcomes are unobserved, it makes sense intuitively that both parent educated households will typically educate the oldest sisters highly regardless of distance to girls' school i.e. be in the set of "always takers"[32]. Ninety-five percent of the oldest sisters in both parent educated households have some schooling and 63 percent have completed primary schooling. Similar figures for mother uneducated households are 69 percent and 39 percent, respectively. It is thus plausible that households with both parents educated are more likely to be "always takers" and hence are relatively less responsive to the instrument.

Table 11a showed that IV estimates will be disproportionately based on households with uneducated mothers because these households comply more strongly to the instrument as compared with households with educated mothers. Next, I explore how treatment effects vary by mother's education using OLS. While the preferred identification strategy is to use IV estimation, it is unfortunately infeasible for the set of households with educated mothers. As just shown in table 11a, the instrument of distance to girls' school is very weak for this subsample with the F statistic on the instrument in the first stage around 1 across all outcomes. Table 11b then relies on OLS to estimate treatment effects for oldest sister's years of schooling separately for households with uneducated mothers and households with educated mothers. It shows that

[31] While the first stage coefficients are not statistically significantly larger for uneducated mother households relative to educated mother households, the estimates are considerably different and the trend is consistent across all outcomes.
[32] Using the language from Angrist, Imbens and Rubin, in this context, "always takers" refers to households which get treated i.e. educate the oldest sister regardless of the value of the instrument i.e. distance to girls school.

treatment impacts of oldest sister's schooling are consistently larger for all outcomes in households with uneducated mothers. While the treatment impacts are statistically significantly different for adding and counting ability only, the impacts for all outcomes are considerably larger in households with uneducated mothers as compared with households with educated mothers. The impacts for reading, writing, schooling and enrollment are 75%, 18%, 46%, and 100% larger in uneducated mother households relative to educated mother households. Significant impacts for adding and counting ability only exist in households with uneducated mothers and they are wiped out to zero in households where the mother has any schooling.

In summary, I have shown evidence that suggests the reason IV effects exceed OLS effects is related to the fact that the LATE estimated by the instrument of distance to girls' school is likely capturing marginal returns for complier households which are higher than the return for non-complying households. Households which comply the most strongly with the instrument are households with uneducated mothers for whom the returns to oldest sister's schooling are considerably higher[33]. Since oldest sister's education is a close substitute for mother's education, it makes sense that the impact of oldest sister's schooling is higher in households with uneducated mothers relative to households with educated mothers.

Putting effects into context

In order to put the treatment effects into context, I compare the estimates to the impact of mother's schooling as found in past studies, and also calculate impact of mother's and father's schooling in my data. Using OLS, I found the impact of an additional year of schooling for the oldest sister on younger brothers' schooling to be 0.151 years. In the literature, the median impact of an additional year of schooling for the

[33] I also divided the households into three categories of i) neither parent educated, ii) father educated but mother uneducated and ii) both father and mother educated. A parent is considered "educated" if he/she has any schooling. I find that the largest treatment effects are for households with neither parent educated and mother uneducated. There is a slight attenuation of treatment effects for households with father educated but mother uneducated relative to the category of households where neither parent is educated. Treatment effects attenuate sharply when comparing the impacts for households with father educated and mother uneducated to impacts from households with both parents educated indicating that oldest sister's education is particularly important when the mother is uneducated. The strongest compliers to the instrument are households with father educated but mother uneducated suggesting that IV captures a LATE which is disproportionately comprised of households with educated fathers but uneducated mothers for whom treatment impacts are quite large.

mother on her children's schooling is 0.23 years[34]. The impact of an additional year of schooling for the oldest sister from my preferred IV specification is 0.42 additional years of schooling for younger brothers. These figures are at the higher end of the range of impacts of mother's education – 0.02-0.65 years – reported by Behrman although it is not clear that these estimates are directly comparable[35].

To make a better comparison, I estimate the effect of parents' education on younger brothers' learning and schooling in the LEAPS data. I use the same sample and control specifications I used in the analysis for the impact of oldest sister's education except that I do not control for oldest sister's education because that would amount to conditioning on an outcome. I find significant positive effects for both father's and mother's education as measured by the indicator for any schooling as well as years of schooling completed in table 12. The OLS impacts of oldest sister's years of schooling on younger brothers in table 3 are larger than the impacts of father's and mother's years of schooling. Since very few mothers have any schooling, one may expect that any meaningful impact of mother's education would be on the extensive margin. I find that having an educated mother is associated with strong positive impacts but the OLS impact of oldest sister having any education is large even compared to the effect of mother having any education. The impact estimates for oldest sister are larger than mother's in 5 out of 6 outcomes. Having a mother with some education is associated with 0.42 more years of schooling for younger brothers while the similar impact of oldest sister is 0.66 additional years of schooling.

These estimates for impact of mother's schooling and the impact of oldest sister's schooling are not directly comparable for a number of reasons. First, I lack an instrumental variable for parental education although a comparison of the OLS results also points to impacts for oldest sister's schooling that are large in comparison to impacts of mother's schooling. Second, I try to keep the specifications and controls similar when analyzing impact of parents' education and oldest sister's education but this means that the specifications are over-controlled for finding the effect of parents' education. Since the specification analyzing

[34] Behrman, 1997

[35] The impact estimates summarized by Behrman include results from OLS and IV specifications in many countries for children of different age ranges, and different instruments move people on different margins for whom treatment impacts may be very different. In this paper, IV estimates were much bigger than OLS estimates suggesting heterogeneous treatment effects

the impact of mother's schooling controls for the father's schooling as well as household wealth and assets, it is controlling for important mechanisms through which mother's schooling impacts her children via assortative mating. I can only conclude that the evidence suggests the impact of the oldest sister's schooling on younger brothers is very important and large even compared to impact of maternal schooling.

There are several reasons why I find impacts of oldest sister's education that are quite large. Oldest sister's education may be a more important input into the learning of children because of the type and frequency of interaction of children with peers close to them in age. The oldest sister was the most important source of help with studies for young children. Oldest sister's education also has the potential to have a particularly meaningful impact in this setting because parents and especially mothers are poorly educated. I showed evidence that suggests the impact of oldest sister's schooling is larger in households with uneducated mothers. In 75% of the households, the mother is uneducated and in 35% of the households, both parents are uneducated. Since the oldest sister is one of the first household members to get any education in a sizable fraction of these households, we would expect her schooling to generate large spillovers.

6. Robustness Checks

The last section briefly mentioned the possibility of IV effects exceeding OLS effects due to the presence of non-classical measurement error. I explore this possibility in this section, and show that the prevalence of measurement error in the binary treatment variables is quite low. I present estimates that explicitly incorporate the presence of non-classical measurement error in all treatment formulations (binary and discrete), and show that although these estimates are relatively attenuated, they remain considerably larger than the OLS effects. Measurement error in the treatment variable does not account for the IV effects exceeding OLS effects.

Two of the three treatment formulations I use are binary variables – indicator for any schooling and indicator for at least primary schooling – and the third uses a discrete variable – years of schooling completed by the oldest sister. Let D* denote the true treatment variable, D the observed treatment variable, and U any measurement error that captures the difference between D and D*. Any measurement error in binary variables is mean-reverting by construction with $\sigma_{D^*,U} < 0$ (Aigner, 1973) which creates upward bias in IV estimates (Black, Berger and Scott, 2000 – BBS hereafter). Any measurement error in the indicator for any schooling and the indicator for at least primary schooling completed by the oldest sister will cause the IV estimates to be biased upwards. Bound, Brown and Mathiowetz point out that if D* has a limited range in a discrete variable as with educational attainment, there will be a tendency for $\sigma_{D^*,U} < 0$ because when D* is at a maximum (minimum) of its range, reporting errors can only be negative (positive). The IV estimates using the discrete variable of years of schooling completed are hence also likely to be upward biased.

Frazis and Loewenstein (F&L hereafter) propose a technique to compute lower bounds of IV estimates for binary explanatory variables under relatively weak assumptions. I implement their method to provide lower bounds for the impact estimates from the binary treatment specifications which incorporate upper bounds on measurement error. With binary explanatory variables, the prevalence of measurement error is more naturally thought of in terms of probabilities of false negatives and false positives. Defining the error probabilities as $\alpha_0 = \Pr(D = 1 | D^* = 0) = \Pr(U = 1 | D^* = 0)$ and $\alpha_1 = \Pr(D = 0 | D^* = 1) = \Pr(U = -1 | D^* = 1)$, F&L and Bound, Brown and Mathiowetz have shown that the IV estimate relates to the true treatment effect, β ,as follows: $\beta_{IV} = \beta(1 - \alpha_0 - \alpha_1)$. F&L show that we can estimate lower bounds of the true treatment effect by finding upper bounds on α_0 and α_1 under assumptions and a procedure that I describe in appendix C. Table 13 shows the estimates of α_0 and α_1, the 95 percent confidence intervals for each, as well as the implied adjustment factors which need to be multiplied to the IV estimates to get the lower bounds on the true treatment impacts. The upper bounds of α_0 and α_1 from their 95% confidence intervals at 11.6 percent and 3.1 percent for the indicator whether the oldest sister has any education show that the prevalence of measurement error is quite low. The relative magnitudes of these bounds are intuitively plausible too. It seems less likely that parents would forget or neglect to report their oldest daughter acquiring any education rather than to report their daughter did receive an education when she did not. The adjustment factors using the upper bound from the 95 percent confidence intervals for α_0 and α_1 yield lower bounds of the treatment effects which are still sizable. The lower bounds of effects for the indicator that oldest sister has any education are at least 85% of the original IV results and those for the indicator that the oldest sister has completed at least primary schooling are at least 89% of the original IV results reported in section 6.

In order to obtain IV estimates that are consistent in the presence of non-classical measurement error when using the discrete formulation of treatment, I follow a technique proposed by BBS who show that one can obtain consistent estimates using a generalized method of moments (GMM) technique when one has two noisy reports on the mismeasured variable of interest. F&L extend this technique to show that the second measure can be replaced by one instrument to obtain consistent estimates using GMM. They also show that allowing for the mismeasured explanatory variable to be endogenous renders the GMM method with one instrument underidentified. I have strong reason to suspect the treatment variable of oldest sister's schooling is endogenous. Therefore, I allow for an endogenous mismeasured treatment variable, and use two instruments to obtain consistent estimates using GMM. The first instrument is distance to closest government girls' school which I have been using. I use an indicator for whether the closest government girls' school offers instruction in classes higher than primary as the second instrument[36]. I describe the model that allows for a general, non-classical measurement error process, and its identification using the two instruments and population moment equations in Appendix C. Table 14 shows the GMM estimates of treatment impact obtained from the model just described which allows for general, non-classical measurement error in the treatment variable of oldest sister's years of schooling, and allows for endogeneity of this treatment variable. The results mirror those found for the binary treatment specifications. The GMM impact estimates are attenuated relative to the IV results but are still considerably larger than the OLS estimates for five out of six outcomes while it is roughly the same as the OLS estimate for the sixth outcome.

[36] There are two types of government schools in my data: those that offer primary classes and those that offer instruction in primary classes and classes higher than primary. While the instrument suffices for the purpose of this robustness check, I do not use this instrument in my preferred identification strategy for two reasons. First, it is a weak instrument and second, with less than two government girls' schools per village, the dominant source of variation in this instrument is inter-village variation in the classes that girls' schools offer.

I also conduct robustness checks by limiting the sample to oldest sisters less than 19 years old to better avoid any selection due to marriage of oldest sister. I avoid most of the bias from marriage because I use oldest sister among the siblings living in the household instead of the oldest according to birth rank. I also restrict the sample to analyze impact of oldest sister according to birth rank only. The impact estimates are robust to both checks and I find significant positive impacts on younger brother schooling, enrollment, reading, adding and counting. A fuller discussion and tables are presented in the appendix.

