Unmet Expectations:

The Impacts of School Construction on Female Outcomes in Rural Punjab, Pakistan

Salman Khan*

Job Market Paper

November 2020

Latest Version Here

Abstract

I study long-term and intergenerational effects of expanding educational opportunities for girls through school construction in Rural Punjab, Pakistan. This is a setting with low levels of education and significant gender inequality. Using a difference-in-differences estimation strategy that exploits variation across birth cohorts and regions in the timing of school construction, I find evidence of educational benefits for both girls and their children. However, I do not find corresponding improvements in female labor force participation and marriage market outcomes. My findings indicate that females in developing countries continue to face barriers in terms of later life outcomes despite improvements to their educational attainment especially in settings where social norms and economic position of males may mediate these effects.

JEL Codes: I2, J13, O15

*Department of Economics, University of Illinois, Chicago. Email: skhan275@uic.edu. I am indebted to my advisor Benjamin Feigenberg for providing advice and support at all the stages of the paper. I would also like to thank Javaeria Qureshi, Darren Lubotsky, Ben Ost, Bhashkar Mazumder, Steve Rivkin, Robert Kaestner, Usman Ghaus, Maciej Wachala, Katherine McElroy, Turab Hassan, Ahsan Farooqui, and Annie Khalid for valuable comments and suggestions.

I Introduction

"When sons go to schools that are far away we don't get worried, but for our daughter we get worried."

— Excerpt from *Pakistan Rural House Survey 2011*

Inadequate access to schools for girls remains a pervasive problem in much of the developing world despite being designated as a basic human right by the Convention on the Rights of the Child (UNESCO, 2017). Around 130 million girls of schooling age, equivalent to the entire population of Mexico or half the population of Indonesia, remain out of school (Evans et al., 2020).[1] Existing evidence, based largely on returns to secondary education for girls, suggests large long-term benefits associated with expanding educational opportunities for girls.[2] However, levels of female education remain quite low in many developing countries with the average education of females aged 25 and over being 0.8 years in Niger, 1.3 years in Yemen, 3.1 years in Haiti, and 3.6 years in India (Barro and Lee, 2013). Given the low levels of female education in many developing countries, it is important to investigate whether efforts to promote primary schooling for girls in these contexts offer similar promise in terms of later life outcomes.

In this paper, I study the short and long-term effects of a large scale primary school construction program in rural Punjab, Pakistan.[3] This program began in the 1960s in

[1]United Nations' figures reveal that the global number of girls out of school at the primary level has actually gone up from 32.0 million in 2015 to 34.3 million in 2018 (UNESCO, 2018).

[2]Studying longer term effects of education for girls in developing countries is crucial but existing evidence is relatively limited (McEwan, 2015). In a recent World Bank study, Wodon et al. (2018) estimate that developing countries stand to gain between \$15 trillion and \$30 trillion in lifetime productivity and earnings by improving educational opportunities for girls to 12 years of education. It is worth noting that these estimates by Wodon et al. (2018) provide orders of magnitude of potential impacts and do not take into account general equilibrium effects.

[3]Punjab, the most populous province in Pakistan, is home to more than half of the country's population. Since most of the school construction happened in rural areas and rural residents were more likely to benefit from improved access to schools, I restrict my analysis to these areas for this study. See Section on

response to low educational levels. The average years of schooling for the population aged 15 or above were 0.44 years for females and 1.8 years for males (Barro and Lee, 2013). Education levels were particularly lower for girls due to traditional customs and safety concerns that restrict girls' access to public spaces, particularly outside of their own community (Jacoby and Mansuri, 2011; Qureshi, 2018). These restrictive customs are enforced more strictly in rural areas, meaning that the majority of the girls are locked out of access to education, healthcare, and labor markets (Jacoby and Mansuri, 2011). High gender disparities in this setting suggest the possibility of multifaceted long-term benefits to increasing educational attainment for girls.

In order to evaluate the impact of the school construction program, I assemble a novel data set that measures the number of new schools available to each birth cohort at their predicted primary schooling age using administrative data from the Punjab Education Management Information System (EMIS). I combine this data with survey data from Demographic Health Surveys (DHS) and the Population Census to study whether investments in infrastructure can cause an increase in educational attainment. I exploit regional variation in the intensity of school construction and birth cohort variation in exposure to new schools. This approach is similar to the Difference-in-Differences specification in Duflo (2001) and adapts it to allow for a staggered approach to school construction in this setting.[4] Since schools are segregated by gender, I construct measures for new school availability separately by gender at the district level.

I find that the construction of an additional primary girls' school per 1000 girls of

Institutional Background for more details.

[4] School construction started in 1960s and continued until 1989. Please see section on Institutional Background for more details on the school construction program.

primary school age at the district level, leads to an increase of 4-5 percentage points in the likelihood that girls complete their primary education and to an overall increase in their years of education by 0.48 years. Given that the mean years of education for the relevant cohorts is only 2.4 years, these effects are quite large and explain around 30 percent of the overall growth in schooling for girls between birth cohorts 1954 to 1958 and 1984 to 1988. I do not find statistically significant effects of an additional boys' school on their educational attainment. Whereas the lack of impacts of the school construction program on male education are somewhat surprising, these results highlight the importance of mobility restrictions, due to social norms, that are only relevant for females in this setting.[5] My findings are therefore consistent with improved access to schooling being the key mechanism behind increased educational attainment for girls.

Given that the school construction program significantly increased girls' education, I use the Labor Force Surveys (LFS) to analyze effects on labor force participation, and the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) to study effects on the marriage market and intergenerational outcomes.[6] I find no significant effects on age at marriage, age at first birth, fertility rate, or child mortality; and find negative effects on female labor force participation for cohorts that are more exposed to the program. However, I do find evidence of intergenerational transmission of education with children born to mothers more exposed to new schools receiving more education with additional benefits to daughters.

[5]Since children walk to school, there is a gender disadvantage for girls as parents are not comfortable sending their daughters alone to walk long distances to school due to safety concerns and social norms (Jacoby and Mansuri, 2011).

[6]Labor Force Survey (LFS) is better suited to study outcomes related to labor force since it has more detailed information on employment-related outcomes over time. DHS and MICS both contain information on marital status for all household members of surveyed individuals above the age of 10.

With any type of quasi-experiment there are concerns that the results are driven by other factors not related to the policy being studied. I provide pre-trend tests by showing that new school construction is uncorrelated with trends in educational outcomes of older cohorts born prior to new school construction. I also run an age-based specification that demonstrates that the results for rural girls are indeed driven by younger cohorts who are more likely to benefit from new schools. Moreover, exploiting the fact that schools in this setting are segregated by gender, I run placebo tests to evaluate the effect of new schools on outcomes for cohorts of the opposite gender and find that area specific trends are not driving my results. As an additional robustness exercise, I further validate my results by using alternative data sets and by focusing on a specification that assesses the importance of endogenous migration as a response to school construction and find that my main results hold.

This study makes important contributions to our understanding of the effects of expanding educational opportunities in developing countries, especially for girls. First, I use quasi-experimental variation to study causal effects of school construction for Pakistan, a low-income country with significant baseline gender disparities in educational attainment (that persist to the present day). This work complements existing work on returns to secondary education by analyzing if the well documented benefits of schooling extend to low levels of education. Moreover, most of the existing work analyzing effects of school construction has focused on Indonesia, a middle income country, following the seminal study on the INPRES program in Indonesia by Duflo (2001).[7] However, there are significant differences in baseline educational attainment as well as in educational attainment by

[7] See for example Akresh et al. (2018), Mazumder et al. (2019a), Mazumder et al. (2019b), Zha (2019))

gender between the two settings.[8] It is therefore important to analyze the short and long term effects of school construction for Pakistan in order to shed light on the generalizability of past findings for low-income settings especially with significant gender disparity in baseline education levels.

Secondly, I contribute to the literature on investments in girls' education by examining impacts on longer term outcomes for females when there are no corresponding changes to the male education distribution. The heterogeneous effects of school construction by gender are different from other contexts such as Indonesia where both male and female education respond to school construction (Duflo, 2001, Akresh et al., 2018, Mazumder et al., 2019a).[9] Whereas I find evidence of intergenerational benefits of improved access to education for females with their children having higher educational attainment on average, I do not find corresponding improvements in female labor or marriage market outcomes as a result of improved educational attainment.[10] I find that without corresponding changes to male education, increasing education at the primary level may be insufficient to improve female outcomes beyond direct impacts on schooling (for both females and their children).

My findings also contribute to a growing literature that finds that females in developing countries continue to face barriers in terms of later life outcomes despite significant

[8]For birth cohorts around 1960, the mean years of education for males and females respectively were 3.0 and 0.8 for Pakistan compared to 4.6 and 4.1 for Indonesia (Barro and Lee, 2013).

[9]Whereas the point estimates of an additional school per 1000 children on years of education for girls' are twice as large compared to the Indonesian context (0.48 vs 0.23), I do not find any significant effects on male education. The effects on male and female education in Indonesia were quite similar (0.27 and 0.23 respectively) (Akresh et al., 2018).

[10]In contrast to the Indonesian context, female labor force participation goes down in this context. Moreover, there is no evidence of getting matched to better educated spouses since male education distribution has not changed as a function of school construction.

improvements to their human capital (Cheema et al., 2019; Edmonds et al., 2020; McKelway, 2020). Whereas low levels of education suggest high potential returns to education, social norms and the economic position of males may mediate the effects of improved access to education for females in such contexts. Taken together, these findings suggest that developing countries may need to ensure improvements to male education and social norms in addition to investing in higher levels of education for girls (Evans and Yuan, 2019, Evans et al., 2020). These findings are likely to be of interest to policy makers interested in returns to investments in girls' education as well as addressing gender inequality in developing countries.

The rest of the paper is organized as follows. Section II provides institutional background on the school construction program, Section III provides a brief description of the data used, and Section IV outlines the empirical methodology used to arrive at the results. Section V presents the results. Section VI outlines the pre-trend (falsification) tests and robustness checks undertaken to ensure the credibility of the results. Section VII concludes and provides guidance for policy work in this area.

II Institutional Background

In this paper, I focus on historical school construction in Punjab, the largest province in Pakistan. Punjab is sub-divided into 36 administrative units (Districts) and 150 sub-district units (Tehsils).[11] Pakistan has had a history of low educational attainment and Punjab is no exception to this. The overall literacy rate for population aged 10 or above

[11] A tehsil in Pakistan is the rough equivalent of a county in the US context.

was only 9.4 percent according to the 1961 census (Akhtar, 1963). There are significant gender disparities in educational attainment as the literacy rate was 14.9% for males and only 3.1% for females around this time (Akhtar, 1963). Public schools, which are segregated by gender, were the only viable schooling option for birth cohorts in Punjab until the late 1990s.[12] Enrollment rates are lower for girls due to concerns about girls' safety and traditional customs that restrict women's and girl's access to public spaces especially in rural areas (Jacoby and Mansuri, 2011; Qureshi, 2018).

