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

Salman Khan[1]

October 2020

Abstract

In this paper, I study the long-term and inter-generational effects of expanding educational opportunities for girls through school construction in rural Punjab, Pakistan. Punjab is a setting characterized by low income and educational levels along with significant gender disparities in educational attainment. 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 but do not find corresponding impacts on female labor or marriage market outcomes. My findings indicate that the downstream effects of school construction and educational attainment for girls may be constrained in the absence of corresponding changes to male educational attainment and social norms, particularly in settings with significant gender inequality.

[1]Department of Economics, University of Illinois at Chicago. Contact: skhan275@uic.edu. I am grateful for helpful comments from Benjamin Feigenberg, Darren Lubotsky, Javaeria Qureshi, and Bhash Mazumder.

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 the Convention on the Rights of the Child establishing educational access as a basic human right (UNESCO, 2015; UNICEF, 2018). While significant progress has been made by many countries towards achieving Universal Primary Education (UPE), United Nations' figures reveal that the global number of girls out of school is actually increasing. Around 130 million girls- equivalent to the entire population of Mexico or half the population of Indonesia- are out of school (Evans et al., 2020). Existing research has shown that expanding educational opportunities for girls can help in increasing the enrollment rates for girls and lead to improvements in later life outcomes.[1] This research is based on the notion that education endows girls with knowledge, skills, and resources to make life choices that can improve their welfare (Duflo et al. (2012), Lundberg and Pollak (1993)). It is, however, important to note that most of this research has focused on returns to secondary education in settings with relatively limited gender disadvantage.[2] In much of the developing world, levels of female education remain quite low. The average woman aged 25 and over in 2000 reported 0.5 years in Niger, 2.6 years in Guatemala and 3.2 years of education in India and Kenya (Barro and Lee, 2001). Given

[1]This notion is perhaps best summarized by the Malala Fund's mission statement: "Secondary Education for girls can transform communities, countries and our world." (Malala Fund).

[2]For example cross-sectional studies on secondary education find that children of mothers with secondary education or higher are twice as likely to survive beyond age 5. In sub-Saharan Africa, an estimated 1.8 million children's lives could have been saved in 2008 if their mothers had at least a secondary education (UNESCO 2012).

the low levels of education that are typical of most developing countries, it is worth investigating whether efforts to promote primary schooling for girls offer similar promise.

In this paper, I study the short and long-term effects of a large scale school construction program in rural Punjab.[3] This program began in the 1960s in response to low educational levels. The overall literacy rate for population aged 10 or above was only 9.4 percent according to the 1961 population census (Akhtar, 1963). Literacy rates for girls were even lower (3.1% for females vs. 14.9 % for males) due to concerns about girls' safety and traditional customs that restrict girls' access to public spaces particularly outside their own community (Jacoby and Mansuri, 2011; Qureshi, 2018). These customs are enforced more rigidly in rural areas meaning that the majority of the girls face particular challenges in accessing education and healthcare (Jacoby and Mansuri, 2011). High gender disparities in this setting therefore suggest the possibility of multi-faceted benefits to increasing educational attainment for girls.

In order to analyze the impact of school construction on educational attainment and later life outcomes, 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 schools. I do so using administrative data on school construction from the Punjab Education Management Information System (EMIS). I combine this data with survey data from Demographic Health Surveys (DHS), as well as 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 cohort-level variation in exposure to

[3]Punjab is the largest of the four provinces in Pakistan and is home to 56% 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 rural areas for this study. See Section on Institutional Background for more details.

new schools to build on the Difference-in-Differences approach similar to Duflo (2001) and adapt it for the local context.[4] Since schools are segregated by gender, I construct measures for new school availability separately by gender at the district level.

In terms of my main results, I find that the construction of one additional primary girls school per 1000 girls of primary school going age at the district level, leads to a 5 percentage point increase in the likelihood of girls completing primary education and increases their years of education by 0.48 years. Given that the mean years of education for the relevant cohorts is only around 2.4 years, these effects are quite large in the context being studied and explain around 60% of the overall growth in schooling for girls. I do not find statistically significant effects of schools for boys on their educational attainment. These findings are consistent with improved access to schooling being the underlying driver of increased educational attainment as social mobility restrictions are more relevant for females in this (rural) setting.