7. Conclusion

This is the first study to conceptualize oldest sister's schooling as an input into younger sibling human capital. I propose a theoretical model that predicts competing effects of increasing oldest sister's schooling on younger brother human capital: a positive quality effect because increasing oldest sister's schooling improves the quality of the time spent with her, and a negative quantity effect because increasing oldest sister's schooling requires more time in school and thus less time with the younger brother. I find that oldest sister's schooling has significant positive impacts on younger brothers' completed years of schooling, enrollment, literacy and numeracy, indicating that the positive quality effect outweighs the negative quantity effect.

This study makes an important contribution to our understanding of human capital production: it is the first to highlight a significant role played by older siblings in addition to that of parents that has long been recognized. These positive impacts of oldest sister's schooling have important implications for policies targeted at increasing girls' education such as gender-targeted conditional cash transfer programs. The model shows that parents will under-invest in girls' schooling if they fail to take the positive externality from oldest sister's schooling into account in schooling investment decisions. Parents in the rural, agrarian villages of Pakistan place low value on schooling, especially for girls as is evidenced by the significant gender gap in schooling. These poorly educated parents are quite unlikely to internalize the spillover benefits from oldest sister's schooling. Since this would lead to underinvestment in girls' education from a social standpoint, there is a role for government intervention to remedy this through gender-targeted education policies. If parents effectively internalize these spillovers on younger siblings, there is no inefficiency due to the externality but governments might still want to intervene for equity concerns or inefficiencies that exist for other reasons. Other potential sources of inefficiency include differences in parents' and girls' own valuation of girls' schooling because daughters move out of the parents' household after marriage and do not support parents in old age, as well as externalities for the girls' future children that may not be internalized. My findings have implications for the design and evaluation of government education policies in this case too. Policymakers aiming to close the gender gap in education should take into account the fact that increasing oldest sister's education will also have the unintended consequence of increasing the education of her younger brothers. The results suggest that evaluations and cost-benefit analyses of policies targeting girls' education which consider only effects on the girls and their children while ignoring potential impacts on younger siblings may systematically underestimate their total benefits.

References

Aigner, Dennis J., 1973, "Regression with a Binary Independent Variable Subject to Errors of Observation," Journal of Econometrics, 1, 49-60.

Alderman, Harold, Behrman, Jere R., Khan, Shahrukh, Ross, David, R., and Richard Sabot, 1997, "The Income Gap in Cognitive Skills in Rural Pakistan." Economic Development and Cultural Change, October 46:1, pp. 97-122.

Alderman, Harold, Orazem, Peter F., and Elizabeth M. Paterno, 2001, "School quality, school cost and the public/private school choices of low-income households in Pakistan", Journal of Human Resources, Spring, 36, pp. 304-326.

Andrabi, Tahir, Das, Jishnu, Khwaja, Asim I., Vishwanath, Tara, and Tristan Zajonc, 2008. "Learning and Educational Achievements in Punjab Schools: Insights to Inform the Policy Debate." World Bank, Washington D.C.

Andrabi, Tahir, Baum Natalie, Das, Jishnu, and Asim I. Khwaja, 2010 "Bad Public Schools are Public Bads: Civic Values and Test-Scores in Public and Private Schools" unpublished manuscript

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, 1996, "Identification of casual effects using instrumental variables", Journal of the American Statistical Association 91:444-472.

Becker, Gary S., 1964, Human Capital, New York, NY: Columbia University Press

Behrman, Jere R., 1997, "Mother's Schooling and Child Education: A Survey", Penn Institute of Economic Research Working Paper

Behrman, Jere R., Andrew Foster and Mark R. Rosenzweig, 1997, "Women's Schooling, Home teaching, and Economic Growth" Journal of Political Economy, 107(4) :632-714

Black, Dan A., Berger, Mark C., and Frank A. Scott, 2000, "Bounding Parameter Estimates with Nonclassical Measurement Error", Journal of the American Statistical Association, September, 95(451): 739-748.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, 2005, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," Quarterly Journal of Economics, May, 120(2), pp. 669-70

Bound, John, Charles Brown, and Nancy Mathiowetz, 2002, "Measurement Error in Survey Data." In *Handbook of Econometrics*, volume 5, ed. James Heckman and Ed Leamer, 3705-3843. New York: Springer-Verlag

Brooks-Gunn, Jeanne, and Lisa B. Markman, 2005, "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness." The Future of Children, 15(1), pp.139-168

Burde, Dana and Leigh L. Linden, 2010, "The Effect of Village-Based Schools: Evidence from a RCT in Afghanistan", unpublished manuscript

Butcher, Kristin F. and Anne Case, 1994,"The effect of sibling composition on women's education and earnings" Quarterly Journal of Economics, 109: 531-563

Currie, Janet and Enrico Moretti, 2003, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings" Quarterly Journal of Economics, November, 118:4, pp.1495-532.

Edmonds, Eric V., 2006, "Understanding Sibling Differences in Child Labor", Journal of Population Economics, 19, pp. 795-821.

Federal Bureau of Statistics, 2008, "Pakistan Social and Living Standards Measurement Survey (PSLM) 2006-07," Islamabad, Pakistan

Frazis, Harley and Mark. A. Loewenstein, 2003, "Estimating Linear Regressions with Mismeasured, Possibly Endogenous, Binary Explanatory Variables" Journal of Econometrics, 117, pp. 151-178

Haveman, Robert and Wolfe, Barbara, 1995, "The Determinants of Children's Attainments: A Review of Methods and Findings" Journal of Economic Literature, December, 33(4), pp. 1829-78.

Hoff, Erika, 2003, "Causes and Consequences of SES-related Differences in Parent-to-Child Speech" In M. H. Bornstein and R. H. Bradley (Editors), *Socioeconomic Status, Parenting, and Child Development*, pp. 147-160, Mahwah, NJ: Erlbaum

Huttenlocher, Janellen, Haight, Wright, Bryk, Anthony, Seltzer, Michael, and Thomas Lyons, 1991, "Early Vocabulary Growth: Relation to Language Input and Gender." Developmental Psychology, 27, pp.236-248.

Ilahi, Nadeem, 2001, "Children's Work and Schooling: Does Gender Matter? Evidence from the Peru LSMS" Working Paper No. 2745, World Bank Policy Research

Levinson, Deborah and Karine S. Moe, 1998, "Household Work as a Deterrent to Schooling: An Analysis of Adolescent Girls in Peru" Journal of Developing Areas 32(3), pp. 339-356

Maccoby, E. Eleanor, 1966, "Sex Differences in Intellectual Functioning." In E. E. Maccoby (Ed.), *The Development of Sex Differences*. Stanford, CA: Stanford University Press

Moore, Quinn and Lucie Schmidt, 2007,"Do Maternal Investments in Human Capital Affect Children's Cognitive Development?", unpublished manuscript, Williams College

Rosenzweig, Mark R., and Kenneth I. Wolpin., 1994, 'Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement', The Journal of Human Resources, Spring, 29(2), pp. 670–93.

Sen, Amartya, 1992, "Missing Women," British Medical Journal, 304, pp. 587-588

Sen, Amartya, 1998, "Mortality as an Indicator of Economic Success and Failure," Economic Journal, 108(1), pp. 1-25

Shrestha, Slesh A., 2011, "Sibling Rivalry in Education: Estimation of Intra-household Trade-offs in Human Capital Investment", working paper, University of Michigan

Strauss, John, and Duncan Thomas, 1995, "Human Resources: Empirical Modeling of Household and Family Decision," *Handbook of Development Economics*, 3, pp.1885-2005.

Summers, Lawrence, 1993, "Foreword" in King, Elizabeth M. and M. Anne Hill, eds, 1993, Women's Education in Developing Countries: Barriers, Benefits and Policies, Baltimore and London: The Johns Hopkins University Press, World Bank, v-vii

Thomas, Duncan, Strauss, John and Maria-Helena Henriques, 1991, "How Does Mother's Education Affect Child Height?" Journal of Human Resources, Spring, 26(2), pp.183-211.

UNDP, 1996, Human Development Report 1996, New York: UNDP

A Model Appendix

A1 Proof of proposition stated in Section 2.

Proposition. *Let parents' investment in older sister's schooling be* y_s^{**} *when net externality of the older sister's schooling on younger brother is positive, and parents take this externality into account when making the schooling decision. Let* y_s^{***} *denote parents' investment in older sister's schooling if they fail to internalize the spillover effect on younger brother. It can be shown that parents will under-invest in the older sister's schooling if they fail to internalize the net positive externality i.e.* y_s^{***} *is lower than* y_s^{**}.

Consider the case where the net externality of the older sister's schooling on younger brother human capital is positive so that $g_{H_S} f_{y_s} - g_{x_s} > 0$. In words, the positive quality effect of increasing older sister's schooling outweighs the associated negative quantity effect.

Suppose parents are unaware of the spillover effects of older sister's schooling on younger brother, and hence do not internalize the spillover benefit and cost associated with sister's schooling in deciding about schooling investments. Such parents would consider the younger brother's human capital production function to simply be a function of his own years of schooling with $H_b = g(y_b)$. Such parents will choose y_s^{***} so as to satisfy the following first-order condition where they only consider the private benefits and costs of the older sister's schooling and ignore the spillover on the brother:

$$a_s V_{w_s} r_s f_{y_s} = U'(c)[p_s + d_s] \quad (4)$$

Now consider the case where parents internalize the net externality of the older sister's schooling on younger brother human capital. These parents will choose y_s^{**} so as to equate the complete marginal benefits of older sister's schooling with the complete marginal cost as indicated by the following condition.

$$a_s V_{w_s} r_s f_{y_s} + a_b V_{w_b} r_b [g_{H_S} f_{y_s} - g_{x_s}] = U'(c)[p_s + d_s] \quad (5)$$

Compared to (4), this equation has an additional positive term on the left-hand side because parents are aware that the positive externality from the older sister's schooling increases their utility by increasing the younger brother's human capital and wealth. In order to equalize the marginal utility from investment in sister's schooling with its marginal cost, parents will increase the investment in the older sister's schooling. This has the effect of decreasing the first term on the left-hand side in equation (5) since $f_{y_s y_s}$ is negative, and hence equalizes the left-hand side with the right-hand side. The equilibrium investment in older sister's schooling in this case, y_s^{**}, is therefore higher than y_s^{***}, the investment that occurred in equilibrium when parents did not internalize the spillovers of sister's schooling on younger brother.

A2 Derivation of comparative static with respect to distance to girl's school

I discuss the comparative static of brother and sister schooling with respect to distance cost of girl's schooling two ways: in a model with spillover and a model without any spillover.

In a model without any spillovers, the SOCs are as follows;

MB of y_b with respect to y_b:

$a_b V''(W_b) r_b^2 g_{y_b}^2 + a_b V'(W_b) r_b g_{y_b y_b}$. Let this equal Y.

MB of y_s with respect to y_s:

$a_s V''(W_s) r_s^2 f_{y_s}^2 + a_s V'(W_s) r_s f_{y_s y_s}$. Let this equal Z.

$\dfrac{\partial y_s}{\partial d_s} = YU' - y_s(p_s + d_s) YU'' + (p_b + d_b)^2 U'U''$ which is negative. Increase in the distance cost of girl's school decreases sister's schooling because sister's schooling is more costly. Also, for a sister who is getting some schooling (non-corner solution for y_s), there is a negative income effect from incurring the distance cost. $\frac{\partial y_b}{\partial d_s} = -(p_b + d_b)(p_s + d_s)U'U'' - y_s(p_b + d_b)ZU''$, where the first term is positive and the second is negative. The first term is a positive substitution effect (brothers and sisters are both competing for schooling investment from a common pool of resources) while the second term is a negative income effect (if there is a non-corner solution for sister schooling i.e. the household is incurring the distance cost, the higher distance cost squeezes household resources thereby leaving less for the brother). Since the distance cost of schooling is just a time cost and schools only charge nominal fees, I expect the income effect of this distance cost to be small so that the substitution effect is expected to dominate.

In a model with spillovers, the SOCs are as follows;

MB of y_b with respect to y_b:

$a_b V''(W_b) r_b^2 g_{y_b}^2 + a_b V'(W_b) r_b g_{y_b y_b}$. Let this equal A.

MB of y_b with respect to y_s

$a_b V''(W_b) r_b^2 g_{y_b} g_{y_s} + a_b V'(W_b) r_b \, g_{y_b y_s}$. Let this equal B.