Large scale school construction by the government in this context started with the Second Five Year Plan in 1960. Following the recommendations of the 1959 National Education Commission in light of the poor state of literacy in the country, the plan allocated allocated 78 million rupees, equivalent to $140 million dollars today, for the opening of 15,200 new primary schools (Commission, 1961). Subsequently, the third, fourth and fifth Five Year Plans also allocated funds for the construction of 42,500 new schools (all over Pakistan) through 1989 in order to improve the access to education (Kaiser, 1999). Figure 1 plots the number of new schools over time for Punjab. Since schools are segregated by gender, I plot female and male schools separately. As Figure 1 shows, most of the new schools were built in rural areas where more than two-thirds of the population lives. Moreover, distance to schools is typically greater in rural areas and educational attainment is on average lower than in urban areas. Since these areas were more likely to benefit from improved access to schools, I restrict my analysis to rural areas for this study.

There was considerable heterogeneity across districts in the number of new schools

[12] The wave of low-fee private school entry took off around 2000 in Pakistan (Andrabi et al., 2008).

built as the government designed the program to target districts in which enrollment was initially lower. Figures 2 (a) through (c) plot the heterogeneity for total new schools, girls' new schools and boys' new schools. In total, around 29,000 primary schools were constructed between 1960 and 1989 in rural Punjab out of which around 16,000 were girls' schools and 13,000 were boys' schools. School construction increased the stock of primary schools in rural Punjab from a baseline of approximately 1,100 primary schools for girls' and 4,400 boys' schools in 1959. It was acknowledged in the first Five Year Plan that girls needed to be provided with much greater opportunities for primary education since only 1.1 million of the 4.7 million children attending primary school were girls (Kaiser, 1999). However, most of the new schools initially constructed were boys' schools. As can be seen in Figure 1, it was only in the later years that greater emphasis was placed on constructing girls' schools.

Given that Pakistan has also experienced rapid growth in its population over this time period, the increase in number of new schools did not automatically translate into improved access to schooling in per capita terms. Table 2 shows the total number of schools per 1000 kids of school going age for the 1934-1989 period. It shows that whereas the construction of girls' schools has led to a monotonic increase in the number of schools per 1000 children, the same is not true for rural males. Schools per capita increased for males with initial school construction but as the focus switched largely towards constructing girls' schools in the 1980s, schools per capita for males started going down. In per capita terms, school construction meant that from a baseline of around 0.3 schools per 1000 girls (for the 1949-53 cohort), the number rose to 2.3 schools for the 1984-88 birth cohort. Similarly, for boys' schools this number went from around 1.8 schools per 1000 boys at baseline to around 2.5 schools for the 1984-88 birth cohort.

Figure 3 plots the educational attainment of rural cohorts over the 1960-1989 time period for the Punjab province. Figure 3 shows that rural females were a disadvantaged group with birth cohorts born around 1960 receiving less than 1 year of education on average. Before the start of school construction, primary school completion rates among primary school-aged girls in the 1954 birth cohort were around 10 percent (See Table 1). For the 1989 birth cohort, primary school completion rates had reached 51 percent. Primary school construction meant that most of the the girls who received any education ended up completing primary education since the fraction of girls getting any education and primary education is quite similar. It is evident from Table 1 that younger female cohorts made progress in bridging the gender gap in educational attainment. The gender gap in primary school enrollment rates decreased from 33 percentage points for the 1954-58 cohort points to 17 percentage points for the 1984-88 cohort. However, education levels in absolute terms remained low relative to the world for both genders and a smaller gender gap has persisted to date in educational attainment.

In short, the low education and per capita income levels along with the significant gender disparity in play make the Pakistani context a novel setting in which to analyze the long term effects of school construction. Studying the effects of primary school construction for Pakistan is therefore likely to be informative for other developing countries with similarly low levels of education particularly for girls.

III Data

To measure the impact of the school construction program, I construct a novel data set which measures the number of new schools available to each birth cohort at the time of their predicted entry into primary school using administrative data on schools in Punjab, Pakistan. The administrative data from Education Management Information System (EMIS) provides information on the school location, year of construction, school gender, school level as well as current information on student enrollment, school resources and number of teachers. I construct my measures of school construction at the district level for the main specifications since I can analyze effects on educational attainment and later life outcomes at this level.[13] I validate this data on school construction using annual reports of Punjab Development Statistics which contain data on new schools constructed over time at the District and Tehsil (sub-district) level.

I combine the administrative data on school construction with household data from multiple waves of the Demographic Health Survey (DHS) (1990-91, 2006-07, 2012-13 and 2017-18) for Pakistan to analyze impacts on educational attainment. I use the DHS as the primary data set for measuring effects on educational attainment since it is representative at the rural level for Punjab and contains information on the education status of all household members. I supplement my analysis on educational outcomes using data from the 10% sample of the 1998 population census. I also use the DHS to collect information on fertility, reproductive health, maternal health and child health.

[13]For my main specification, I construct my measure of new schools at the district level since I can combine different data sets at this level. Moreover, migration rates at this level are low. However, as a robustness exercise, I also run analysis at the Tehsil level and find similar results. Please see section on Robustness for more details.

Since the DHS does not include labor force participation questions for all household members, I use the annual Labor Force Survey (LFS) for the years 1990-91 to 2012-13 to analyze labor market outcomes. The LFS is therefore better suited to capture labor market participation rates and has more detailed information on employment-related outcomes. For analyzing marriage market outcomes, I use both the DHS and Multiple Indicator Cluster Survey (MICS 2003-04, 2006-07, 2010-11, 2013-14, 2017-18) as they both contain information on marital status of the entire household of the surveyed individuals. I additionally use MICS for intergenerational education and health-related outcomes as this survey is representative at the sub-district level for Punjab and has a large sample size for studying these effects. For information on the population of school-going children over time at the district level, I use data from the 1973, 1981 and 1998 population censuses of Pakistan (10% sample). Using these data sets, each of which has particular strengths, allows me to study the impact of school construction on a range of outcomes including education, employment, fertility, health, and marriage markets as well as human capital investments for the next generation. The next section outlines the empirical methodology used to analyze these effects.

IV Empirical Specification

I estimate a specification in which an individual's year of birth and region of residence jointly determine their exposure to the school construction program.[14] Given the stag-

[14] Ideally, I would have liked to use region of birth but I do not have information on region of birth in most of my data sets. However, I check the robustness of my results using the sub-sample (See Table A.5 in appendix) for which I do have information on region of birth and the results are quantitatively similar.

gered implementation of school construction, I build on the Difference-in-Differences approach similar to Duflo (2001) and adapt it for this context. Since students of age 5-13 typically attend primary school in Pakistan, I use the relevant number of new schools available to each birth cohort at their predicted primary school age as my measure of exposure to school construction.

I construct my measure of new schools at the district level.[15] This approach has two advantages. Firstly, given the migration rate between birth district and current district of residence are quite low at around 5% (Barkley, 1991), this approach allows me to evaluate the effect of the school construction without having to worry about endogenous migration as a function of school construction. Secondly, defining the treatment variable at the district level allows me to use data from multiple waves of the DHS as well as from other data sets such as LFS, Census, and MICS for analysis on longer-term outcomes and robustness analysis.

I start off by analyzing the effect of school construction on educational attainment.[16] I run a pooled regression of the following form:

$$Schooling_{ijk} = \alpha + \delta NewSchools(OwnGender)_{jk} + \beta_j + \gamma_k + \epsilon_{ijk} \tag{1}$$

where $Schooling_{ijk}$ refers to the schooling outcome of interest for individual i, residing in region j, born in birth cohort k. $NewSchools_{jk}$ is defined as the total number of new

[15] I also run an alternate specification at the Tehsil (sub-district) level and find similar results to my main specification. Please see the section on Robustness Checks and Table A6 for more details.

[16] We can think of school construction as affecting education outcomes through three main channels. Firstly, new primary schools reduce the distance that children have to walk to school. Secondly, new schools can lead to lower student teacher ratios as more teachers are added. Lastly, school construction can change expectations/ norms about returns to education for boys and girls.

schools (per 1000 kids of relevant gender) available to birth cohort j in region k at the time of their predicted age at entry to schooling. γ_k, and β_j refer to birth cohort and district fixed effects respectively. Robust Standard errors are clustered at the individual's district of residence. Since schools are segregated by gender, I run specification (1) separately by gender.

For my main schooling measures, I focus on whether an individual received any education, if they completed primary education, and their years of education using data from the DHS. As a supplementary exercise, I also evaluate the effect on secondary education using the Census data.

Since the ultimate goal of education is improvements in later life outcomes, I study the long-term effects of exposure to school construction on outcomes such as labor force participation, marriage markets, and fertility-related outcomes. I re-estimate equation (1) with alternative outcomes to study the relationships between longer term outcomes and school construction. Therefore, I run regressions of the following form:

$$Y_{ijk} = \alpha + \delta NewSchools(OwnGender)_{jk} + \beta_j + \gamma_k + \epsilon_{ijk} \tag{1'}$$

where Y_{ijk} refers to the relevant outcome of interest in the labor or marriage market for individual *i*, residing in region *j*, born in birth cohort *k*.

In terms of my outcome measures for labor force participation, I focus on employment status and whether the individual is working in the agriculture sector or not.[17] It is worth

[17]Given the lack of significant effects on educational attainment for rural males, I run longer term outcomes as a function of female schools. See the results section for more details.

noting that in this rural context, agriculture is the largest sector in terms of employment levels even for females who have much lower labor force participation rates compared to men. Focus on employment in general as well as in agriculture specifically allows me to evaluate the impact of school construction on females' likelihood of work and whether it leads to any shifts in the share of employment in agriculture in this setting.

For effects on the marriage market, I analyze effects on age at marriage, age at first birth, and total number of children. Females in this setting typically become part of the spouse's household upon marriage. Given that I have information on education, age, and wealth levels of spouses for rural females, I also analyze spousal characteristics as a function of the female school construction.

Since school construction started in the 1960s, I am able to study the effects of school construction on the next generation's outcomes as well. Whereas the main results on the marriage market and fertility are presented in the results section, it is worth noting before moving onto the empirical specification for intergenerational effects that I do not find any significant effects of female school construction on total fertility, age at marriage or age at first birth. Therefore, I can estimate the impact on children's schooling and other health-related outcomes based on the mother's extent of exposure to school construction in the absence of concerns associated with selection along these margins. I estimate reduced-form relationships between second generation outcomes and school construction using the following specification:

$$Y_{icjk} = \alpha + \delta NewSchools(Mother)_{jk} + \beta_j + \gamma_k + \omega_c + \epsilon_{ijk} \tag{2}$$

where Y_{icjk} refers to outcomes for child i, born to a mother in birth cohort j, in district k in birth year c. β_j, γ_k, and ω_c refer to mother birth cohort, district of residence, and child year of birth fixed effects respectively. Standard errors are clustered at the mother's district of residence.

V Results

This section describes the impact of school construction on educational attainment, and longer term impacts on labor and marriage markets as well as inter-generational effects on education and health.

A Results on Educational Attainment

Table 3 presents results on the relationship between school construction and educational attainment for rural cohorts using specification (1). For girls, an additional school at the district level per 1000 children of primary school age increased the likelihood of receiving any education and completing primary school by 5 percentage points and increased their years of education by 0.48 years on average (Table 3 Panel A). These estimates imply an increase in years of schooling of almost 1 year for females since the mean number of new schools at the district level is 2. Moreover, the mean education for females in the sample is around 2.4 years implying that these effects are quite large and can explain around 30 percent of the overall growth in years of schooling for girls between birth cohorts 1954 to 1958 and 1984 to 1988.