Since the ultimate goal of education is to produce improvements in later life outcomes, I also analyze if school construction leads to changes in later life outcomes for rural females. Studying longer term effects of education for girls in developing countries (with significant gender disparity) is crucial but existing evidence is relatively limited.[5] In this spirit, I not only look at educational outcomes but also look at labor force participation, age at marriage, spousal characteristics, and fertility outcomes. In order to better capture the effects on labor and the marriage market, I use the Labor Force Surveys (LFS) to ana-

[4]School construction in this context followed a staggered approach starting in 1960s and continuing until 1988. Please see sections on Institutional Background and Empirical Specification for more details on school construction and empirical methodology used.

[5]For example, McEwan (2015) in a meta-analysis of more than hundred primary school interventions in developing countries find that only 10 percent have a evaluation component longer than one month after the intervention.

lyze effects on the labor force participation and use the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), to look at effects on the marriage market and inter-generational outcomes.[6] I find no significant effects on age at marriage, age at first birth, fertility rate, or child mortality and somewhat negative effects on female labor force participation. However, I do find evidence of inter-generational transmission of education driven largely by daughters of mothers more exposed to new schools.

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 (falsification) 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 am able to use the results to support the assertion 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.

In this paper, I aim to contribute to the literature in three respects. Firstly, I use quasi-experimental variation to study the causal effects of expanding primary education opportunity for girls on their educational attainment and later life outcomes in a setting

[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.

where baseline education levels are low. Studying these effects in this context is likely to be informative for other developing countries with similarly low levels of education for girls. This work therefore complements existing work on returns to secondary education by analyzing if the well documented benefits of schooling extend to low levels of education. Secondly, I study short and long term effects of school construction for Pakistan, a low-income country with significant baseline gender disparities in educational attainment (that persist to the present day) in educational attainment. Much of the existing work looking at long-term effects of school construction has focused on Indonesia following the seminal study on the INPRES program in Indonesia by Duflo (2001).[7] However, it is worth noting that the socio-economic[8] and cultural setup[9] in this context differ from Indonesia's (Hamid, 2016). There are also significant differences in baseline educational attainment[10] as well as in educational attainment by gender between the two settings.[11] Given the socio-economic and cultural differences between the two settings, it is worth analyzing the short and long term effects of school construction in Pakistan to shed light on the generalizability of past findings. Lastly, I contribute to the literature on investments in girls education by examining the effects of education on various longer term outcomes for females when there are no corresponding changes to the male education distribution. The heterogeneous treatment effects of school construction by gender in this context are different from other contexts such as Indonesia where both male and female outcomes re-

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

[8] Whereas Indonesia is a middle income country (ranked 7th in the world in terms of total GDP at PPP), Pakistan is a low-income country (it ranks 132nd in the world in terms of total GDP at PPP).

[9] For example, Islamist parties who rigidly advocate against female access to public spaces are not as strong in Indonesia compared to Pakistan (Hamid, 2016).

[10] Whereas the mean years of education for birth cohorts born around 1960 was 7.7 for Indonesia, the corresponding number is much lower at 2.7 for Pakistan.

[11] For birth cohorts around 1960, the mean years of education for males and females are 4.3 and 1.1 for Pakistan compared to 8.0 and 7.1 years of education for males and females in Indonesia.

spond to school construction (Duflo, 2001, Akresh et al., 2018, Mazumder et al., 2019a). 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 mother and children). This finding is likely to be of interest to policy makers in developing countries who are interested in returns to girls' education. My findings also support recent research (Evans and Yuan, 2019, Evans et al., 2020) that finds that in order to truly reap the benefits of increased education attainment, developing countries may need to ensure improvements to male education (and social norms) in addition to investing in higher levels of education for girls.

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 that is home to 56 percent of Pakistan's population and is sub-divided into 36 administrative units (Districts) and 150 sub-district units (Tehsils).[12] Pakistan has had a history of low educational attainment and Punjab is no exception to this. The overall

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

literacy rate for population aged 10 or above 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 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.[13] 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 upon the recommendations of the 1959 National Education Commission in light of the poor state of literacy in the country, the plan allocated Rs 78 million for the opening of 15,200 new primary schools (Commission, 1961).[14] 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(a) 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(a) 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.

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

[14] Roughly equivalent to $ 140 Million at purchasing power parity in 2020.

There was considerable heterogeneity across districts in the number of new schools built as the government designed the program to target districts in which enrollment was initially lower. Figures 2 (a)-(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' school. School construction increased the stock of primary schools in rural Punjab from a baseline of approximately 1100 primary schools for girls' and 4400 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(a), 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. Figure 1(b) plots the total number of new schools per 1000 kids of school going age for the 1960-1989 period. It shows that whereas the construction of girls' schools has led to a monotonic increase in the number of schools per capita, 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 0.6 schools per 1000 girls (for the 1949-53 cohort), the number rose to around 5 schools for the 1984-88 birth cohort (See Table 2). Similarly, for boys schools this number went from around 4 schools per capita at baseline to around 5.5 schools for the 1984-88 birth cohort.