MB of y_s with respect to y_s:

$a_s V''(W_s) r_s^2 f_{y_s}^2 + a_s V'(W_s) r_s f_{y_s y_s} + a_b V''(W_b) r_b g_{y_s}^2 + a_b V'(W_b) r_b g_{y_s y_s}$. Let this equal C. Notice that C = Z + $a_b V''(W_b) r_b g_{y_s}^2 + a_b V'(W_b) r_b g_{y_s y_s}$. Define K = $a_b V''(W_b) r_b g_{y_s}^2 + a_b V'(W_b) r_b g_{y_s y_s}$.

I assume the first- and second-order conditions described in section 2, that the net externality of sister's schooling on younger brother human capital is positive with diminishing returns ($g_{y_s} > 0$ and $g_{y_s y_s} < 0$). These assumptions yield that

$g_{y_s y_s} = g_{H_s} f_{y_s y_s} + g_{H_s H_s} f_{y_s} - g_{H_s x_s} f_{y_s} + g_{x_s x_s} - g_{x_s H_s} f_{y_s}$ which is negative because $g_{H_s} > 0$, $f_{y_s} > 0$, $f_{y_s y_s} < 0$, $g_{H_s H_s} < 0$, $g_{x_s x_s} < 0$ due to the assumption of positive and diminishing marginal returns of the inputs, and $g_{H_s x_s} > 0$ due to the assumed complimentarity between sister's human capital and time spent with brother. I assume that the term B i.e. the differential of the MB of y_b with respect to y_s is negligible but also separately describe the results that would hold if there was a strong complimentarity between y_b and y_s such that B>0.

$$\frac{\partial y_s}{\partial d_s} = AU^{'} + y_s(p_b + d_b)BU'' - y_s(p_s + d_s)AU'' + (p_b + d_b)^2 U'U''$$ which is negative as in the model without spillovers[37].

$$\frac{\partial y_b}{\partial d_s} = -(p_b + d_b)(p_s + d_s)U'U'' - y_s(p_b + d_b)ZU'' - y_s(p_b + d_b)KU''$$
$$+[y_s(p_s + d_s)U'' - U']B$$

which is ambiguous as in the model without spillovers. Notice that while the positive substitution effect is the same in both models, even if we assume B=0, the model with spillovers picks up an additional term due to the spillover. Now, an increase in the distance cost of girl's schooling has three effects: the positive substitution effect, the negative income effect as well as a negative spillover effect. This new term captures the fact that distance-induced reductions in older sister's schooling lead to a reduction in the spillover benefits for the younger brother. Although the overall sign of the comparative static is still ambiguous, relative to the model without spillovers, an increase in the distance cost of girl's schooling in a world with positive spillovers is associated with a more negative impact on brother's schooling[38].

A3 Derivation of comparative statics with respect to parental wealth, parental altruism, returns to schooling and productivity in household production

Next I calculate comparative statics with respect to the exogenous parameters, W_p, a_b, a_s, r_s, r_b, p_s, and p_b which yield the following propositions.

---

[37] This expression is negative if B is negligible or if B>0.

[38] If B>0, then the expression is even more negative because there is a strong complimentarity between sister and brother schooling such that there is pressure for brother schooling to fall in response to the distance-induced decrease in sister schooling.

1. $\frac{\partial y_b}{\partial W_p} > 0$ and $\frac{\partial y_s}{\partial W_p} > 0$: Wealthier parents invest more in children's schooling

2. $\frac{\partial y_s}{\partial r_s} > 0$ and $\frac{\partial y_b}{\partial r_s} < 0$: Parents substitute away from brother's schooling and towards sister's schooling when market returns to sister's human capital increase

3. $\frac{\partial y_b}{\partial r_b} \gtrless 0$ and $\frac{\partial y_s}{\partial r_b} \gtrless 0$: The impact on brother and sister schooling is ambiguous when market returns to brother's human capital increase.

4. $\frac{\partial y_b}{\partial a_s} < 0$ and $\frac{\partial y_s}{\partial a_s} > 0$: Parents substitute away from investing in brother's schooling towards investing more in the sister's schooling when they care more about the sister's utility.

5. $\frac{\partial y_b}{\partial a_b} \gtrless 0$ and $\frac{\partial y_s}{\partial a_b} \gtrless 0$: The impact on parents' investment in brother and sister schooling is ambiguous when parents' altruism toward brother increases.

6. $\frac{\partial y_b}{\partial p_b} \gtrless 0$ and $\frac{\partial y_s}{\partial p_b} > 0$: Parents invest more in sister's schooling when brother's productivity in household production is higher, ceteris paribus. The impact of investment in brother's own schooling is ambiguous when brother's productivity in household production increases.

Proof

Differentiating the FOCs from the optimization of brother and sister's schooling yields following SOCs

MB of y_b with respect to y_b:

Define A = $a_b V''(W_b) r_b^2 g_{y_b}^2 + a_b V'(W_b) r_b g_{y_b y_b}$

MB of y_b with respect to y_s

Let B = $a_b V''(W_b) r_b^2 g_{y_b} g_{y_s} + a_b V'(W_b) r_b g_{y_b y_s}$

MB of y_s with respect to y_s :

Let C = $a_s V''(W_s) r_s^2 f_{y_s}^2 + a_s V'(W_s) r_s f_{y_s y_s} + a_b V''(W_b) r_b g_{y_s}^2 + a_b V'(W_b) r_b g_{y_s y_s}$

1. $\frac{\partial y_b}{\partial W_p} = C(p_b + d_b)U'' - B(p_s + d_s)U''$ and

$\frac{\partial y_s}{\partial W_p} = A(p_s + d_s)U'' - B(p_b + d_b)U''$. Both expressions are positive given the assumptions outlined above..

2. $\frac{\partial y_s}{\partial r_s} = -Aa_s V'(W_s) f_{y_s} - (p_b + d_b)^2 a_s V'(W_s) f_{y_s} U''$ which is positive. Since parents respond to higher market return to sister human capital by investing more in the sister, this leaves less resources for the brother, and puts downward pressure on brother's schooling. The first term in

$\frac{\partial y_b}{\partial r_s} = (p_b + d_b)(p_s + d_s)(a_s V'(W_s) f_{y_s})U'' + Ba_s V'(W_s) f_{y_s}$ is negative if B is negligible. This highlights the usual competition effect between sister and brother schooling since they both draw from the same pool of household resources, an increase in one usually entails a reduction in the other. The expression also highlights a special case in which y_b may increase in response to an increase in r_s. If B>0 (the marginal benefit from y_s is increasing y_b) and this is sufficiently large to overcome the negative competition effect, it may be optimal for the parents to increase y_b as well since the investment of each unit y_b is now more productive with higher y_s .

3. Let $F = -a_b V'(W_b) g_{y_b}$ and $G = -a_b V'(W_b) g_{ys}$, then

$$\frac{\partial y_b}{\partial r_b} = CF + (p_s + d_s)^2 FU'' - (p_s + d_s)(p_b + d_b)GU'' - BG$$

The impact of an increase in younger human capital is ambiguous without further assumptions because both y_b and y_s can be increased to increase younger brother human capital. If I impose that y_b is more effective at creating younger brother human capital than y_s (i.e. $g_{y_b} > g_{y_s}$) as seems plausible, then we get the prediction that brother schooling increases in response to an increase in market return to brother human capital. $\frac{\partial y_s}{\partial r_b} = -BF + AG - p_b(1+d_b)p_s(1+d_s)FU'' + p_b(1+d_b)^2 GU''$ remains ambiguous because the resulting increase in brother schooling crowds out sister schooling but sister schooling is also productive in increasing brother human capital, the return to which has gone up.

4. Let $O = -V'(W_s)r_s f_{y_s}$, then

$$\frac{\partial y_s}{\partial a_s} = (p_b + d_b)^2 OU'' + AO \text{ and } \frac{\partial y_b}{\partial a_s} = -(p_b + d_b)(p_s + d_s)OU'' - BO$$

Parents increase sister's schooling when altruism towards sister is higher. This crowds out investment in schooling for brothers and leads to a reduction in y_b. If B>0, the strong complimentarity between y_b and y_s means that there is a positive pressure on y_b because the increased y_s means y_b is more productive.

5. Let $M = -V'(W_b)r_b g_{y_b}$ and $N = -V'(W_b)r_b g_{y_s}$, then

$$\frac{\partial y_b}{\partial a_b} = CM - (p_s + d_s)^2 MU'' - (p_b + d_b)(p_s + d_s)NU'' - BN$$

and $\frac{\partial y_s}{\partial a_b} = AN - (p_b + d_b)(p_s + d_s)MU'' + (p_b + d_b)^2 NU'' - BM$

The impact of an increase in altruism toward younger brother is ambiguous because both y_b and y_s can be increased to increase younger brother human capital. If I impose that y_b is more effective at creating younger brother human capital than y_s (i.e. $g_{y_b} > g_{y_s}$) as seems plausible, then we get the prediction that brother schooling increases if parents' altruism towards brother increases. The impact on sister schooling remains ambiguous because the increased brother schooling crowds out sister schooling but sister schooling is also productive in creating brother human capital.

6: $\frac{\partial y_s}{\partial p_s} = AU' - (1 - y_s)(p_b + d_b)BU'' + (1 - y_s)(p_s + d_s)AU'' + (p_b + d_b)^2 U'U''$ and

$$\frac{\partial y_b}{\partial p_s} = -BU' + (1-y_s)(p_b+d_b)CU'' - (1-y_s)(p_s+d_s)BU'' - (p_b+d_b)(p_s+d_s)U'U''$$

Higher household productivity for the sister has an ambiguous effect on her own schooling because there is a a negative substitution effect (because the sister is now more productive in the household) and a positive income effect (from her enhanced household productivity which increases income). Increasing the household productivity of the sister increases brother's schooling because for the brother the substitution effect and income effect both are positive. The The relationship is analogous for an increase in brother's household productivity with $\frac{\partial y_b}{\partial p_b} \gtrless 0$ and $\frac{\partial y_s}{\partial p_b} > 0$.

B Data Appendix

Construction of Sample of Interest

The identification of oldest sisters and younger siblings is complicated by the survey format and timing. Age reports contain significant amounts of measurement error in developing countries. For this reason, I choose to identify older and younger pairs of siblings based on complete fertility histories that were asked of mothers. As part of the fertility histories, mothers were asked to rank the birth order of all their children ever born. I use this history to determine who is the older and the younger individual in all sibling pairs. Since the fertility histories were only administered in round 3, any children who were not living in the household in round 3 do not get tagged as either the oldest sister or her younger sibling. I describe the procedure by which I determined the relationships for such individuals in detail in the appendix.

Instead of limiting my focus to the absolute oldest sister (in terms of birth order), I look at the oldest sister among the sisters still living in the household. In 73 percent of the households I use in my analyses, the oldest sister in the household is in fact the oldest daughter that was ever born to that household. The reason for defining oldest as oldest living in the household is two-fold. I have education information only for individuals who have lived in the household at some point during the panel. If the oldest sister in a household moved out of the household before data collection, I know of her existence from the fertility history but I lack information about her education. Secondly, it is not very interesting to ask about the impact of an oldest sister with whom the younger sibling would have spent very few years interacting since she moved out of the household a long time ago.

Age of oldest sister

The effective sample of interest includes 1211 households in which I have identified the oldest sister and at least one sibling younger than her aged 5-12 years old inclusive. I limit the sample to households in which the oldest sister is between 8 and 30 years old in round 1. Only 1 percent of the sisters have age greater than 30 in round 1, so dropping these observations is reasonable trimming of extreme outliers. On the lower bound of age, I drop sisters who are too young to have acquired any schooling.