Table 3 also shows considerable gender differences in the effect of the new schools. In contrast to the large effects for females, the effects for males are close to zero and insignificant across all education levels implying that new male schools do not have an impact on male educational outcomes on average. These findings are consistent with improved access to schooling being the underlying mechanism for the results as customs that restrict female access to public spaces are not relevant for males in this (rural) setting.

Similar estimates on receiving any education and primary education for girls suggest that girls who enroll in school at any point in time are less likely to drop out and tend to complete primary education. There are smaller but significant effects on secondary education with secondary school completion rates increasing by 1.7 percentage points for rural girls. These results indicate that new primary schools allowed a fraction of the girls to transition to higher education levels. Given that completion rates at the secondary level are even lower than at the primary (4% vs. 22%), these represent larger percentage increases for secondary schools.

It is worth noting here that the effects of school construction for children in this context differ from those in Indonesia where both male and female outcomes respond to school construction (Duflo, 2001; Akresh et al., 2018; Mazumder et al., 2019a). For reference, an additional school per 1000 children through the INPRES program in Indonesia raised the mean years of education by around 0.23 years for females and 0.27 years for males (Akresh et al., 2018). In terms of point estimates, the effect size is twice as large for girls in this context. Moreover, the average years of schooling also differ between the two contexts. At baseline, 57% of females and 71% of males complete primary school in Indonesia and the mean years of schooling were 4.1 and 4.6 for females and males respectively (Barro and Lee, 2013). However, the average years of schooling for females and males in my sample are around 0.6 and 3.6 years and only 10% of the female sample and 43% of the male sample completes primary education at baseline in this context. Given these differences in education, the effect sizes are stronger in percentage terms for girls in this context compared to the Indonesian setting.

Since there there are no changes to male education as a result of school construction in this setting, I examine the effects of exposure to school construction on various long term outcomes as a function for female schools. This approach will allow me to shed light on whether complementarity in male and female education affect women's long-term outcomes. I present these long term effects on the labor and the marriage market for rural females in the next sub-section.

B Long-run labor market and Marriage Market impacts

Having observed increases in education in response to school construction for rural female cohorts,[18] Table 4 studies the effect of school construction on the labor market. Typically low or uneducated females work in the agriculture sector in the rural Punjab with overall labor participation rates being low at around 20 percent. Table 4 results show that rural females are less likely to be working in areas that received more schools. More specifically, I find that labor force participation rates decrease by around 4 percentage points for each additional school per 1000 children at the district level. Increased exposure to

[18]In this context, school construction is likely to affect individuals both through the direct channel whereby they get more education (Direct Treatment) or indirectly through outcomes of their peers who get treated (General Equilibrium Effects). Impacts on later life outcomes as a function of school construction should therefore be interpreted as an Intention-to-Treat effect (ITT) since they are a weighted average of both the direct and the indirect effects.

school construction suggests movement away from low or unpaid farm work. This effect seems to be driven by married females who are now 5 percentage points less likely to be working in the agriculture sector. Given the low labor force participation rates of around 20%, these declines in labor force participation are quite strong. The finding that females move away from the agriculture sector is consistent with Indonesian context where males move away from the agriculture sector as a result of school construction (Akresh et al., 2018). However, my results differ from the Indonesian context since male labor force participation increased significantly and there were smaller but insignificant increases for females as well in that context (Akresh et al., 2018). It is worth reemphasizing that LFP rates differ for females and males across the two countries. LFP for males and females in Indonesia are 95% and 64% respectively whereas in Pakistan, LFP rates are around 68% for men and around 21% for women. One possible explanation for the reduced labor force participation results for females in this context is the U-shaped female labor force participation function with education in developing countries such as India (Goldin, 1994; Fletcher et al., 2017; Chatterjee et al., 2018).

In terms of the marriage market, similar to the Indonesian context, women marry almost five years earlier than men, but there is no effect of exposure to school construction on the age of first marriage (See Table 5). Coefficients are slightly negative and statistically insignificant. I do not find evidence that women match with spouses who have significantly higher levels of education and education differences between spouses go down on average. This is perhaps unsurprising given that the average education levels for males in regions with new female schools did not increase. I do find some evidence of getting matched to wealthier spouses which can potentially explain part of the reduced labor force participation effects that we saw in Table 4 for married females.

I also do not find any significant effects on the number of total children,[19] age at first birth or child mortality (See Table 5). The lack of downstream effects on fertility are not surprising given the lack of effect on age at marriage. It would be interesting to look at inter-generational effects on the marriage market since parents traditionally make marriage decisions for their children. However, as most children born to these cohorts are not old enough to be married, data limitations prevent me from analyzing intergenerational effects on the marriage market.

C Intergenerational Impacts on Investments in Human Capital

Having observed effects of the school construction program on a wide range of outcomes, including education, employment, and the marriage market, I now investigate whether the effects extend to the next generation and affect the children of those exposed to the program. As explained in Section IV, second generation impacts are measured using the same difference-in-differences framework as first generation effects. The main explanatory variable is an interaction of the intensity of school construction in a mother's residence district with the degree with which the mother benefited from the program. Since female and male children might move outside of the their parent's family, I restrict my children below age of 17 since at this age children are unlikely to have left their parent's household for work or marriage reasons. I also include birth cohort fixed effects for children in order to ensure that comparisons take place across children of the same

[19] Whereas data sets such as the DHS only carry information on children living in the same household, for roughly a 10% sub-sample of the data I have detailed information on total births.

age. For effects on child health, I use information on children below the age of 5 since anthropometric measures are measured for all children in this range both in the DHS and MICS.

Table 6 (and Table 2 in the appendix) show the reduced form results of a mother exposure to school construction on the next generation's educational outcomes. In order to compare the educational outcomes of children of similar ages, I include additional controls for birth year fixed effects. The results show there are indeed benefits of the mother's education that extend to the next generation. These results are consistent with (Andrabi et al., 2012) who find evidence of returns to low levels of mother education in the rural Pakistani context.

I also check for effects on health of children as well as healthcare utilization by mothers more exposed to school construction. I do not find any positive effects on child health as well as healthcare utilization. In fact, I find somewhat adverse average effects on weight and height of children below the age of 5 for mother-cohorts that were more exposed to new schools (See Appendix Tables 3-4). The lack of positive effects on healthcare utilization are consistent with the limited results on women empowerment in the Indonesian context (Samarakoon and Parinduri, 2015), which has fewer restrictions on women in the public space than Pakistan.

VI Threats to Identification and Robustness checks

My main empirical specification exploits variation across regions and birth cohorts in exposure to school construction. This implies that my results on educational attainment and later life outcomes could be driven by trends in educational attainment or by some omitted factor such as a higher demand of education that I have failed to control for. In this section, I discuss possible threats to my identification strategy and provide evidence on the robustness of my results. Given that I use a Difference-in-Differences specification for my main empirical specification, I start by showing evidence on pre-trends.

A Pre-Trend test for Difference-in-Differences

Since school construction followed a staggered design, I first check to see if school construction is correlated with trends in educational attainment of older cohorts unlikely to benefit from the new schools. Typically students aged 5-13 attend primary school in Pakistan. Given that school construction started around 1960, cohorts born prior to 1950 are unlikely to benefit from these schools.

I start by running a regression of the following form:

$$S_{ijk} = \alpha + \beta_j + \gamma_k + \sum_{l=1934-38}^{1984-88} (NewSchools_k * d_{il})\gamma_{1l} + \epsilon_{ijk} \tag{3}$$

where S_{ijk} refers to schooling of individual *i*, in birth cohort *j*, in district *k*. β_j refers to birth cohort fixed effect, γ_k refers to district fixed effect d_{il} is a dummy that indicates whether individual *i* belongs to birth cohort l. $NewSchools_k$ is defined as the total number of new schools in region k between 1960-1989, scaled by population (in thousands) of school aged children in district k. I cluster standard errors at the district level, and given that the data allows estimating the effects of school construction separately by gender, I estimate Equation (3) separately for males and females.

The identifying assumption behind this specification is that counterfactual trends are uncorrelated with school construction. I test this by looking at the relationship between school construction and educational attainment of older cohorts. Figures 1(a) through 1(d) (in the appendix) plot the coefficients obtained from gender-specific regressions of educational outcomes on new schools. These figures provide visual evidence that trends in educational attainment of older cohorts are uncorrelated with subsequent school construction for both rural male and rural female schools.

Further, with the continued school construction, younger cohorts should benefit more from the school construction happening between 1960 and 1989. Based on Figures 1(a) and 1(c) which plot effects on receiving any education and completing primary education, the effect size does get larger for more recent female cohorts. Moreover, Figures 1(b) and 1(d) show that the effect size remains insignificant for male cohorts both for receiving any education as well as for completing primary education. In short, the figures provide evidence that the new schools matter more for younger rural female cohorts and that they are not correlated with trends in educational attainment of older cohorts. Therefore, these figures present evidence in support of school construction being a plausibly exogenous supply shock to access to schooling for the relevant female cohorts.

B Age-based Specification: Who Benefits from School Construction

There may be concerns that the results are driven by some omitted factor which leads to spurious correlation between school construction and educational attainment. I use an age-based specification to investigate if effects are indeed driven by children of primary-school age and not older cohorts.

For this purpose, I use data from the DHS and the Population Census to create a retrospective history of the number of schools available to each individual at different ages in their life cycle based on their birth cohort and district of residence. Thereafter, I analyze the effect of number of schools of own gender per 1000 children on schooling outcomes for the individual while controlling for year and district fixed effects. I run regressions of the following form separately for each age, $a_i \in$ {0,20}

$$Schooling_{ikt} = \alpha + \delta NewSchools(Female)_{kt} + \gamma_t + \omega_k + \epsilon_{ikt} \tag{4}$$

where $Schooling_{ikt}$ refers to the schooling outcome of interest for individual *i*, residing in region *k*, in year t. $NewSchools_{kt}$ is defined as the total number of new schools (per 1000 kids) available to the individual on the basis of his region k in year t at age a. γ_t, ω_k refer to year and district fixed effects respectively.[20]

This specification compares the effect of being exposed to more schools at different ages on educational attainment while controlling for fixed differences between districts over time as well as any time trends in educational attainment. The idea of this specifi-

[20] I use robust standard errors that are clustered at the district level.

cation test is that more schools per 1000 children should only matter for younger cohorts who are likely to benefit from school construction. Cohorts older than 13 are likely to be too old to benefit from more primary schools. This specification therefore serves as a test to validate my results.

Figure 4(a) and 4(b) plot the coefficients on the likelihood of completing primary and any education respectively for rural females using the age-based specification. Figure 4 shows clear evidence that it is indeed children younger than seven who benefit the most from the newly constructed schools. Girls aged seven to ten also benefit, although to a lesser extent, as propensity to enroll or stay on in school likely decreases with their age at the time of the school construction. In short, the age-based specification presents evidence in support of the research design by showing that effects for girls' education are indeed being driven by the younger female cohorts. These cohorts are more likely to benefit from improved access to schooling given their age and the mobility restrictions that females face in this setting.