Recall that the sample includes oldest sisters who are the oldest among the sisters still living in the household. In case the absolute oldest sister never lived in the household during the data collection period, I substitute with the oldest of the sisters that does appear in the data. Although the substitution with oldest of the sisters living in the household makes the sample selection less problematic, the sample still conditions on the sister not having moved out of the household before the panel starts. Girls in Pakistan move out of the household to live with their husband's family at the time of marriage. Marriage accounts for 99 percent of the girls' moves out of the household observed in my data. Since the length of a girl's stay in her household is dictated by her marital status, my sample based on older sisters who are still living in the house may be a non-random sample. If education improves the probability of getting married, the better quality, more educated oldest sisters are already married off by the time data collection starts, and dropped from my analysis because I don't know their education. In this case, my sample of oldest sisters is adversely selected, has lower education than the population and the positive estimates I obtain from my sample are likely lower bounds of the true treatment effect for the population[39]. I also conduct a robustness check by limiting to oldest sisters who are aged less than 20 years old in section 8.

The complete fertility histories which I use to determine birth order of siblings were administered in round 3 of the panel, and they also list the identifier, name, gender, age, whether the child still lives in the household, and reason for the child not living in the household. If the oldest sister no longer lives in the household in round 3, the survey data does not report her member ID – the within-household identification number – in the fertility history section. For this reason, some of the oldest sisters do not get flagged as such even though they may appear somewhere else in the panel data because they were living in the household in a different round. Due to this problem, I can only identify oldest sisters in 1060 of the 1646 households that report having at least one daughter born to them.

There are two types of oldest sisters with missing ID in the fertility history section: i) some of these have lived in the household recently enough to have been captured in the panel data at least in one round, and ii) other oldest sisters have moved out of the household or been dead for long enough such that they are never captured in the data. For i), I am interested in identifying these girls as the oldest sister because I have valid education information for them since they appear in the panel at some point. I matched these girls' reported names in the fertility histories to the female names listed in the household roster in all rounds. Since there are no uniform rules for the transliteration of Urdu names into English (the data is in English), I had to match these names manually on the basis of the phonetics. Using this procedure, I was able to identify another 129 oldest sisters.

In case ii), the oldest sister has either been married or dead for a sufficiently long time so that she never appears as living in the household during the duration of a four-year panel. For these households, I flag the next oldest daughter who does appear in the data at some point as the oldest sister. By thus flagging later-

[39] One could also imagine higher education competing with marriage for girls if, as in urban Pakistan, most girls discontinue their schooling and take on household and family responsibilities after they get married. Education and the incidence of marriage could have a negative relationship. Given that the median age of marriage is 20 while the average schooling for the 20-year olds is 5 years (achieved at roughly 13/14 years of age), it does not seem that marriage should compete with schooling for these rural girls in this way.

born daughters who are the oldest among the children still in the household, I identify the oldest sister in an additional 374 households.

Finally, not all households have valid fertility histories filled out. 1727 of 1807 households have a valid fertility history section filled out. For these households, I determine who is older and younger in sibling pairs based on the reported age of children of the household head. At the end of all these steps, I am able to flag 1630 households with the oldest sister. Out of the 1727 households that had valid fertility histories, 5 percent never had a daughter born to them. Having identified 1630 households with an oldest sister means I have captured 90 percent of my sample households. While the 10 percent of households I lose in my sample is bigger than the 5 percent we expect not to have any daughters born to them, this is very reasonable attrition considering that attrition can be caused due to several reasons including not having a daughter ever born to you, having daughters born to you but having them all die or move out of the household before the panel starts.

Turning to the other side, I also use the birth order from the fertility history to identify younger siblings. For 193 younger siblings who had moved out of the house in round 3 but did appear in the data at some point, I match them by name so they can be flagged as younger and used in the analyses. Even though I use birth order from the fertility history to determine who is older and younger among siblings, I find that the reported age for 14 percent of the sibling-year observations in my panel are equal to or greater than the age of the oldest sister. In order to ensure that I'm using cases where I can cleanly identify oldest sister and younger siblings, I drop these cases from my analyses.

Cleaning of education variables

I utilize the multiple reports of education in the panel dataset to conservatively clean egregious cases of detectable measurement error in the education variables. I describe the procedure I follow to clean after describing the questions relevant to education. The household respondent was asked for each household member whether the individual had never attended school, used to go to school, or is currently attending. If the response is “used to go” or “currently attending”, the follow-up question asked is “what is the highest grade completed by that individual”. If the response was “never attended”, the highest grade question is skipped.

I have three rounds of reports in my panel. Generally, I “correct” an outlier response if it disagrees with the majority response, while paying special attention to the relative timing. I make the following updates to get a clean variable for highest grade completed:

1) If the last report indicates you never went to school and have highest grade completed (hgc) missing/zero but the first two reports say either “used to go” or “currently attending”, I correct the last hgc and update it to the hgc from the second round.
2) If the middle report indicates you never went to school and has hgc equal to missing/zero but the other two reports say either “used to go” or “currently attending”, I correct the middle hgc and update it to the hgc from the first round.
3) If the first report indicates you used to go or shows you currently attending, but the last two reports both say never and hgc equal to 0/missing, I update the hgc report in first round to equal 0.
4) If the second report indicates you used to go or shows you currently attending but the first and last report both say never and hgc equal to 0/missing, I update the hgc report in the second round to equal 0.
5) If all three reports show completed enrollment for the individual and no fresh enrollment during panel, and the hgc report deviates in one round but is consistent across the remaining two rounds, I update the deviant hgc to equal the hgc from the other two rounds.

Selection of oldest sister

This study estimates the impact of the oldest sister’s education on younger sibling learning and education where the oldest sister is defined as the oldest of the sisters living in the household during the panel. One reason behind considering the oldest sister conditional on her living in the household rather than analyzing impact of the absolute oldest sister based on birth order was that the dataset only includes education information for those family members who are residing in the household at some point during the panel. Since girls move to the husband's house after marriage in Pakistan, if an (absolute) oldest sister got married before the data collection, I lack information on her education. Second, since I am estimating impact on outcomes for primary school age children, it is not very meaningful to estimate the impact of an oldest sister who may have married and moved away from the household a long time ago. Thus, for data availability and for theoretical reasons, it made sense to analyze the impact of the oldest of the sisters living in the household.

One may expect that the role model effects of the education of the absolute oldest sister may be more important than that of the oldest among the sisters living in the household. A higher age gap between the absolute oldest sister and the younger sibling may mean the oldest sister had a greater role in taking care of the younger sibling. On the other hand, it may imply less interaction overall than a younger sibling may have with another older sister who is closer to him/her in age, particularly because the absolute oldest sister gets married earlier. For these reasons, the impact of the absolute oldest sister may diverge from that of the oldest sister as defined in the paper so far i.e. the oldest of the sisters living in the household. In 73 percent of the cases, the oldest of the sisters living in the household is also the absolute oldest sister. Here I present results after limiting the sample to just the absolute oldest sisters as a robustness check.

Table A3 shows the IV results for the absolute oldest sister's years of completed schooling. The treatment effects are about the same for adding and counting, bigger for writing and schooling, and smaller for reading and enrollment than those found in the IV specification in table 10a. The coefficients for read, add, count, schooling and enrollment are statistically significant as before. It seems that there is no systematic variation across the two specifications so it is hard to discern whether the greater role model effects of the absolute oldest sister and any effects from the increased care-taking role she takes on for her younger siblings outweigh the effect of decreased interaction with younger siblings. It is important to also realize that limiting attention to the absolute oldest sisters entails a necessarily selected sample because I know the absolute oldest sister's education only if she is still living in the household and not married.

The next robustness check deals with selection due to marriage. All girls move out of their parents' household and into the husband's family after marriage. I only observe education of the oldest sister if she is still living in the household and is therefore still single. If education improves a girl's marriage prospects, my treatment impacts may be estimated from an adversely selected sample. If education competes with a girl's transition into marriage (which is unlikely given the low educational attainment of girls), the impact may be estimated from a positively selected sample. Defining the oldest sister as the oldest sister among the sisters residing in the household helps mitigate the selection to some extent because we can substitute for the oldest sisters that got married really early. As an additional check, I restrict the sample to oldest sisters less than 20 years old. Only 15.6 percent of the oldest daughters aged 15-19 years old were married in my sample. Table A4 shows the results from the sample of oldest sisters aged 19 and under. I find that there is no qualitative difference in the estimates after restricting the sample. The impact estimates for reading, writing, and adding are smaller while the estimates for counting, schooling and enrollment are bigger than those found using the more general sample.

C Measurement Error Appendix

Estimating lower bounds of the treatment effect with binary mismeasured explanatory variables

F&L describe a procedure that allows us to estimate lower bounds of the true treatment effect β by finding upper bounds on α_0 and α_1 under the following assumptions: i) these probabilities are assumed to be independent of X and ε, and ii) $Cov(D, D^*) > 0$ (if this is not the case, measurement error is so severe that (1-D) is a better measure of D* than D is). Then independence of X and the measurement error process yields that

$$\begin{aligned}\Pr(D=1 \mid X) &= (1-\alpha_1)\Pr(D^*=1 \mid X) + \alpha_0(1-\Pr(D^*=1 \mid X) \\ &= \alpha_0 + (1-\alpha_0-\alpha_1)\Pr(D^*=1 \mid X)\end{aligned}$$

This equation implies that $\alpha_0 \le \Pr(D=1 \mid X) \le 1-\alpha_1$ for all X. F&L propose that one can obtain the tightest possible bound for α_0 (α_1) by estimating $E(T \mid X \in S)$ over the subset of sample S having the lowest (highest) expected value of T. In order to get the lowest (highest) expected value of T, they propose estimating Pr(D=1|X) by regressing D on X and the instrument Z, and then calculating $E(T \mid X)$ over the observations with percentile rank of Pr(D=1|X) less (more) than q. The optimal choice of q is left as an open question for future research but the authors use q = 5 themselves. Since $\alpha_0 \le \Pr(D=1|X) \le 1-\alpha_1$ for all X, an incorrect functional form only affects the tightness of the bounds, not their validity.

GMM estimation which is consistent with non-classical measurement error

My model is $Y_i = \beta D_i^* + \delta X_i + \varepsilon_i$ for observation i, where Y_i, D_i^* and ε_i are scalars, and X_i is a 1 x k row vector. I am only interested in the parameter β, and consider the vector δ a nuisance parameter which must be estimated in order to control for the observables but which is not an object of interest itself. I use partial regression to condition on X_i in all estimation, but I suppress the covariate notation so as to conserve notation in the remainder of the discussion. One can estimate this model by purging the impact of observables X from all other variables using partial regression (Yule, 1907) as suggested by Black and Smith (2006). I regress all outcomes, the treatment variable and the two instruments on the vector X. I work with the residuals obtained from these regressions – called yulized residuals by Black and Smith – which will provide results that are numerically identical to estimating the model in one step.

The model for the outcome then is $Y_i = \beta D_i^* + \varepsilon_i$ and the model allowing for a general measurement error process is $D = \gamma D^* + \upsilon$ where it is assumed that $Cov(D^*, \upsilon) = 0$ as before. I assume the availability of two instruments, Z_1 and Z_2: $Z_1 = \alpha_1 D^* + \omega_1$ and $Z_2 = \alpha_2 D^* + \omega_2$ such that $Cov(D^*, \omega_1)$, $Cov(D^*, \omega_2)$,

$Cov(\omega_1, \upsilon)$, $Cov(\omega_2, \upsilon)$, and $Cov(\omega_1, \omega_2)$ are all equal to zero. The first instrument, Z_1, is distance to closest government girls' school which I have been using in this paper. The second instrument, Z_2, is an indicator for whether the closest government girls' school offers instruction in classes higher than primary[40].

[40] There are two types of government schools in my data: those that offer primary classes and those that offer instruction in primary classes and classes higher than primary. While the instrument suffices for the purpose of this robustness check, I do not use this instrument in my preferred identification strategy for two reasons. First, it is a weak

With this structure, the system is exactly identified because I obtain 6 population moment conditions and 6 parameters, including $\beta, \gamma, \alpha_1, \alpha_2, Var(D^*)$ and $Cov(D^*, \varepsilon)$. The population moment conditions are:

$$
\begin{aligned}
s_{Y,D} &= \beta\gamma \operatorname{var}(D^*) + \gamma Cov(D^*, \varepsilon) \\
s_{Y,Z_1} &= \beta\alpha_1 \operatorname{var}(D^*) + \alpha_1 Cov(D^*, \varepsilon) \\
s_{Y,Z_2} &= \beta\alpha_2 \operatorname{var}(D^*) + \alpha_2 Cov(D^*, \varepsilon) \\
s_{D,Z_1} &= \gamma\alpha_1 \operatorname{var}(D^*) \\
s_{X,D_2} &= \gamma\alpha_2 \operatorname{var}(D^*) \\
s_{Z_1,Z_2} &= \alpha_1\alpha_2 \operatorname{var}(D^*)
\end{aligned}
$$

instrument and second, with less than two government girls' schools per village, the dominant source of variation in this instrument is inter-village variation in the classes that girls' schools offer.

Figure 1

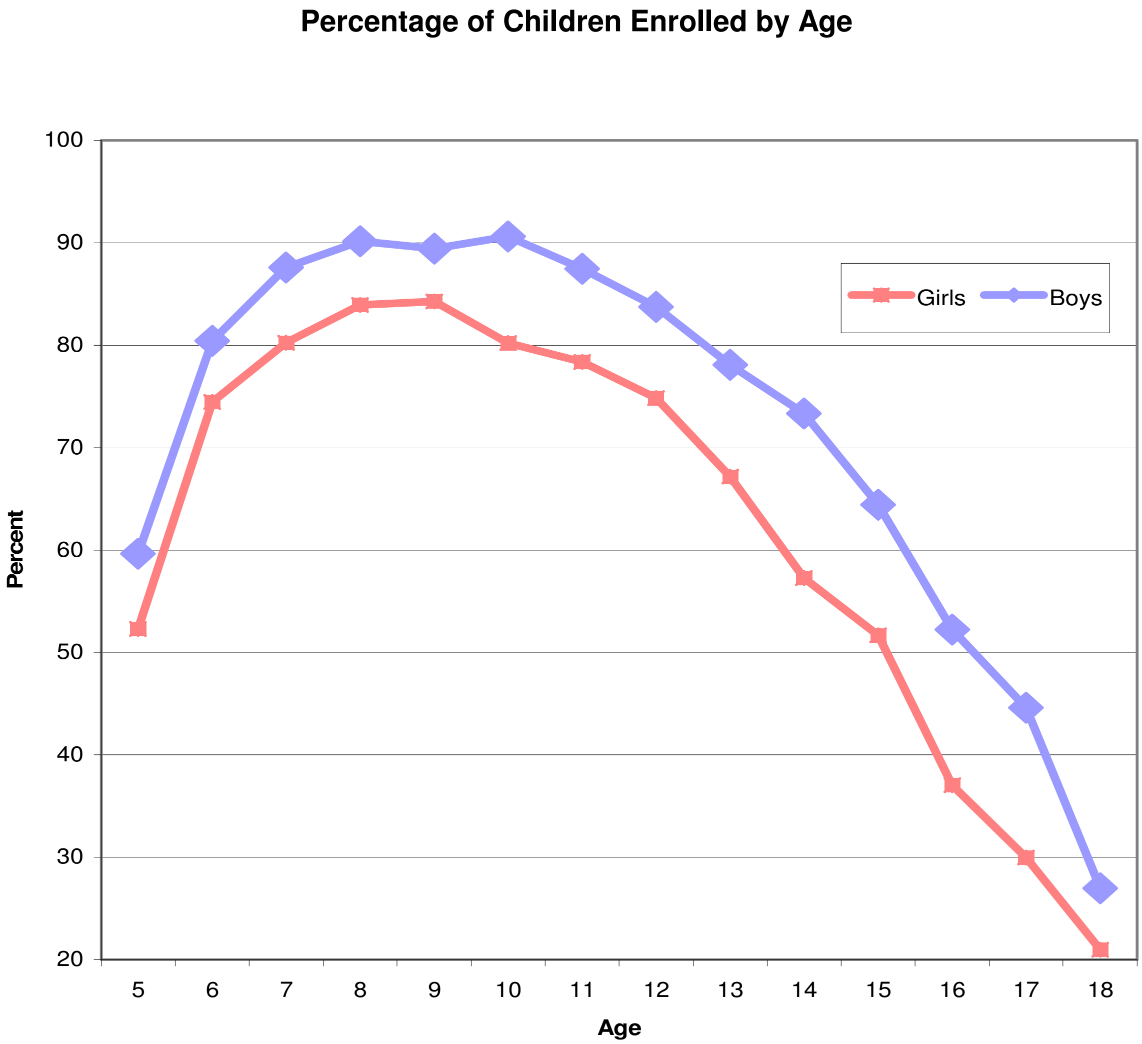
Percentage of Children Enrolled by Age
100
90
80
70
60
50
40
30
20
Percent
Girls
Boys
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Age

Figure 2

Distribution of Years of Schooling for Boys and Girls aged 16-20 years old

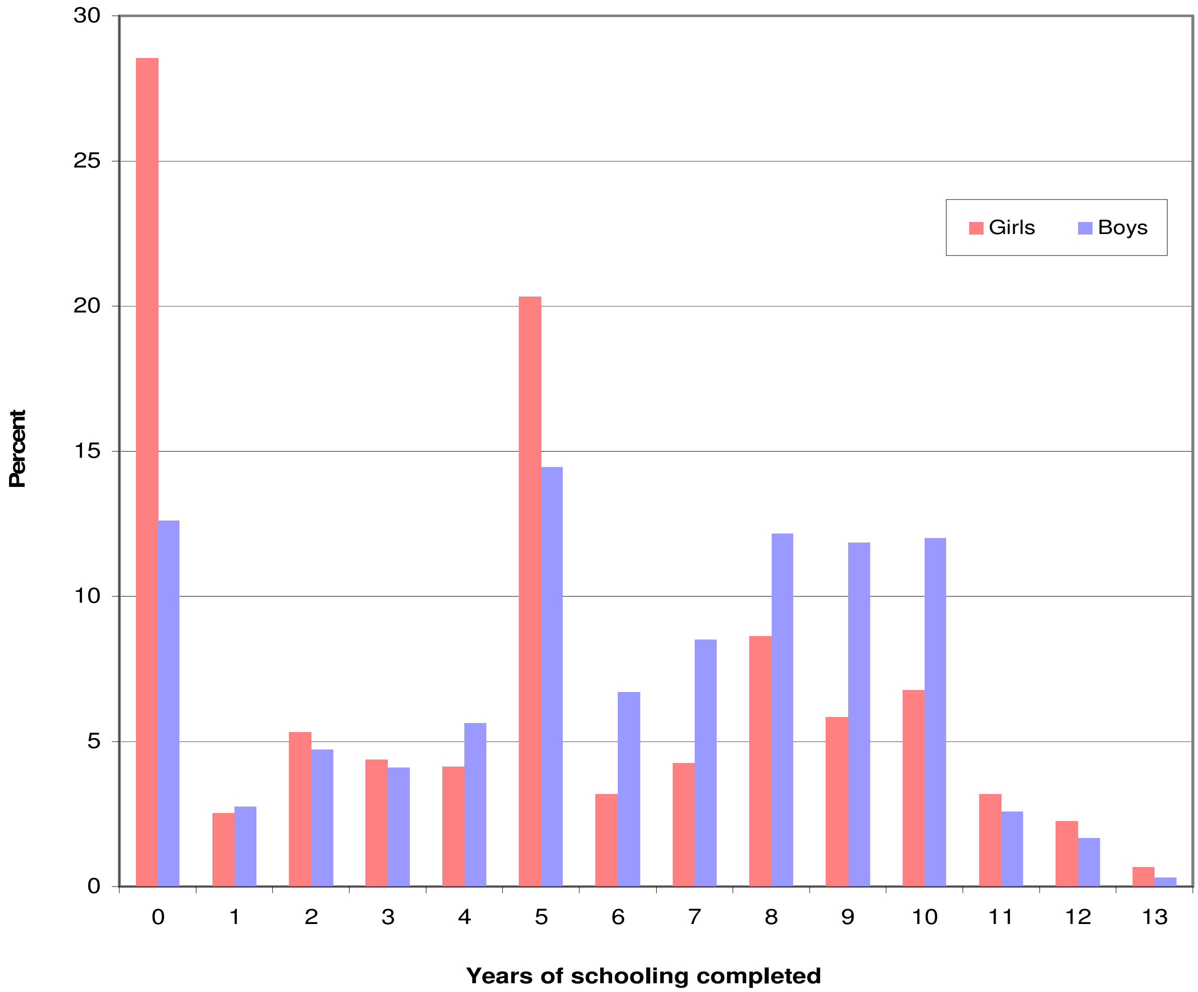

Figure 3

Literacy and numeracy capabilities of younger siblings aged 5-12 years old

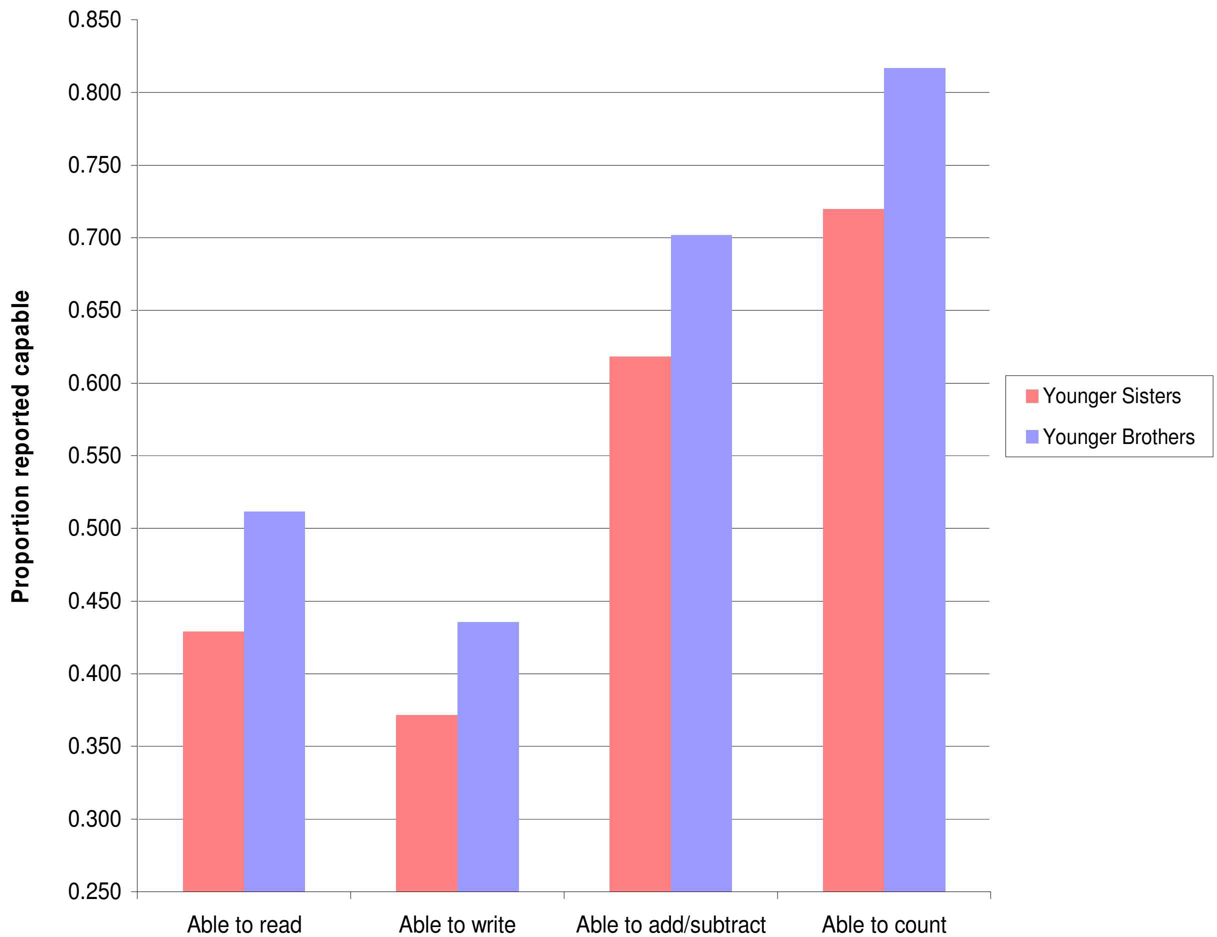

Figure 4
Proportion of children with reported literacy and numeracy capabilities by class