C Placebo regressions to check the effect of school construction

Another potential threat to my identification strategy could be that effects on female education are driven by area trends in education (such as parental demand for education) and not by an exogenous increase in supply of new schools. In order to mitigate concerns in this regard, I exploit the fact that schools in this setting are segregated by gender. I run placebo tests to evaluate the effect of new schools on outcomes for cohorts of the opposite gender. More specifically, I run the following specifications:

$$Y_{ijgk} = \alpha + \delta NewSchools(Othergender)_{g'jk} + \beta_j + \gamma_k + \epsilon_{ijk} \tag{5}$$

where Y_{icjk} refers to educational outcomes for individual i, of gender g born in birth cohort j, and residing in district k. Schooling measure is defined in terms of schools per capita of opposite gender g'. β_j and γ_k, refer to birth cohort, and district of residence fixed effects respectively.[21]

Table 1 in the Appendix shows these results. I do not find any statistically significant effects of new male schools on female cohorts and vice versa. The results of these placebo tests, examining the effect of new schools on outcomes for cohorts of the opposite gender, allow me to rule out area specific trends as driving my results. Running these placebo specifications therefore allows me to evaluate more credibly that the effects on female education are likely to be driven by the supply of new schools and not due to some omitted factor such as parental demand for education.

D Using District of Birth instead of District of Residence

An additional concern with my empirical specification is regarding using district of residence at the time of the survey for construction of my measure of exposure to school construction since information on district of birth is not available for most of the data sets used (DHS, MICS, and Population Census 1998). Using district of residence at the time of the survey instead of district of birth leads to concerns that my results could be affected

[21] Standard errors are clustered at the district level.

by migration of individuals between primary schooling age and when they are measured in the survey.[22] Whereas, I argue in my main specification that these concerns are unlikely to be first-order since migration rates between districts are low in this setting, I provide a more formal test here by focusing on data from an earlier Population Census for which I have information both on district of residence and district of birth.

I therefore check if my results could be driven by endogenous migration by focusing on a specification that uses birth district at the time of the survey. As Table 5 (in the Appendix) shows, results of specifications using the district of birth and the district of residence are quite similar. Given the low migration rates between districts and my results in Table 5, endogenous migration as a function of school construction is less likely to be a concern in this context.[23]

E Additional Robustness Tests

I construct my measure of school construction at the district level for my main specifications. This might lead to a concern that I am wrongly assuming linearity in aggregating outcomes at this level. In order to mitigate concerns in this regard, in additional robustness tests, I use waves of MICS data that are representative at the Tehsil(sub-district) level. Using variation in school construction at the Tehsil level, I analyze impacts of school construction on educational attainment (See Table 6 in the Appendix). I find significant effects of girls' schools on their education but do not find effects of boys' schools

[22] Additionally, it could also be a concern if parents moved to districts where more new schools were constructed. However, existing evidence suggests that migration for educational reasons is very low (Memon, 2005).

[23] It is worth noting here that in the Indonesian context, district sizes are smaller and therefore inter-district mobility is higher.

as in my primary specification.

Moreover, in order to rule out effects being driven by secondary schools, I also run a specification in which I control for the construction of new secondary schools. My results are robust to the inclusion of controls for new secondary schools and are therefore unlikely to be driven by them. The series of robustness tests allow me to check various threats to the identification strategy and enable me to more credibly interpret the causal impacts of school construction program on educational attainment and later life outcomes for rural females in this setting.

VII Summary and Concluding Remarks

This paper studies the long-term and intergenerational effects of a large supply side intervention aimed at expanding primary school access in a setting where education levels are low particularly for females (Rural Punjab, Pakistan). Using administrative data, I construct a novel measure for exposure to new schools at the district level that allows for staggered school construction in this context. Using a difference-in-differences estimation strategy that exploits variation across birth cohorts and regions in their exposure to school construction, I analyze the impact of the program on educational attainment as well as later life outcomes. Since schools are segregated by gender, I analyze impacts of new schools separately by gender using data from household surveys and the population census. I find that new schools for girls have a significant impact on improving their educational attainment as measured by their likelihood of receiving any education, completing primary education, and years of schooling. I do not find any significant effects of school construction for males on their educational outcomes. These findings indicate that improved access to schooling is driving the results on educational attainment since mobility restrictions, due to social norms, are only relevant for females in this setting.

In terms of long-term outcomes, my results suggest that improved educational attainment does not lead to significant effects on the marriage market (as measured by likelihood of marriage, age at marriage, education of spouse) or fertility related outcomes (age at first birth, total children born, child mortality). Exposure to school construction does alter marriage market outcomes for males with their spouses being more educated. However, I do not find evidence of similar effects for females since the male education distribution has not changed as a function of school construction. Moreover, as adults, females who are more exposed to the program are less likely to be part of the labor force. This effect is driven by the lower likelihood of them working in the agriculture sector. The benefits of increased education for females are, however, transmitted to the next generation with additional benefits accruing to daughters, a finding consistent with past work (Thomas, 1994; Akresh et al., 2018).

The lack of impacts on later life outcomes such as age at marriage, age at first birth, total children and labor force participation for females despite the increased educational attainment suggest the importance of other constraining factors. A possible reason for the lack of desired effects on later life outcomes is that the baseline education levels are quite low. Existing work on returns to secondary schooling is based on the notion that education equips girls with the knowledge and ability to make life choices that improve their welfare (Duflo et al., 2012; Lundberg and Pollak, 1993). However, it is unclear whether increases in primary education for girls can lead to similar effects particularly without corresponding changes to male education and norms (Evans and Yuan, 2019; Evans et al., 2020.)

My findings are also consistent with and contribute to a growing literature that finds that females in developing countries continue to face barriers in their ability to make important decisions related to the labor market (Cheema et al., 2019; McKelway, 2020) and the marriage market (Edmonds et al., 2020) on their own. These barriers persist despite significant improvements in human capital through education (Edmonds et al., 2020), skill acquisition (Cheema et al., 2019) or increased generalized self-efficacy (McKelway, 2020). More research is therefore needed in order to better understand interventions that can be designed to offer feasible ways to improve long-term outcomes for females in settings where social norms and the economic position of males may mediate the effect of increased educational attainment on longer term outcomes.

References

Akhtar, J. (1963). Literacy and education: Fifth release from the 1961 census of pakistan.

Akresh, R., Halim, D., and Kleemans, M. (2018). Long-term and intergenerational effects of education: Evidence from school construction in indonesia. Technical report, National Bureau of Economic Research.

Alderman, H., Kim, J., and Orazem, P. F. (2003). Design, evaluation, and sustainability of private schools for the poor: the pakistan urban and rural fellowship school experiments. *Economics of Education review*, 22(3):265–274.

Andrabi, T., Das, J., and Khwaja, A. I. (2008). A dime a day: The possibilities and limits of private schooling in pakistan. *Comparative Education Review*, 52(3):329–355.

Andrabi, T., Das, J., and Khwaja, A. I. (2011). *Students today, teachers tomorrow? Identifying constraints on the provision of education*. The World Bank.

Andrabi, T., Das, J., and Khwaja, A. I. (2012). What did you do all day? maternal education and child outcomes. *Journal of Human Resources*, 47(4):873–912.

Andrabi, T. R., Das, J., and Khwaja, A. I. (2010). Education policy in pakistan: A framework for reform.

Andriano, L. and Monden, C. W. (2019). The causal effect of maternal education on child mortality: Evidence from a quasi-experiment in malawi and uganda. *Demography*, 56(5):1765–1790.

Assembly, U. G. (1989). Convention on the rights of the child. *United Nations, Treaty Series*, 1577(3).

Barkley, A. P. (1991). The determinants of interdistrict labour in-migration in pakistan, 1971-1980. *The Pakistan Development Review*, pages 275–296.

Barro, R. J. and Lee, J.-W. (2001). International data on educational attainment: updates and implications. *oxford Economic papers*, 53(3):541–563.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104:184–198.

Chatterjee, E., Desai, S., and Vanneman, R. (2018). Indian paradox: Rising education, declining womens'employment. *Demographic Research*, 38:855.

Cheema, A., Khwaja, A. I., Naseer, F., and Shapiro, J. N. (2019). Glass walls: Experimental evidence on access constraints faced by women. Technical report, Mimeo, Harvard University.

Commission, P. P. (1961). *The Second Five Year Plan, 1960-65: Including Revised Estimates*. Manager of Publications.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review*, 91(4):795–813.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78.

Edmonds, E. V., Feigenberg, B., and Leight, J. (2020). Advancing the agency of adolescent girls. Technical report, National Bureau of Economic Research.

Evans, D. K., Akmal, M., Jakiela, P., et al. (2020). Gender gaps in education: The long view. *Center for Global Development Working Paper*, 523.

Evans, D. K. and Yuan, F. (2019). *What We Learn about Girls' Education from Interventions that Do Not Focus on Girls*. The World Bank.

Fletcher, E., Pande, R., and Moore, C. M. T. (2017). Women and work in india: Descriptive evidence and a review of potential policies.

Glewwe, P. and Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier.

Goldin, C. (1994). The u-shaped female labor force function in economic development and economic history. Technical report, National Bureau of Economic Research.

Hamid, S. (2016). What's different about islam in malaysia and indonesia? vol. 2018. *World Affairs: The Aspen Institute*.

Jacoby, H. G. and Mansuri, G. (2011). *Crossing boundaries: gender, caste and schooling in rural Pakistan*. The World Bank.

Kaiser, B. (1999). History of education policy making and planning in pakistan.

Lundberg, S. and Pollak, R. A. (1993). Separate spheres bargaining and the marriage market. *Journal of political Economy*, 101(6):988–1010.

Mahmud, M., Musaddiq, T., Said, F., and Sabir, M. (2010). Internal migration patterns in pakistan—the case for fiscal decentralisation [with comments]. *The Pakistan Development Review*, pages 593–607.

Male, C. and Wodon, Q. (2018). Girls' education and child marriage in west and central africa: Trends, impacts, costs, and solutions. In *Forum for Social Economics*, volume 47, pages 262–274. Taylor & Francis.

Mazumder, B., Rosales-Rueda, M., and Triyana, M. (2019a). Intergenerational human capital spillovers: Indonesia's school construction and its effects on the next generation. In *AEA Papers and Proceedings*, volume 109, pages 243–49.

Mazumder, B., Triyana, M., and Fernanda Rosales, M. (2019b). Social interventions, health and wellbeing: The long-term and intergenerational effects of a school construction program.

McCleary-Sills, J., Hanmer, L., Parsons, J., and Klugman, J. (2015). Child marriage: a critical barrier to girls' schooling and gender equality in education. *The Review of Faith & International Affairs*, 13(3):69–80.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, 85(3):353–394.

McKelway, M. (2020). Women's employment in india: Intra-household and intra-personal constraints. Technical report, Working paper, Massachusetts Institute of Technology.

Memon, R. (2005). Determinants and trends of internal migration in pakistan. *Internal Labour Migration in Pakistan. Institute of Developing Economies, Japan External Trade Organisation, Chiba, Japan*.

Osili, U. O. and Long, B. T. (2008). Does female schooling reduce fertility? evidence from nigeria. *Journal of development Economics*, 87(1):57–75.

Psacharopoulos, G. and Patrinos, H. A. (2018). *Returns to investment in education: a decennial review of the global literature*. The World Bank.

Qureshi, J. A. (2018). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal*, 128(616):3285–3319.