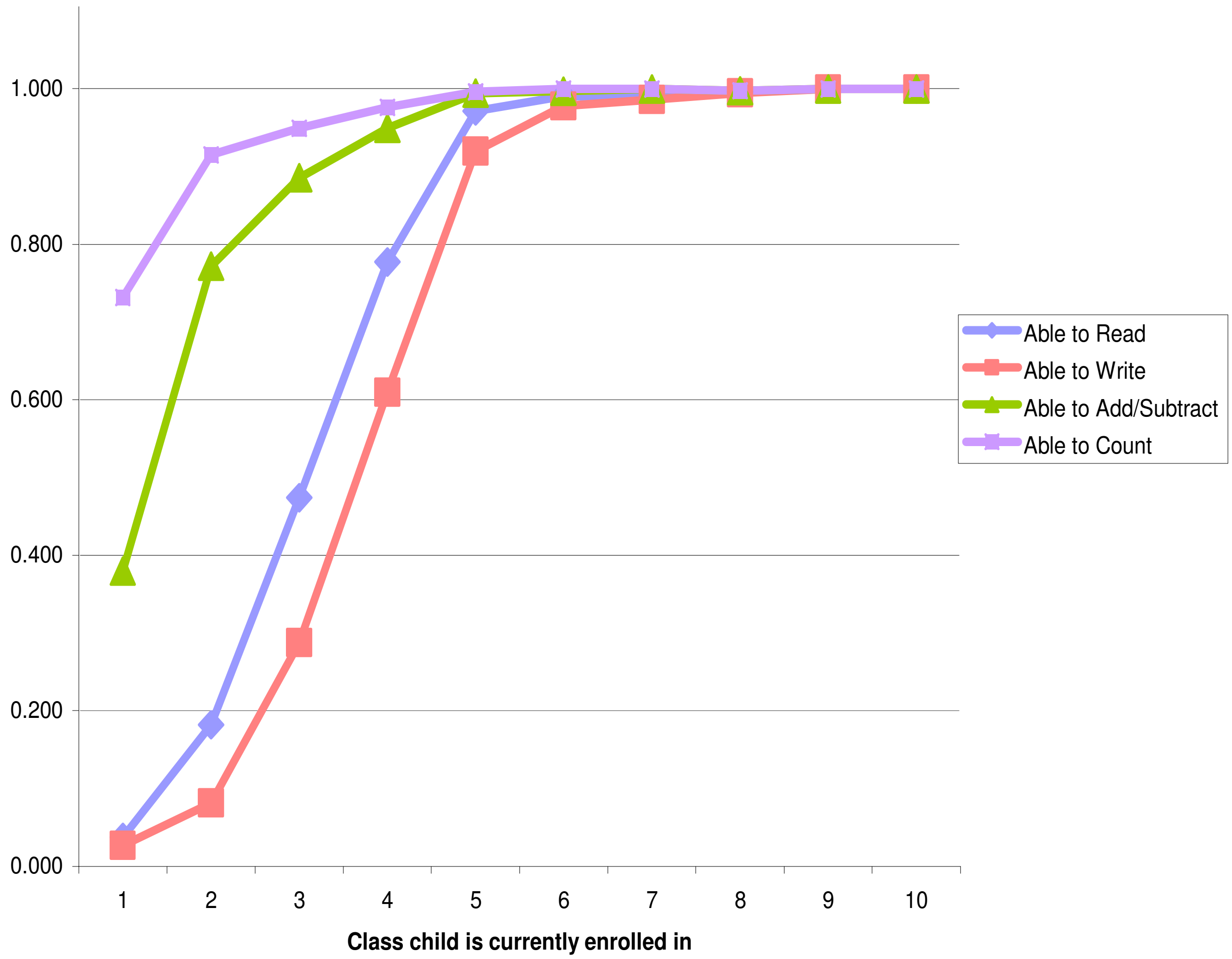

Table 1: Summary statistics of households in data 2003-2006

	Mean	Std Dev	N
Household size	7.854	2.634	1,211
# of children 0 to 18 yrs old	5.095	1.921	1,211
# of boys 0 to 18 yrs old	2.394	1.364	1,211
# of girls 0 to 18 yrs old	2.620	1.422	1,211
Mother is educated	0.237	0.425	1,202
Father is educated	0.616	0.486	1,087
Mother's Ed Yrs	1.457	2.768	1,202
Father's Ed Yrs	4.404	4.169	1,087
Own any land	0.478	0.460	1,210
Own house living in	0.937	0.212	1,210
Hse is pukka	0.508	0.375	1,210
Hse is kucha	0.108	0.236	1,210
Hse is kucha-pukka	0.384	0.336	1,210
Water source is a hand pump	0.241	0.270	1,211
Water source is tap water	0.141	0.208	1,211
Water source is a motor pump	0.191	0.268	1,211
Water source is external like pond, stream etc.	0.100	0.222	1,211
Food expenditure per capita (Rs/month)	1,852.459	3,852.013	1,210
District: Attock	0.336	0.473	1,211
District: Faisalabad	0.386	0.487	1,211
District: RYK	0.277	0.448	1,211

The universe is households with an oldest sister aged 8-30 in round 1 with non-missing education, and who has at least one younger sibling aged 5-12 years old inclusive. At the time of the survey, 1 USD = 60 Pakistani rupees on average.

Table 2a: OLS results - Years of completed schooling of oldest sister

	1	2	3	4	5	6
Read						
Oldest sister's years of schooling	0.032***	0.027***	0.026***	0.026***	0.027***	0.027***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	3,553	3,553	3,553	3,553	3,553	3,553
R-squared	0.360	0.367	0.372	0.374	0.375	0.387
Mean of dependent variable			0.410			
Write						
Oldest sister's years of schooling	0.025***	0.020***	0.019***	0.019***	0.019***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	3,542	3,542	3,542	3,542	3,542	3,542
R-squared	0.335	0.342	0.350	0.356	0.356	0.364
Mean of dependent variable			0.325			
Add						
Oldest sister's years of schooling	0.025***	0.020***	0.018***	0.018***	0.018***	0.018***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	3,523	3,523	3,523	3,523	3,523	3,523
R-squared	0.380	0.386	0.393	0.394	0.394	0.395
Mean of dependent variable			0.642			
Count						
Oldest sister's years of schooling	0.018***	0.013***	0.012***	0.012***	0.012***	0.013***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	3,561	3,561	3,561	3,561	3,561	3,561
R-squared	0.289	0.296	0.303	0.304	0.304	0.313
Mean of dependent variable			0.780			
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Parents' education controls	No	Yes	Yes	Yes	Yes	Yes
Wealth controls	No	No	Yes	Yes	Yes	Yes
District F.E.	No	No	No	Yes	Yes	Yes
Year F.E.	No	No	No	No	Yes	Yes
District*Year F.E.	No	No	No	No	No	Yes

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level. Household controls include age of oldest daughter, age of younger brother, number of daughters & sons, family size, and language. Parents' education controls include indicators for mother and father having any education, controls for years of education and indicators if missing. Wealth controls include whether family owns the house they live in, type of house, type of water source, whether they own land, food expenditure per capita

Table 2b: OLS results - Years of completed schooling of oldest sister

	1	2	3	4	5	6
Education (yrs)						
Oldest sister's years of schooling	0.179***	0.160***	0.156***	0.154***	0.152***	0.151***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Number of observations	5,333	5,333	5,333	5,333	5,333	5,333
R-squared	0.595	0.599	0.601	0.606	0.606	0.607
Mean of dependent variable	2.942					
Enrollment Status						
Oldest sister's years of schooling	0.025***	0.019***	0.019***	0.019***	0.019***	0.019***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Number of observations	5,349	5,349	5,349	5,349	5,349	5,349
R-squared	0.201	0.211	0.216	0.223	0.223	0.224
Mean of dependent variable	0.781					
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Parents' education controls	No	Yes	Yes	Yes	Yes	Yes
Wealth controls	No	No	Yes	Yes	Yes	Yes
District F.E.	No	No	No	Yes	Yes	Yes
Year F.E.	No	No	No	No	Yes	Yes
District*Year F.E.	No	No	No	No	No	Yes

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level. Household controls include age of oldest daughter, age of younger brother, number of daughters & sons, family size, and language. Parents' education controls include indicators for mother and father having any education, controls for years of education and indicators if missing. Wealth controls include whether family owns the house they live in, type of house, type of water source, whether they own land, food expenditure per capita

Table 3: OLS results - all treatment specifications

	Read	Write	Add	Count	Schooling	Enrollment
Oldest sister's years of schooling	0.027***	0.019***	0.018***	0.013***	0.151***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)	(0.002)
Number of observations	3,553	3,542	3,523	3,561	5,333	5,349
Indicator oldest sister has any educ	0.106***	0.057***	0.085***	0.064***	0.664***	0.126***
	(0.023)	(0.022)	(0.022)	(0.021)	(0.098)	(0.020)
Number of observations	3,553	3,542	3,523	3,561	5,333	5,349
Indicator oldest sister educ >=5 yrs	0.132***	0.085***	0.097***	0.072***	0.359***	0.077***
	(0.019)	(0.018)	(0.018)	(0.017)	(0.058)	(0.014)
Number of observations	3,553	3,542	3,523	3,561	5,333	5,349
Mean of dependent variable	0.410	0.325	0.642	0.780	2.942	0.781

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
The dependent variables read, write, add, count are indicator variables for whether younger brothers are reported as capable of reading, writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the brother is currently enrolled. The regression results reported are from the fullest specification which controls for household variables, parents' education, wealth controls as well as district*year fixed effects.

Table 4: OLS results by gender of younger sibling

	Read	Write	Add	Count	Schooling	Enrollment
Oldest sister's years of schooling	0.027***	0.019***	0.018***	0.013***	0.150***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)	(0.002)
Oldest sister's years of schooling*	0.008*	0.012***	0.005	0.011***	0.092***	0.019***
indicator younger sister	(0.004)	(0.004)	(0.004)	(0.003)	(0.025)	(0.003)

	Read	Write	Add	Count	Schooling	Enrollment
Indicator oldest sister has any education	0.106***	0.057***	0.085***	0.064***	0.649***	0.126***
	(0.023)	(0.022)	(0.022)	(0.021)	(0.100)	(0.020)
Indicator oldest sister has any education*	0.082***	0.081***	0.085***	0.122***	0.292**	0.165***
indicator younger sister	(0.029)	(0.027)	(0.029)	(0.030)	(0.146)	(0.029)

	Read	Write	Add	Count	Schooling	Enrollment
Indicator oldest sister education >= 5 yrs	0.132***	0.086***	0.097***	0.071***	0.332***	0.077***
	(0.019)	(0.018)	(0.018)	(0.016)	(0.057)	(0.014)
Indicator oldest sister education >= 5 yrs*	0.059**	0.080***	0.035	0.049**	0.466***	0.076***
indicator younger sister	(0.025)	(0.025)	(0.024)	(0.022)	(0.084)	(0.020)
Number of observations	6,765	6,744	6,733	6,775	10,152	10,179

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
The main effect gives treatment effects for younger brothers while the interaction term indicates the amount by which the effect for younger sisters differs relative to younger brothers. The dependent variables read, write, add, count are indicator variables for whether younger siblings are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The regressionresults reported are from the fullest specification which controls for household variables, parents' education, wealth controls as well as district*year fixed effects.

Table 5: OLS results for impact on older brothers

	Read	Write	Add	Count	Schooling	Enrollment
Oldest sister's years of schooling	0.028***	0.026***	0.012***	0.009***	0.168***	0.021***
	(0.006)	(0.007)	(0.003)	(0.003)	(0.039)	(0.008)
Number of observations	1,565	1,562	1,566	1,566	1,907	988
R^2	0.181	0.195	0.080	0.063	0.348	0.350
Indicator oldest sister education >=5 yrs	0.062***	0.042	0.048***	0.031**	0.436***	0.009
	(0.023)	(0.029)	(0.015)	(0.015)	(0.115)	(0.031)
Number of observations	1,565	1,562	1,566	1,566	1,907	988
R^2	0.155	0.172	0.070	0.054	0.321	0.342
Indicator oldest sister has any education	0.194***	0.231***	0.099***	0.071***	0.664***	0.098**
	(0.038)	(0.039)	(0.024)	(0.020)	(0.153)	(0.039)
Number of observations	1,565	1,562	1,566	1,566	1,907	988
R^2	0.196	0.217	0.098	0.074	0.333	0.351
Mean of dependent variable	0.789	0.751	0.949	0.965	5.786	0.633

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
The dependent variables read, write, add, count are indicator variables for whether older brothers are reported as capable of reading, writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The regression results reported are from the fullest specification which controls for household variables, parents' education, wealth controls as well as district*year fixed effects.

Table 6: Summary statistics on school characteristics

A. Household-level characteristics	Mean	Std Dev	Min	Max	N
Distance to closest government girls' school	0.598	0.676	0	5.917	1,178
Distance to closest government boys' school	0.607	0.635	0	5.573	1,190
Distance to closest private school	0.573	0.711	0	5.644	1,190
Distance to village center	0.625	1.004	0	14.228	1,210
B. Village-level characteristics					
Number of government girls' schools in village	1.786	1.166	0	6	112
Number of government boys' schools in village	2.089	1.504	1	8	112
Number of private schools in village	3.134	2.338	0	13	112

Panel A provides summary statistics on household-level variables. Distance to closest government girls' school is missing for 12 more households than distance to closest government boys' school is missing for. This is due to the two villages that do not have a government girls' school. Panel B provides summary statistics for the 112 villages in the dataset.

Table 7: IV results from impact of oldest sister's education on older brother

	Read	**Write**	**Add**	**Count**	**Schooling**	**Enrollment**
Oldest sister's years of schooling	-0.005	-0.024	-0.033	-0.020	0.163	-0.087
	(0.062)	(0.070)	(0.037)	(0.030)	(0.324)	(0.074)
First stage results						
Distance to closest girls' school	-0.339***	-0.337***	-0.356***	-0.356***	-0.333***	-0.389***
	(0.119)	(0.119)	(0.120)	(0.120)	(0.107)	(0.112)
F statistic	8.123	8.066	8.821	8.821	9.734	12.110
Number of observations	1,501	1,498	1,502	1,502	1,819	945
Indicator oldest sister education >=5 yrs	-0.038	-0.179	-0.209	-0.134	-0.264	-1.260
	(0.350)	(0.372)	(0.218)	(0.174)	(2.113)	(1.280)
First stage results						
Distance to closest girls' school	-0.055***	-0.056***	-0.056***	-0.056***	-0.057***	-0.031
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.023)
F statistic	7.618	7.728	7.784	7.840	8.237	1.716
Number of observations	1,501	1,498	1,502	1,502	1,819	945
Mean of dependent variable	0.787	0.748	0.947	0.965	5.745	0.624

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
This analysis is limited to the sample of older brothers only.
The dependent variables read, write, add, count are indicator variables for whether younger siblings are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.

Table 8: Distance penalty for girls and boys' schooling

Distance penalty for girls						
	Years of schooling		**Indicator any education**		**Indicator education >= 5 years**	
Distance to closest girls' school (km)	-0.187***	-0.115**	-0.046***	-0.023*	-0.034***	-0.027***
	(0.050)	(0.054)	(0.011)	(0.012)	(0.007)	(0.009)
Distance to center		-0.256***		-0.052***		-0.013
		(0.063)		(0.012)		(0.008)
Number of observations	7,869	7,869	7,869	7,869	7,869	7,869
R^2	0.431	0.433	0.312	0.316	0.347	0.347

Distance penalty for boys						
	Years of schooling		**Indicator any education**		**Indicator education >= 5 years**	
Distance to closest boys' school (km)	-0.164*	-0.074	-0.045***	-0.035***	-0.015	-0.015
	(0.091)	(0.098)	(0.012)	(0.012)	(0.012)	(0.012)
Distance to center		-0.188***		-0.020**		-0.016**
		(0.057)		(0.009)		(0.008)
Number of observations	7,292	7,292	7,292	7,292	7,292	7,292
R^2	0.545	0.547	0.314	0.315	0.414	0.414

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
This table shows the distance penalty for schooling for girls and boys aged 5-18 years old. The distance penalty shows how an increase in distance to the closest girls/boys' school by an additional km affects the schooling of girls/boys, with and without controlling for the distance from the household to the village center. The regression results reported also control for household variables, parents' education, wealth controls as well as district*year effects.

Table 9: Distance penalty for girls and boys' schooling by age

	Distance penalty for girls			
	Years of schooling		**Indicator for enrollment**	
	Age < 13	**Age>=13**	**Age<13**	**Age>=13**
Distance to closest girls' school (km)	-0.058	-0.239***	-0.036**	-0.045***
	(0.045)	(0.078)	(0.018)	(0.017)
Distance to center	-0.131***	-0.420***	-0.031*	-0.046***
	(0.042)	(0.104)	(0.017)	(0.016)
Number of observations	4,329	3,540	4,370	3,436
R^2	0.513	0.303	0.202	0.290

	Distance penalty for boys			
	Years of schooling		**Indicator for enrollment**	
	Age < 13	**Age>=13**	**Age<13**	**Age>=13**
Distance to closest boys' school (km)	-0.079	-0.136	-0.015	-0.002
	(0.051)	(0.161)	(0.013)	(0.020)
Distance to center	-0.114**	-0.263**	-0.027**	-0.003
	(0.045)	(0.102)	(0.012)	(0.016)
Number of observations	4,299	2,993	4,329	2,958
R^2	0.543	0.218	0.128	0.235

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
This table shows how the distance penalty for education and enrollment varies by age. The distance penalty shows how an increase in distance to the closest girls/boys' school by an additional km affects the schooling and enrollment of girls/boys respectively, while controlling for the distance from the household to village center. The results are shown for boys/girls younger than 13 and aged 13 and older separately. The regression results reported also control for household variables, parents' education, wealth controls as well as district*year effects.

Table 10a: IV results for younger brothers - Oldest sister's years of schooling

	Read	Write	Add	Count	Schooling	Enrollment
Second stage IV results						
Oldest sister's years of schooling	0.077**	0.039	0.076**	0.055*	0.420***	0.075***
	(0.034)	(0.030)	(0.032)	(0.032)	(0.125)	(0.028)
First stage results						
Distance to closest girls' school	-0.391***	-0.406***	-0.400***	-0.386***	-0.398***	-0.405***
	(0.083)	(0.083)	(0.083)	(0.083)	(0.074)	(0.074)
F statistic	22.373	23.912	23.136	21.902	29.052	30.030
Distance to closest boys' school	0.002	0.003	0.003	0.002	-0.003	-0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Mean of dependent variable	0.408	0.323	0.642	0.779	2.935	0.780
Number of observations	3,413	3,405	3,386	3,422	5,100	5,115
R^2	0.337	0.353	0.331	0.270	0.547	0.162

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
The dependent variables read, write, add, count are indicator variables for whether younger brothers are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.

Table 10b: IV results for younger brothers - Indicator oldest sister has completed any schooling

	Read	Write	Add	Count	Schooling	Enrollment
Second stage IV results						
Indicator oldest sister has any	0.463*	0.282	0.434**	0.326*	3.125***	0.465**
education	(0.238)	(0.203)	(0.203)	(0.193)	(0.899)	(0.182)
First stage results						
Distance to closest girls' school	-0.061***	-0.065***	-0.066***	-0.064***	-0.059***	-0.064***
	(0.011)	(0.012)	(0.012)	(0.012)	(0.010)	(0.010)
F statistic	28.837	32.036	32.376	30.803	38.069	42.120
Distance to closest boys' school	-0.001	-0.001	0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of dependent variable	0.408	0.323	0.642	0.779	2.935	0.780
Number of observations	3,413	3,405	3,386	3,422	5,100	5,115
R^2	0.318	0.327	0.338	0.274	0.507	0.182

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
The dependent variables read, write, add, count are indicator variables for whether younger brothers are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.

Table 10c: IV results for younger brothers - Indicator oldest sister has completed at least primary schooling

	Read	Write	Add	Count	Schooling	Enrollment
Second stage IV results						
Indicator oldest sister education	0.508**	0.250	0.500**	0.384**	2.799***	0.563***
>= 5 yrs	(0.220)	(0.196)	(0.213)	(0.195)	(0.920)	(0.200)
First stage results						
Distance to closest girls' school	-0.060***	-0.062***	-0.061***	-0.060***	-0.056***	-0.056***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.011)	(0.011)
F statistic	20.160	21.344	20.612	19.714	23.814	24.305
Distance to closest boys' school	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of dependent variable	0.408	0.323	0.642	0.779	2.935	0.78
Number of observations	3,413	3,405	3,386	3,422	5,100	5,115
R^2	0.316	0.340	0.316	0.256	0.515	0.131

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
The dependent variables read, write, add, count are indicator variables for whether younger brothers are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.

Table 11a: Heterogeneity in first stage impact of instrument by maternal education

Mother uneducated						
	Read	**Write**	**Add**	**Count**	**Schooling**	**Enrollment**
Oldest sister's years of schooling						
First stage results						
Distance to closest girls school (km)	-0.340***	-0.368***	-0.358***	-0.343***	-0.336***	-0.363***
	(0.094)	(0.096)	(0.096)	(0.095)	(0.080)	(0.082)
F statistic	13.25	14.746	13.913	13.032	17.808	19.536
Observations	2,352	2,349	2,334	2,359	3,574	3,582

Mother educated						
	Read	**Write**	**Add**	**Count**	**Schooling**	**Enrollment**
Oldest sister's years of schooling						
First stage results						
Distance to closest girls school (km)	-0.270	-0.265	-0.247	-0.270	-0.122	-0.181
	(0.268)	(0.268)	(0.268)	(0.268)	(0.236)	(0.230)
F statistic	1.020	0.980	0.846	1.020	0.270	0.624
Observations	630	629	627	630	888	890

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.

This table shows the heterogeneity in the first stage impact of the instrument of distance to closest girls school on oldest sister's years of schooling by maternal education. I show the first stage results for the instrument separately for the subsample of households with uneducated mothers, and the subsample of households with educated mothers. The coefficient on distance to closest girls school indicates how an increase in distance to closest girls school of 1 km impacts oldest sister's years of schooling completed. The top panel shows that the instrument has a strong, highly statistically significant impact on oldest sister's schooling in households where the mother is uneducated, and that the distance penalty is considerably attenuated and lacks statistical significance for households with educated mothers. The first stage results are shown for each of the 6 outcomes considered and are from regressions using the fullest specifications which control for household variables, father's education, wealth controls as well as district*year fixed effects.

Table 11b: Heterogeneity in treatment impact by maternal education - OLS estimates

Mother Uneducated	Read	Write	Add	Count	Schooling	Enrollment
Oldest sister's years of schooling	0.028***	0.020***	0.023***	0.017***	0.164***	0.022***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.016)	(0.003)
Mean of dependent variable	0.378	0.297	0.618	0.761	2.828	0.749
Observations	2,352	2,349	2,334	2,359	3,574	3,582

Mother Educated	Read	Write	Add	Count	Schooling	Enrollment
Oldest sister's years of schooling	0.016**	0.017**	0.003	0.003	0.112***	0.011**
	(0.007)	(0.008)	(0.006)	(0.006)	(0.032)	(0.005)
Mean of dependent variable	0.478	0.391	0.708	0.835	3.324	0.894
Observations	630	629	627	630	888	890

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
This table shows the OLS results for treatment impact of oldest sister's schooling estimated separately for households with uneducated mothers and households with educated mothers. The term "educated" refers to an indicator for whether the mother has had any formal schooling. The results show that oldest sister's schooling has larger treatment impacts on younger brother learning and education in households where the mother is uneducated relative to households where the mother has some education. The regression results use the fullest specification which controls for household variables, wealth controls as well as district*year fixed effects.