Samarakoon, S. and Parinduri, R. A. (2015). Does education empower women? evidence from indonesia. *World Development*, 66:428–442.

Thomas, D. (1994). Like father, like son; like mother, like daughter: Parental resources and child height. *Journal of human resources*, pages 950–988.

UNESCO (2017). Reducing global poverty through universal primary and secondary education.

UNESCO (2018). One in five children, adolescents and youth is out of school.

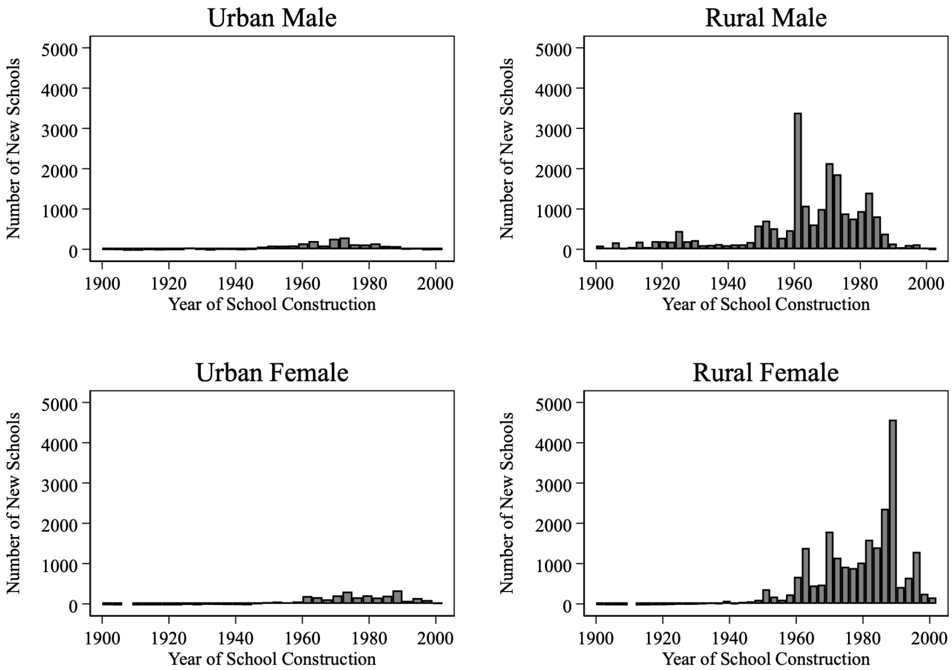
Watch, H. R. (2015). *World report 2015: Events of 2014*. Policy Press.

Wodon, Q., Montenegro, C., Nguyen, H., and Onagoruwa, A. (2018). *Missed Opportunities: The high cost of not educating girls*. World Bank.

Zha, D. (2019). Schooling expansion and the female marriage age: Evidence from indonesia. Technical report, Technical report.

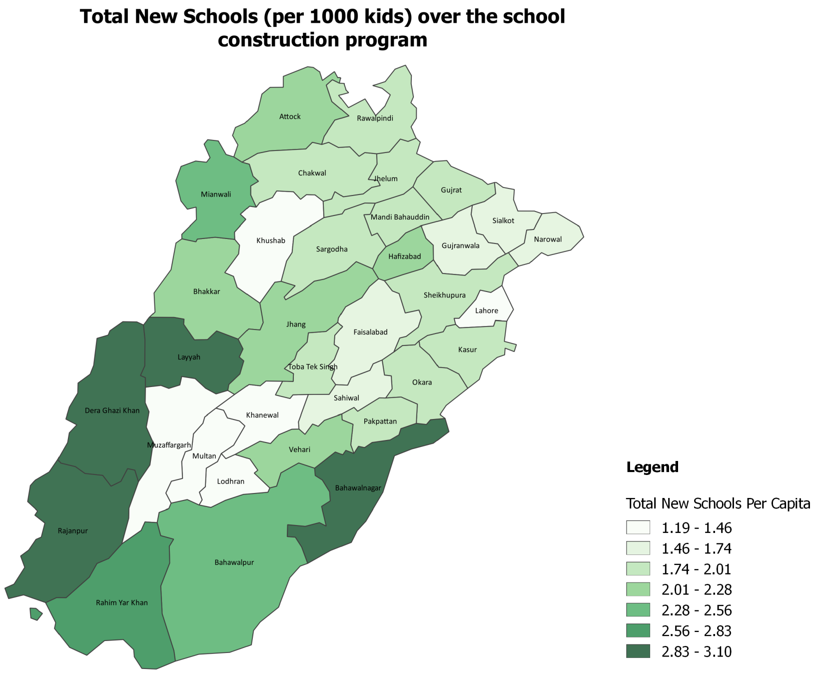
Figures

Figure 1: **Number of New Schools by Gender and region for Punjab, Pakistan (1900-2000)**



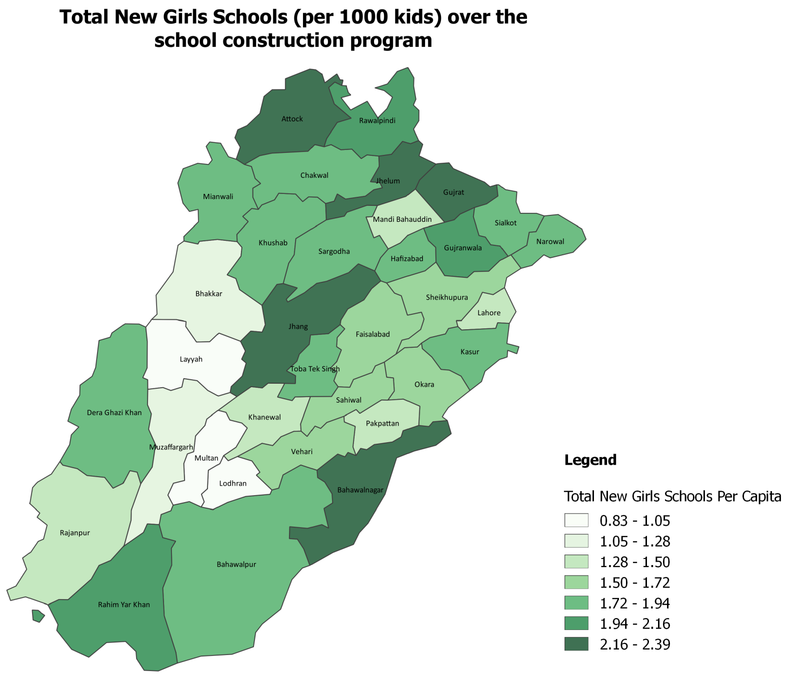
Note: Figure 1 plots the total number of new schools constructed between 1900-2000 for Punjab, Pakistan. Number of schools of own gender are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS).

Figure 2a: **Heterogeneity of School Construction at the District Level for Rural Punjab**



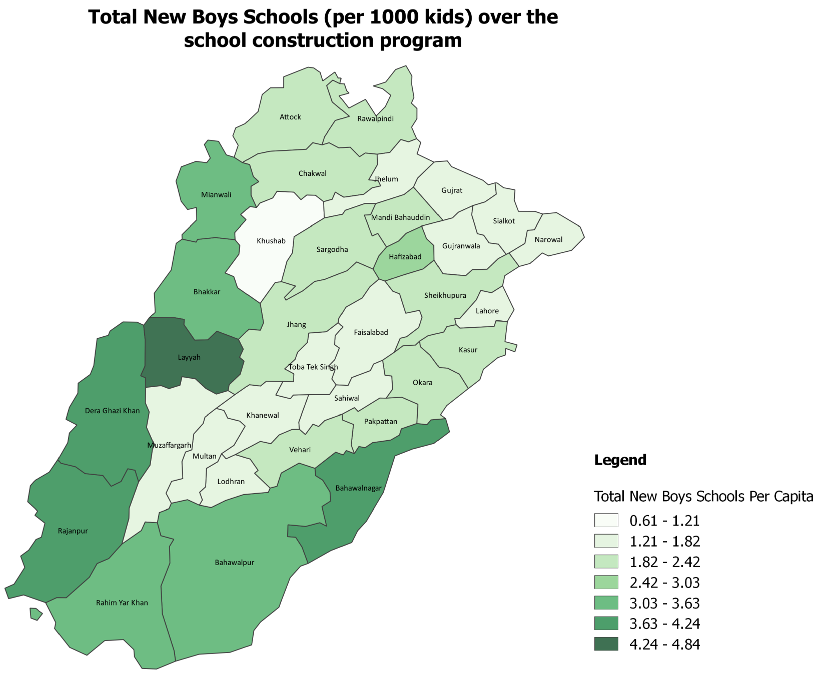
Note: Figure 2a plots the total number of new schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per capita measure of school construction is obtained by dividing new schools at the district level by the total population of primary school-going age children in thousands.

Figure 2b: **Heterogeneity of Female School Construction at the District Level**



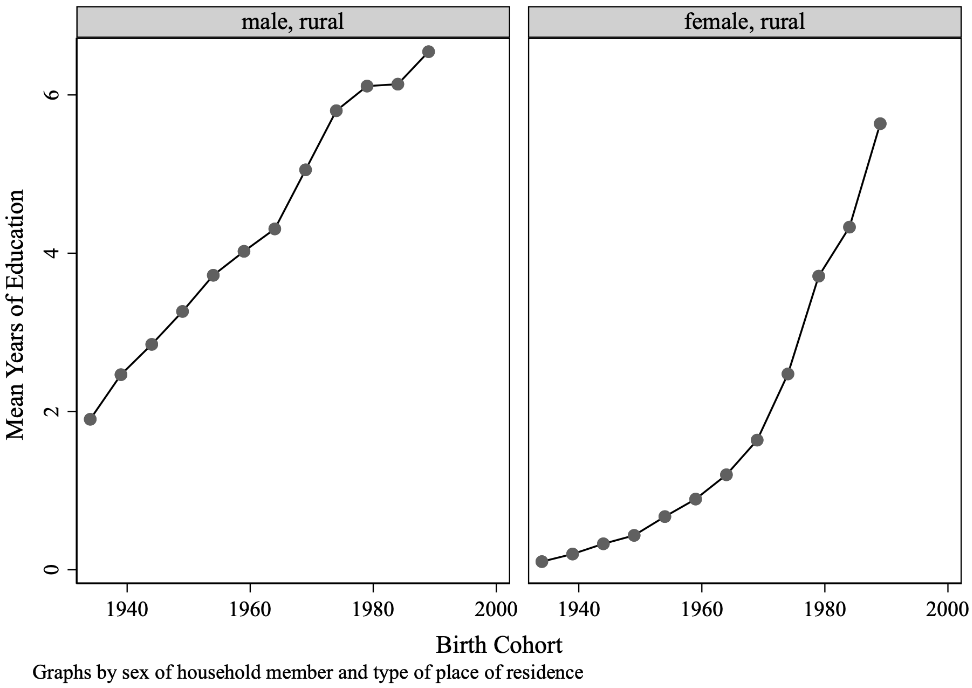
Note: Figure 2b plots the total number of new girls' schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per capita measure of school construction is obtained by dividing new schools at the district level by the total population of primary school-going age females in thousands.

Figure 2c: **Heterogeneity of Male School Construction at the District Level**



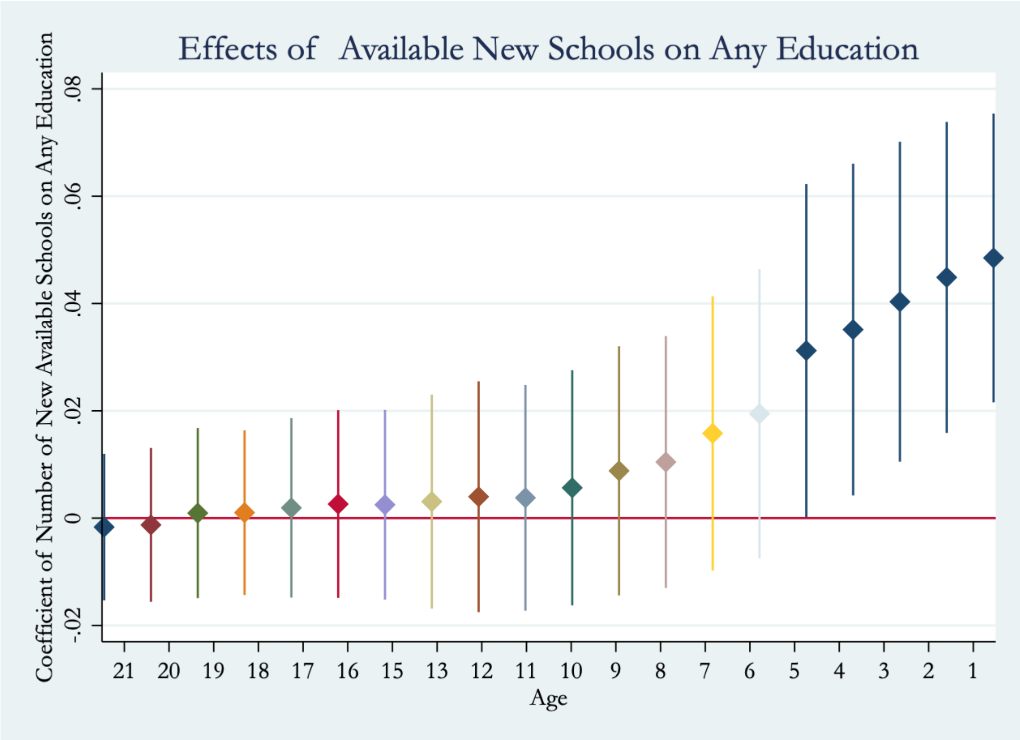
Note: Figure 2c plots the total number of new boy's schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per capita measure of school construction is obtained by dividing new male schools at the district level by the total population of primary school-going age males in thousands.

Figure 3: **Educational Attainment Over time (Punjab, Pakistan)**



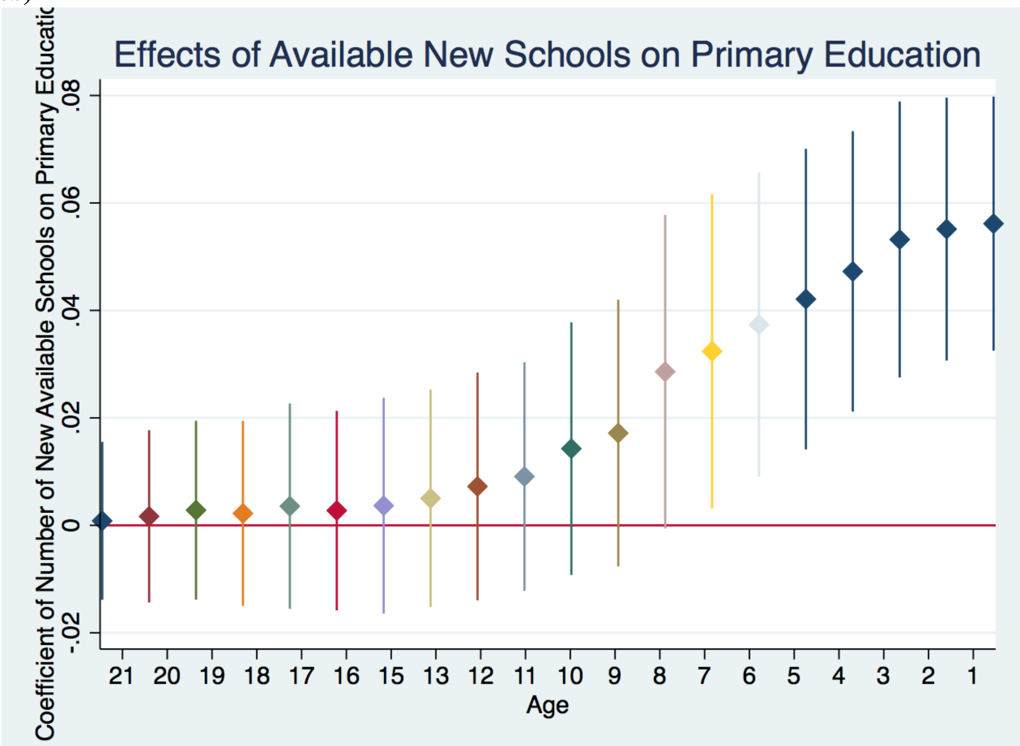
Note: Figure 3 plots the mean years of education for rural cohorts by gender in Punjab, Pakistan over the 1934-1989 period. It uses data on educational attainment from multiple waves of DHS. Rural regions are defined as per the DHS definition.

Figure 4a: **Coefficient of Schools of Own Gender on Any Education (Rural Females)**



Note: Figure 4a plots the coefficients corresponding to specification 4 in Robustness Checks subsection B for the any schooling measure for rural females. Any education is defined as a binary variable equal to 1 if the person received at least 1 year of education. Number of female schools at the district level at age is constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS) and using reported age in Population Census 1998 (10% sample). Per capita measure of school construction is obtained by dividing new female schools at the district level by the total population of primary school-going age females in thousands. Solid lines represent 95% Confidence Intervals.

Figure 4b: **Coefficient of Schools of Own Gender on Primary Education (Rural Females)**



Note: Figure 4b plots the coefficients corresponding to specification 4 in Robustness Checks subsection B for the primary schooling measure for rural females. Primary Education is a binary variable equal to 1 if the person completed 5 years of education. Number of female schools at the district level at age is constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS) and using reported age in Population Census 1998 (10% sample). Per capita measure of school construction is obtained by dividing new female schools at the district level by the total population of primary school-going age females in thousands. Solid lines represent 95% Confidence Intervals.

Tables

Table 1: **Educational Attainment of Cohorts over time for Punjab, Pakistan**

Birth Cohort	Rural Male			Rural Female		
	Any Education	Primary Education	Mean Education	Any Education	Primary Education	Mean Education
1934-38	0.286	0.226	1.93	0.020	0.012	0.103
1939-43	0.349	0.303	2.54	0.034	0.028	0.200
1944-48	0.384	0.340	2.87	0.054	0.047	0.312
1949-53	0.426	0.382	3.34	0.078	0.063	0.439
1954-58	0.467	0.428	3.75	0.120	0.097	0.694
1959-63	0.511	0.466	4.07	0.150	0.125	0.910
1964-68	0.538	0.493	4.36	0.198	0.168	1.257
1969-73	0.610	0.558	5.08	0.253	0.214	1.683
1974-78	0.683	0.631	5.84	0.341	0.299	2.549
1979-83	0.744	0.678	6.16	0.481	0.435	3.793
1984-88	0.767	0.685	6.11	0.570	0.510	4.454
1989-93	0.815	0.700	6.66	0.660	0.588	5.641

Note: Table 1 shows the mean educational attainment for rural birth cohorts by gender using data from multiple waves of the DHS. Any education is defined as a binary variable equal to 1 if the person received any education. Primary Education is a binary variable equal to 1 if the person completed 5 years of education. Mean Education captures years of education in single years.

Table 2: **Schools per 1000 children over time for Rural Punjab, Pakistan (By Gender)**

Schools (per 1000 kids)	1934-38	1939-43	1944-48	1949-53	1954-58	1959-63	1964-68	1969-73	1974-78	1979-83	1984-88
Male	1.38	1.21	1.24	**1.77**	1.84	2.97	2.92	3.49	3.23	2.97	**2.48**
Female	0.11	0.12	0.15	**0.27**	0.34	0.78	0.90	1.33	1.47	1.67	**2.32**

Note: Number of schools of own gender are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). In order to get per 1000 children measure, I divide schooling measure by total children (in thousands) of own gender of primary school age (5-14) measured using Population Census 1998.

Table 3: **Effect of New Schools on Educational Attainment of Rural Cohorts**

Panel A: Results using DHS

	Any Education		**Primary Education**		**Years of Education**	
Dependent variable	Male	Female	Male	Female	Male	Female
New Schools Available per 1000 kids (Own Gender)	0.00319	0.0485	-0.000729	0.0503	-0.0372	0.480
	(0.0121)	(0.0250)*	(0.0116)	(0.0255)*	(0.118)	(0.267)*
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
1954-59 Birth Cohort Mean	0.467	0.120	0.428	0.097	3.750	0.694
Observations	49115	52180	49115	52180	49115	52180

Panel B: Results using Census

	Any Education		**Primary Education**		**Secondary Education**	
Dependent variable	Male	Female	Male	Female	Male	Female
New Schools Available per 1000 kids (Own Gender)	0.00895	0.0501	0.00471	0.0399	0.00257	0.0156
	(0.0120)	(0.0208)**	(0.00933)	(0.0157)**	(0.00206)	(0.00434)***
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
1954-59 Birth Cohort Mean	0.449	0.120	0.386	0.093	0.038	0.005
Observations	1749614	1674949	1605926	1541133	1210386	1184551

Note: Table 3 presents the results from Specification (1). Sample used in this table are individuals aged 18 and above. Panel A uses data from multiple waves of DHS, whereas Panel B uses data from Population Census 1998. Number of New Schools is constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). Cohort Size is defined as the population of primary school age children in thousands. Robust Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: **Effect on Labor Market using Labor Force Surveys (Rural Females)**

	Working Status				**Work on Farm**				**Non-Farm Working Status**			
Dependent variable	Full Sample	Married	Spouse	Not Married	Full Sample	Married	Spouse	Not Married	Full Sample	Married	Spouse	Not Married
New Schools Available per 1000 kids (Female)	-0.0423	-0.0643	0.00385	-0.0124	-0.0359	-0.0468	0.00214	-0.0212	-0.00765	-0.0176	0.00171	0.00879
	(0.0172)**	(0.0188)***	(0.00962)	(0.0214)	(0.0123)***	(0.0131)***	(0.000644)***	(0.0157)	(0.00674)	(0.00985)*	(0.00963)	(0.00954)
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	181921	129568	73791	52353	181923	129569	73819	52354	181921	129568	73791	52353

Note: Table 4 shows results on labor market for rural female aged 18 or older. Working status defined as binary variable equal to 1 if the individual is employed and 0 otherwise. Work on farm is defined as a binary variable equal to 1 if the individual works in the agriculture sector. Non-farm working status is a binary variable equal to 1 if the individual works in a non-agriculture sector. Columns 1, 5, 9 show results for total sample. Columns 2, 6, 10 show results for the sample of rural females who are married. Columns 3, 7, 11 show results for spouses of rural females. Columns 4, 8, 12 show effects for rural females who are not married. Robust Standard errors at clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: **Effect of New Schools on Marriage Market Outcomes of Rural Females**