Table 12: OLS results - Impact of Mother's and Father's Education

	Indicator variables for any education of mother and father					
	Read	**Write**	**Add**	**Count**	**Schooling**	**Enrollment**
Indicator mother has any education	0.072***	0.069***	0.073***	0.053***	0.423***	0.071***
	(0.021)	(0.022)	(0.019)	(0.017)	(0.091)	(0.017)
Indicator father has any education	0.074***	0.054***	0.037*	0.060***	0.563***	0.106***
	(0.021)	(0.019)	(0.019)	(0.017)	(0.099)	(0.019)
Number of observations	3,553	3,542	3,523	3,561	5,333	5,349
Mean of dependent variable	0.410	0.325	0.642	0.780	2.942	0.781
	Years of schooling of mother and father					
	Read	**Write**	**Add**	**Count**	**Schooling**	**Enrollment**
Mother's years of schooling	0.012***	0.011***	0.011***	0.007**	0.053***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.015)	(0.003)
Father's years of schooling	0.010***	0.010***	0.008***	0.010***	0.082***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.011)	(0.002)
Number of observations	3,553	3,542	3,523	3,561	5,333	5,349
	Mean	**Std Dev**	**N**			
Indicator mother is educated	0.237	0.425	1,202			
Indicator father is educated	0.616	0.486	1,087			
Mother's education (years)	1.457	2.768	1,202			
Father's education (years)	4.404	4.169	1,087			

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
This table shows the effect of mother's and father's education on learning and education outcomes for the sample of younger brothers used to analyze the impact of oldest sister's education. The sample thus consists of boys aged 5-12 for literacy and numeracy outcomes, and aged 5-18 for schooling and enrollment. The dependent variables of read, write, add and count refer to indicators for whether the boy can read, write, add/subtract or count. The dependent variable of schooling is years of completed schooling and enrollment is an indicator for current enrollment status. The regression results use the fullest specification which controls for household variables, wealth controls as well as district*year fixed effects.

Table 13: Estimates and bounds for treatment impacts and measurement error parameters

Indicator oldest sister has completed any schooling						
	α_0	α_1			α_0 95% CI	α_1 95% CI
Upper bound estimates		0.082	0.016		[0, .116]	[0, 0.031]
Adjustment factor ($1-\alpha_0-\alpha_1$)		0.902			0.853[a]	

	Read	Write	Add	Count	Schooling	Enrollment
IV estimate, β	0.463*	0.282	0.434**	0.326*	3.125***	0.465**
	(0.238)	(0.203)	(0.203)	(0.193)	(0.899)	(0.182)
Adjusted IV estimate, $\beta*(1-\alpha_0-\alpha_1)$	0.418	0.254	0.391	0.294	2.819	0.419
Adjusted IV estimate, $\beta*(1-\alpha_0-\alpha_1)$[a]	0.395	0.241	0.370	0.278	2.666	0.397

Indicator oldest sister has completed at least primary schooling						
	α_0	α_1			α_0 95% CI	α_1 95% CI
Upper bound estimates		0.008	0.059		[0, 0.019]	[0, .088]
Adjustment factor ($1-\alpha_0-\alpha_1$)		0.933			0.894[a]	

	Read	Write	Add	Count	Schooling	Enrollment
IV estimate, β	0.508**	0.250	0.500**	0.384**	2.799***	0.563***
	(0.220)	(0.196)	(0.213)	(0.195)	(0.920)	(0.200)
Adjusted IV estimate, $\beta*(1-\alpha_0-\alpha_1)$	0.474	0.233	0.467	0.358	2.612	0.525
Adjusted IV estimate, $\beta*(1-\alpha_0-\alpha_1)$[a]	0.454	0.223	0.447	0.343	2.501	0.503

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level.
[a] Lower bound for adjustment factor using the upper bounds of α_0 and α_1 estimates from the 95% confidence intervals α_0 is the probability that the indicator for treatment status equals zero but the reported treatment status equals 1, and α_1 is the probability that the indicator for treatment status equals one but the reported treatment status equals 0.
This table shows the original IV treatment estimates for oldest sister having any education and oldest sister having at least primray schooling, the upper bounds on the prevalence of measurement error in these two binary treatment variables, as well as the lower bounds on the IV treatment effects after incorporating the upper bounds on the measurement error. I also show the 95% confidence interval for the estimated measurement error probabilities, and calculate the lower bounds of the IV treatment effects using the upper bound from the 95% confidence interval of the estimated measurement error prevalence.

Table 14: Robustness check: GMM estimates for impact of oldest sister's years of schooling

	Read	Write	Add	Count	Schooling	Enrollment
OLS estimate	0.027***	0.019***	0.018***	0.013***	0.150***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)	(0.002)
IV estimate	0.077**	0.039	0.076**	0.055*	0.420***	0.075***
	(0.034)	(0.030)	(0.032)	(0.032)	(0.125)	(0.028)
GMM estimate	0.034***	0.016*	0.042***	0.020***	0.293***	0.040**
	(0.009)	(0.008)	(0.012)	(0.007)	(0.039)	(0.019)

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at household level. The table shows GMM estimates for the impact of oldest sister's years of schooling on younger brother outcomes. Measurement error in the oldest sister's years of schooling variable is likely mean-reverting because it is discrete and has a limited range. Mean-reverting measurement error biases the IV estimates of the impact of the mismeasured variable upwards. GMM estimation can provide consistent estimates in the presence of measurement error. Here, I use two instruments in a GMM framework to yield a system of equations that is exactly identified. The two instruments are distance to closest girls' school and an indicator for whether the closest government girls' school offers classes higher than primary. The table also lists the OLS, IV and GMM estimates for younger brothers for comparison purposes.

Appendix Table A1: Indicator mother educated and reported child literacy and numeracy

	Read	Read	Read	Read	Write	Write	Write	Write
Indicator mother educated	0.004	0.006	0.011	0.012	0.004	0.006	0.011	0.011
	(0.016)	(0.016)	(0.016)	(0.016)	(0.020)	(0.020)	(0.019)	(0.020)
English test score	0.074***	0.078***			0.100***	0.105***		
	(0.007)	(0.008)			(0.009)	(0.011)		
Urdu test score			0.077***	0.080***			0.118***	0.120***
			(0.007)	(0.008)			(0.009)	(0.010)
Interaction English test score		-0.017				-0.015		
		(0.016)				(0.021)		
Interaction Urdu test score				-0.012				-0.007
				(0.016)				(0.020)
Number of observations	2,360	2,360	2,360	2,360	2,341	2,341	2,341	2,341

	Count	Count	Add	Add
Indicator mother educated	0.001	0.001	0.001	0.000
	(0.005)	(0.005)	(0.008)	(0.008)
Math test score	0.013***	0.011***	0.021***	0.019***
	(0.002)	(0.003)	(0.004)	(0.004)
Interaction Math test score		0.006		0.008
		(0.006)		(0.009)
Number of observations	2,369	2,369	2,348	2,348

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
This table explores whether there are any systematic differences in literacy and numeracy capability reporting across uneducated and educated mothers. The concern is that mother's education may play a role in how she reports her child's ability to read/write/add/count, holding fixed the child's true capabilities. I regress the indicators for reading, writing, counting and adding on an indicator for mother's education to see if they are significantly correlated after controlling for child test scores. The idea is that test scores are an objective measure of the child's accumulated learning and if mother's education significantly predicts child's reported abilities even after controlling for test scores, educated mothers answer the child ability question significantly differently than uneducated mothers do. When considering mother report of child reading and writing abilities, I control for English and Urdu (Pakistani vernacular) test scores separately while for adding and counting abilities, I control for Math test score. I interact the indicator for mother's education with test score to see if there is a gradient to mother's report of child ability.

Appendix Table A2: Mother's education in years and reported child literacy and numeracy

	Read	Read	Read	Read	Write	Write	Write	Write
Mother's schooling (years)	0.002	0.003	0.004	0.004*	0.001	0.001	0.003	0.003
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
English test score	0.073***	0.077***			0.100***	0.098***		
	(0.007)	(0.008)			(0.009)	(0.010)		
Urdu test score			0.076***	0.081***			0.118***	0.117***
			(0.007)	(0.008)			(0.009)	(0.010)
Interaction English test score		-0.003				0.001		
		(0.003)				(0.003)		
Interaction Urdu test score				-0.003				0.001
				(0.003)				(0.003)
Number of observations	2,360	2,360	2,360	2,360	2,341	2,341	2,341	2,341

	Count	Count	Add	Add
Mother's schooling (years)	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Math test score	0.013***	0.013***	0.021***	0.022***
	(0.002)	(0.003)	(0.004)	(0.004)
Interaction Math test score		-0.001		-0.001
		(0.001)		(0.001)
Number of observations	2,369	2,369	2,348	2,348

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
This table explores whether there are any systematic differences in literacy and numeracy capability reporting across mothers with different years of schooling. The concern is that mother's education may play a role in how she reports her child's ability to read/write/add/count, holding fixed the child's true capabilities. I regress the indicators for reading, writing, counting and adding on an indicator for mother's education to see if they are significantly correlated after controlling for child test scores. The idea is that test scores are an objective measure of the child's accumulated learning and if mother's education significantly predicts child's reported abilities even after controlling for test scores, educated mothers answer the child ability question significantly differently than uneducated mothers do. When considering mother report of child reading and writing abilities, I control for English and Urdu (Pakistani vernacular) test scores separately while for adding and counting abilities, I control for Math test score. I interact the indicator for mother's education with test score to see if there is a gradient to mother's report of child ability.

Appendix Table A3: IV results for younger brothers - Absolute oldest sister

	Read	Write	Add	Count	Schooling	Enrollment
Second stage IV results						
Oldest sister's years of schooling	0.065**	0.054	0.072**	0.053*	0.468***	0.063**
	(0.029)	(0.033)	(0.029)	(0.029)	(0.136)	(0.030)
First stage results						
Distance to closest girls' school	-0.528***	-0.496***	-0.523***	-0.508***	-0.432***	-0.440***
	(0.100)	(0.100)	(0.101)	(0.100)	(0.089)	(0.089)
F statistic	27.563	24.602	27.040	25.705	23.523	24.305
Number of observations	2,685	2,677	2,667	2,691	3,835	3,845
R^2	0.386	0.358	0.361	0.286	0.581	0.184

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
This analysis limits the sample to the absolute oldest sisters i.e. the first born daughter of the household unlike the rest of the analysis where I defined oldest sister as the oldest out of those still living in household during some point in the panel. The dependent variables read, write, add, count are indicator variables for whether younger siblings are reported as capable of reading writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.

Appendix Table A4: IV results for younger brothers - Oldest sisters aged less than 20

	Read	Write	Add	Count	Schooling	Enrollment
Second stage IV results						
Oldest sister's years of schooling	0.060*	0.031	0.070**	0.079**	0.432***	0.083***
	(0.033)	(0.031)	(0.034)	(0.033)	(0.130)	(0.030)
First stage results						
Distance to closest girls' school	-0.413***	-0.428***	-0.417***	-0.412***	-0.406***	-0.404***
	(0.087)	(0.087)	(0.087)	(0.087)	(0.082)	(0.082)
F statistic	22.658	24.305	23.040	22.563	24.503	24.404
Number of observations	2,841	2,834	2,819	2,848	3,753	3,765
R^2	0.378	0.364	0.366	0.259	0.560	0.159

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.
This analysis limits the sample of oldest sisters to those under the age of 20, analyzing impact on younger brothers only. The dependent variables read, write, add, count are indicator variables for whether younger siblings are reported as capable of reading, writing, add/subtracting or counting. The dependent variable schooling refers to years of completed schooling as an outcome and enrollment is an indicator for whether the child is currently enrolled. The IV regression results reported are from the fullest control specification which controls for household variables, parents' education, wealth controls, distance to center, distance to closest government boys' school as well as district*year fixed effects.