Panel A: Results on Marriage Outcomes using DHS

Dependent variable	Ever Married	Age at Marriage	Child Marriage	Number of Children	Age at First Birth
New Schools Available per 1000 kids (Female)	0.00418	-0.000461	0.0315	0.161	-0.0641
	(0.0177)	(0.304)	(0.0255)	(0.177)	(0.259)
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes
1954-59 Birth Cohort Mean	0.988	18.9	0.436	6.01	21.1
Observations	50832	5985	5983	6074	6074

Panel B: Results on Marriage Outcomes using MICS

Dependent variable	Ever Married	Age at Marriage	Child Marriage	Number of Children	Age at First Birth
New Schools Available per 1000 kids (Female)	-0.00253	-0.291	0.0253	0.0396	-0.177
	(0.00565)	(0.282)	(0.0222)	(0.202)	(0.238)
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes
1954-59 Birth Cohort Mean	0.986	19.9	0.365	6.4	22.0
Observations	121684	64397	64397	63905	58559

Panel C: Results on Spousal Outcomes using MICS

Dependent variable	Education (Spouse)	Age Difference with Spouse	Wealth Index Score	Spouse in top 2 Quartiles	Blood relation with Spouse
New Schools Available per 1000 kids (Primary)	0.187	-0.0873	0.0596	0.0225	0.0161
	(0.118)	(0.141)	(0.0197)***	(0.0132)*	(0.0169)
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes
1954-59 Birth Cohort Mean	3.44	6.7	-0.22	0.296	0.683
Observations	32180	32220	12437	12437	6073

Note: Results correspond to specification (1') in long-term outcomes. Panel A uses data from DHS, whereas Panel B uses data from MICS. Panel C uses data from MICS except last column is made using DHS since MICS does not carry information on blood relationship between spouses. Sample used in these tables are rural girls aged 18 and above in column 1 and aged 18-49 in columns 2-6. Sample size drops after column 1 particularly for DHS since questions on marital outcomes are only asked to a sub-set of female (aged 15-49) in the long questionnaire. Ever Married is a binary variable equal to 1 if the individual has ever been married. Age at Marriage measures age in completed years at the time of first marriage. Child marriage is a binary variable equal to 1 if an individual married before the age of 18. Number of Children measures total children ever born. Age at first birth measures age in completed years at the time of first birth. Wealth Index Score is used as calculated by MICS. Spouse in top 2 quartiles is a binary variable equal to 1 if spouse of rural female has above median wealth. Robust Standard errors are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

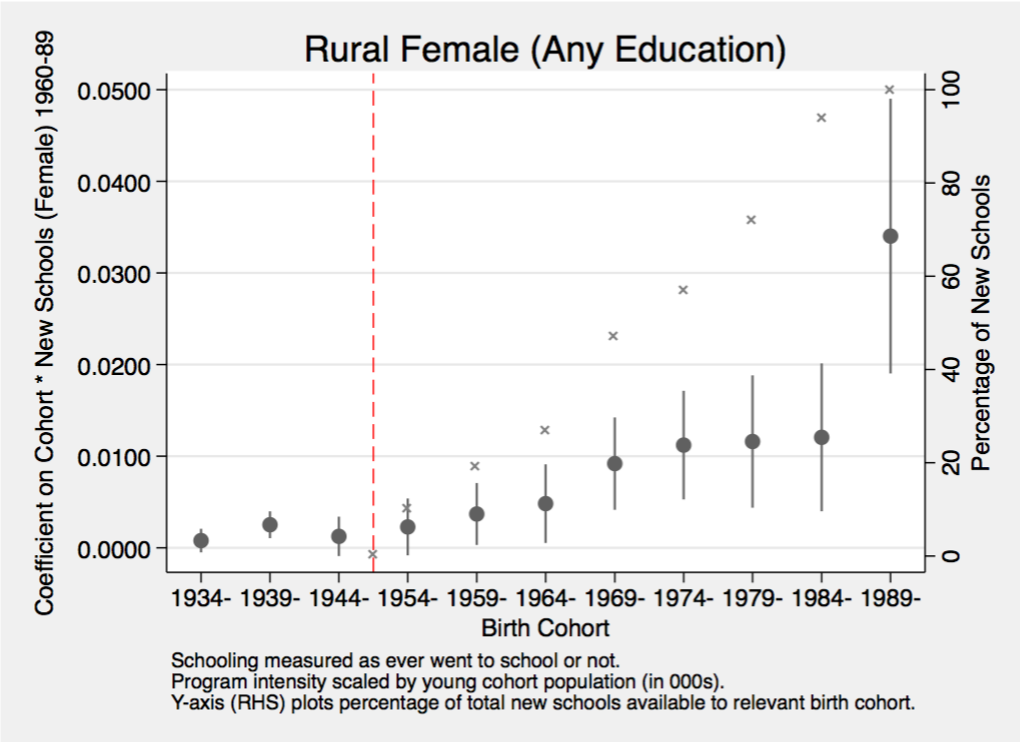
Table 6: **Effect on Intergenerational Education**

	Any Education (Children)			Primary Education (Children)		
Dependent variable	Full Sample	Male Children	Female Children	Full Sample	Male Children	Female Children
New Schools Available per 1000 kids (Mother)	0.0137	0.0113	0.0267	0.0109	0.00628	0.0192
	(0.00883)	(0.00519)**	(0.0114)**	(0.00792)	(0.00644)	(0.00890)**
Birth Cohort Fixed Effect (Mother)	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year Fixed Effect (Child)	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	471366	250533	220833	471366	250533	220833

Note: Sample size is children aged 10-17 using Population Census 1998. Effects calculated using Population Census 1998. Child birth year fixed effects included so that the marginal benefit to children's years of schooling is estimated across different households but among children of the same age. Robust Standard errors are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

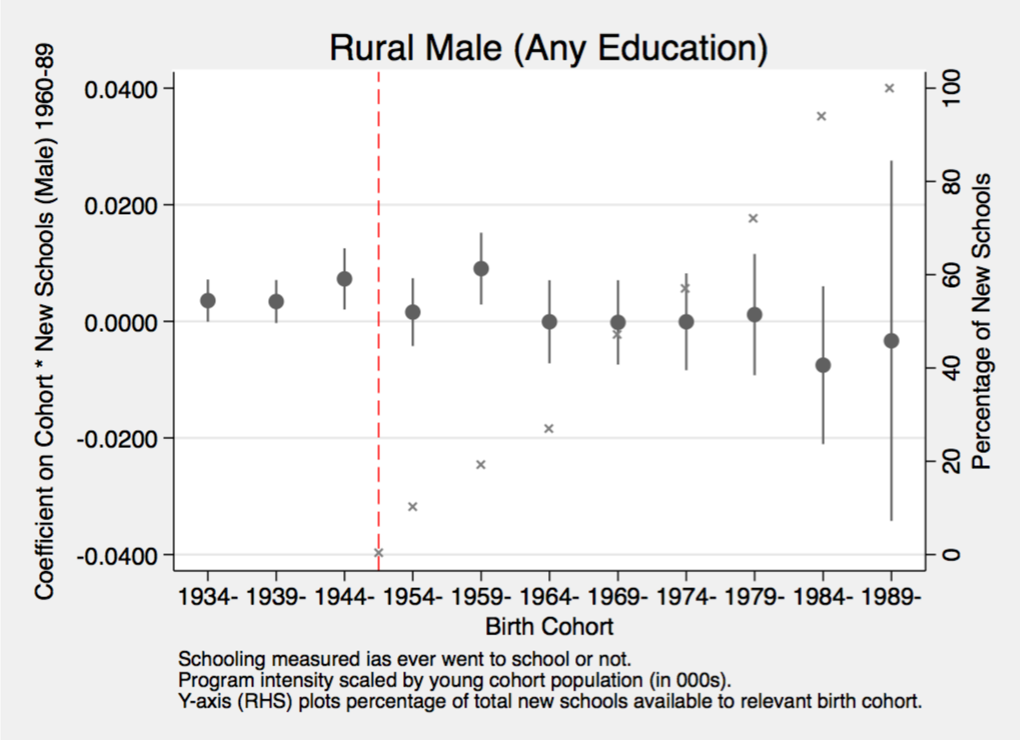
Appendix

Figure 1a: **Coefficient of Schools of Own Gender on Any Education (Rural Females)**



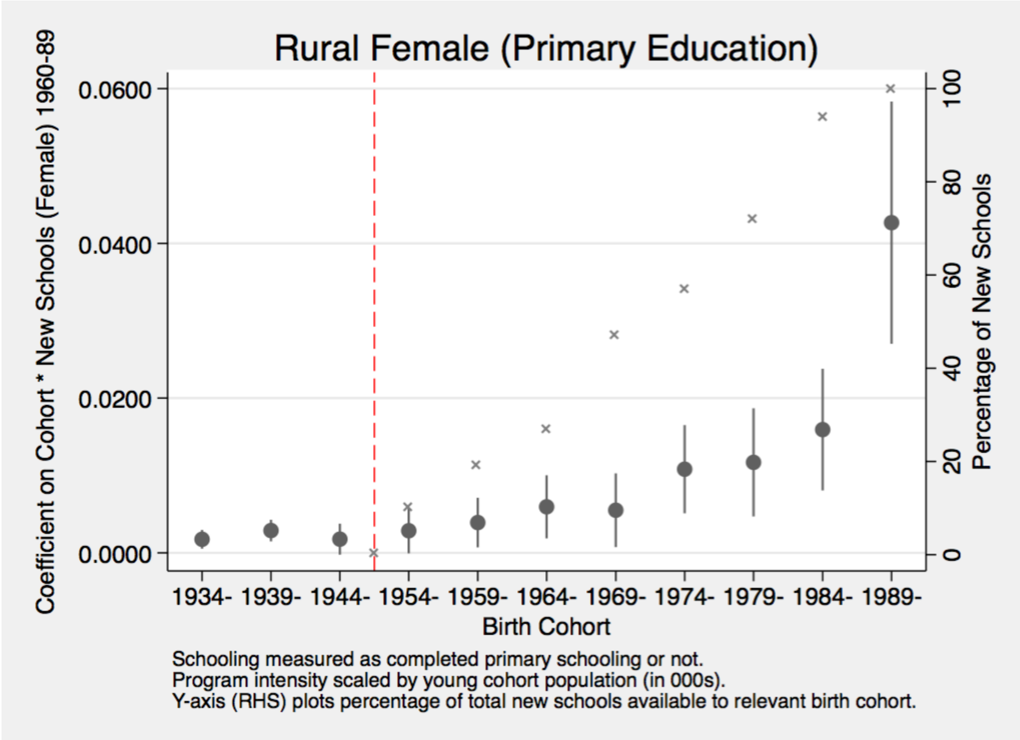
Note: Estimates plotted correspond to equation 4 for Rural Females with Any Schooling measure. Any Schooling measured as a binary variable equal to 1 if received at least 1 year of education. Birth cohorts split into 5-year bins. RHS axis plots the percentage of new female schools that have been constructed that are available to the relevant birth cohort. Administrative data on total new schools in the entire construction period from Punjab EMIS scaled by population of primary-school aged children (in 000s) using Population Census 1998. Data on educational attainment from DHS. Solid lines represent 95% Confidence Intervals.

Figure 1b: **Coefficient of Schools of Own Gender on Any Education (Rural Males)**



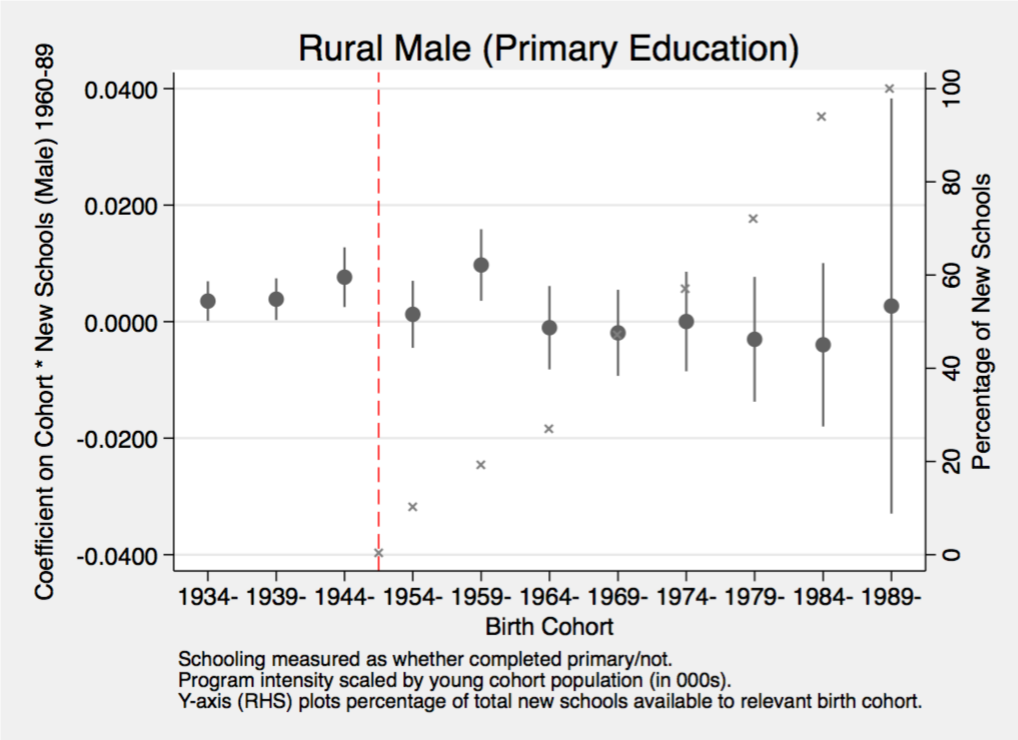
Note: Estimates plotted correspond to equation 4 for Rural Males with Any Schooling measure. Any Schooling measured as a binary variable equal to 1 if received at least 1 year of education. Birth cohorts split into 5-year bins. RHS axis plots the percentage of new male schools that have been constructed that are available to the relevant birth cohort. Administrative data on total new schools in the entire construction period from Punjab EMIS scaled by population of primary-school aged children (in 000s) using Population Census 1998. Data on educational attainment from DHS. Solid lines represent 95% Confidence Intervals.

Figure 1c: **Coefficient of Schools of Own Gender on Primary Education (Rural Females)**



Note: Estimates plotted correspond to equation 4 for Rural Females with Primary Schooling measure. Primary Schooling measured as a binary variable equal to 1 if completed primary education (5 years of education). Birth cohorts split into 5-year bins. RHS axis plots the percentage of new female schools that have been constructed that are available to the relevant birth cohort. Administrative data on total new schools in the entire construction period from Punjab EMIS scaled by population of primary-school aged children (in 000s) using Population Census 1998. Data on educational attainment from DHS. Solid lines represent 95% Confidence Intervals.

Figure 1d: **Coefficient of Schools of Own Gender on Primary Education (Rural Males)**



Note: Estimates plotted correspond to equation 4 for Rural Males with Primary Schooling measure. Primary Schooling measured as a binary variable equal to 1 if completed primary education (5 years of education). Birth cohorts split into 5-year bins. RHS axis plots the percentage of new male schools that have been constructed that are available to the relevant birth cohort. Administrative data on total new schools in the entire construction period from Punjab EMIS scaled by population of primary-school aged children (in 000s) using Population Census 1998. Data on educational attainment from DHS. Solid lines represent 95% Confidence Intervals.

Table 1: **Placebo effects of School of Opposite Gender on outcomes for Rural Cohorts**

	Effect of Male Schools on Rural Females			**Effect of Female Schools on Rural Males**		
Dependent variable	Any Education	Primary Education	Years of Education	Any Education	Primary Education	Years of Education
New Schools Available per 1000 kids (Other Gender)	-0.00422	-0.00673	-0.108	0.00344	0.0127	0.0994
	(0.0250)	(0.0249)	(0.238)	(0.0139)	(0.0141)	(0.126)
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52180	52180	52180	45266	45266	45266

Note: Estimates correspond to equation 5 in sub-section C of Section VI (Robustness Checks). Columns 1-3 analyze the effect of male schools on educational outcomes for rural females whereas Column 4-6 analyze the effect of female schools on educational outcomes for rural males. School construction measures constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1998 population census. Education measures are obtained from multiple waves of the DHS data. Any education and Primary education are binary variables equal to 1 if an individual received any education or if they completed primary school (5 years of education) respectively. Years of education measures single years of education. Sample of individuals aged 18 or above used in the calculations. Robust Standard errors are clustered at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: **Effect on Intergenerational Years of Education (Using DHS/MICS)**

	Using DHS			**Using MICS**		
Dependent variable	Full Sample	Male Children	Female Children	Full Sample	Male Children	Female Children
New Schools Available per 1000 kids (Mother)	0.606	0.360	0.834	0.298	0.276	0.326
	(0.353)*	(0.294)	(0.427)*	(0.0951)***	(0.145)*	(0.145)**
Birth Cohort Fixed Effect (Mother)	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year Fixed Effect (Child)	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8527	4532	3995	47058	24365	22693

Note: Estimates correspond to equation 2 in section on Empirical Specification. Columns 1-3 analyze the effect of mother exposure to school construction on children outcomes using DHS. Column 4-6 analyze the effect of mother exposure to new schools per capita using MICS. School construction measures constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1998 population census. Years of education measures single years of education. Sample of individuals who are living with parents in the household and not older than 16 are used in these calculations. Robust Standard errors are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: **Effect on Intergenerational Height (Using DHS/MICS)**

Dependent variable	Height for Age Percentile (NCHS) for Kids Age Below 5						Height for Age Z-score (WHO) for Kids Age Below 5					
	Full Sample		**Male Sample**		**Female Sample**		**Full Sample**		**Male Sample**		**Female Sample**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
New Schools Available per 1000 kids (Mother)	-0.351	-0.584	-0.390	-0.823	-0.329	-0.365	-0.0207	-0.0272	-0.0228	-0.0482	-0.0194	-0.00594
	(0.123)***	(0.258)**	(0.137)***	(0.338)**	(0.254)	(0.436)	(0.00771)**	(0.0121)**	(0.00795)***	(0.0191)**	(0.0160)	(0.0227)
Birth Cohort Fixed Effect (Mother)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year Fixed Effect (Child)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Time Trend $\times District$	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	38594	38594	19713	19713	18881	18881	38143	38143	19453	19453	18690	18690

Note: Children aged 0-5 born to mothers exposed to school construction are used in these calculations. Height for age percentile and Height for age scores are according to National Center for Health Statistics (NCHS) and World Health Organization (WHO) standards (as measured in MICS). School construction measures constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1998 population census. Robust Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: **Effect on Intergenerational Weight (Using DHS/MICS)**

	Low Birth Weight		Weight for Age Percentile (NCHS) for Kids Age Below 5						Weight for Age Z-score (WHO) for Kids Age Below 5					
Dependent variable	Full Sample		Full Sample		Male Sample		Female Sample		Full Sample		Male Sample		Female Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
New Schools Available per 1000 kids (Mother)	0.00805	0.0150	-0.131	-0.581	-0.0417	-0.658	-0.232	-0.497	-0.00712	-0.0237	-0.0000968	-0.0363	-0.0149	-0.0106
	(0.00316)**	(0.00606)**	(0.109)	(0.163)***	(0.145)	(0.231)***	(0.168)	(0.237)**	(0.00779)	(0.0106)**	(0.00956)	(0.0192)*	(0.0114)	(0.0151)
Birth Cohort Fixed Effect (Mother)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year Fixed Effect (Child)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Time Trend $\times District$	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15432	15432	38722	38722	19782	19782	18940	18940	38621	38621	19729	19729	18892	18892

Note: Children aged 0-5 born to mothers exposed to school construction are used in these calculations. Weight for age percentile and Weight for age scores are according to National Center for Health Statistics (NCHS) and World Health Organization (WHO) standards (as measured in MICS). Low birth weight is a binary variable equal to 1 if child was deemed to have below average weight at birth. Columns 1-2 use data from DHS whereas Columns 3-14 use data from MICS. School construction measures constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1998 population census. Robust Standard errors are clustered at the mother district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: **Effects on educational attainment for Rural Females Using district of Birth**

	Any Education		Primary Education	
Dependent variable	Using Residence District	Using Birth District	Using Residence District	Using Birth District
New Schools Available per 1000 kids (Own Gender)	0.0147	0.0150	0.0215	0.0248
	(0.00827)*	(0.00806)*	(0.0117)*	(0.0110)**
District Fixed Effects	Yes	Yes	Yes	Yes
Year of Birth Fixed Effects	Yes	Yes	Yes	Yes
Mean (All years)	0.269	0.264	0.199	0.194
Observations	131616	116333	88973	74841

Note: This table uses data from the 1973 Population Census for which data on district of birth and district of residence is both available. Columns 1-2 analyze impacts on any education measures whereas Columns 3-4 analyze impacts on primary education. Any Education and Primary Education are defined as binary variables equal to 1 if the individual received any education and if they completed primary education respectively. School construction measures constructed at the district level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1973 population census. Robust Standard errors are clustered at mother's district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: **Effect of School construction on Education attainment of Rural Cohorts Using Tehsil (sub-district) level variation**

	Any Education			Primary Education			Years of Education		
Dependent variable	Full Sample	Male	Female	Full Sample	Male	Female	Full Sample	Male	Female
New Schools Available per 1000 kids (Own Gender)	0.0186	0.00281	0.0324	0.0175	-0.00159	0.0309	0.169	0.0125	0.252
	(0.00444)***	(0.00424)	(0.00826)***	(0.00413)***	(0.00260)	(0.00829)***	(0.0378)***	(0.0270)	(0.0811)***
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tehsil Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	278673	140343	138330	278135	139970	138165	278135	139970	138165

Note: This table uses data from the 2008 and 2011 waves of Multiple Indicator Cluster Survey which are representative at the Tehsil (sub-district) level. Any Education and Primary Education are defined as binary variables equal to 1 if the individual received any education and if they completed primary education respectively. Years of Education measures single years of education. School construction measures constructed at the tehsil level using administrative data from Punjab Education Management Information System (EMIS). In order to convert schooling measure in per capita term, I use population of school-going children from the 1998 population census. Robust Standard errors are clustered at the tehsil level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$