Figure 3 plots the educational attainment of rural and urban cohorts over the 1960-1990 time period for the Punjab province. Figure 3 shows that rural females were the the most disadvantaged group with birth cohorts born around 1960 receiving less than 1 year of education on average. Primary school completion rates in the 1954 birth cohort before the start of school construction were around 10 percent among primary school-age girls (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 quite 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 1954-58 cohort points to 17 percentage points for 1984-88 cohort. However, education levels in absolute terms remained low relative to the world for both genders and there still persisted a relatively smaller gender gap 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 novel from the contexts where longer term effects of school construction have been analyzed. 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 schools using administrative data on schools in Punjab, Pakistan from Education Management Information System (EMIS). Administrative data from EMIS provides information on school's 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.[15] I validate this data on school construction using annual reports of Punjab Development Statistics which contain data on new schools constructed over time with their break-up at the District and Tehsil (sub-district) level.

I combine the administrative data on school construction in Punjab, Pakistan with household data from multiple waves of the Demographic Health Surveys (DHS) data (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 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

[15]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 sub-district (tehsil) level and find similar results. Please see sections on Empirical Specification and Robustness for more details.

health and child health. To analyze labor market outcomes, I use annual Labor Force Surveys (LFS) from 1990-91 to 2012-13 since the DHS does not include labor force participation questions for all members. LFS is therefore better suited than the DHS to capture labor market participation rates and has more detailed information on employment-related outcomes over time. For analyzing marriage market outcomes, I use both 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 inter-generational education and health-related outcomes as this survey is representative at the sub-district level for Punjab and has 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 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, marriage market 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

For my 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.[16] Given that the new schools constructed followed a staggered design, I build

[16] 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

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

I construct my measure of new schools at the district level.[17] This approach has two main advantages. Firstly, given that migration rates between birth district and current district of residence are quite low (around 5%), 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. 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 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, β_j refer to birth cohort and district

in appendix) for which I do have information on region of birth and the results are quantitatively similar.

[17] 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.

fixed effects respectively. 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 also study the long-term effects of exposure to the 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:

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

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*. $NewSchools_{jk}$ is defined as the total number of new schools (per 1000 kids of own gender) available to birth cohort j in region k at the time of their predicted age at entry to schooling. β_j, γ_k refer to birth cohort and district fixed effects respectively.

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.[18] It is worth

[18]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 whether employment is in the agriculture sector or not 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.

The lag between school construction that started in the 1960s and data collection allows me to study the effects of school construction on the next generation's outcomes. 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:

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

where Y_{icjk} refers to outcomes for child i, born to mother 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, the program increased the likelihood of 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). At the mean number of schools built per 1,000 children (2), these estimates imply an increase in years of schooling of almost 1 year for females. Given that the mean education for females in the sample is around 2.4 years, these effects are quite large. Moreover, new schools for girls can explain around 60% of the overall growth in schooling for girls in this context between the baseline cohort (1954-58) and the 1984-88 cohort.

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 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 women's and girl's access to public spaces are not relevant for boys in this (rural) setting.

Similar estimates on any education and primary education for girls suggest that girls who do end up enrolling tend to complete primary education. There are smaller but significant effects on secondary education and secondary school completion rates increase by 1.7 percentage points. Given that completion rates at secondary level are even lower than at 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, school construction through the INPRES program in Indonesia raised mean years of education by around 0.46 years for females and 0.53 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, 73% of females and 81% of the males complete primary schools in Indonesia and mean years of schooling were 7.1 and 8.0 for females and males respectively (Akresh et al., 2018). However, the average years of schooling for females and males in my sample are around 2.4 and 4.2 years and only 30% of the female sample and 52% 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.

Given that I only find effects on educational attainment for females in this context, I examine the effects of education on various longer term outcomes such as effects on marriage and labor markets as a function of new female schools only. This approach will allow me to shed light on whether long-term effects for females depend on male impacts given that there are no corresponding changes to male education as a function of school construction in this setting. 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, Table 4 studies the effect of school construction on labor market. Typically low or uneducated females work in the agriculture sector in the rural Punjab with overall labor participation rates being low 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 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 point 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 agriculture sector is (somewhat) consistent with Indonesian context where males move away from agriculture sector as a result of school construction (Akresh et al., 2018). However, we do not see significant increases in labor force participation for males and smaller but insignificant increases for females that we see in the Indonesian 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 for 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 Indonesia 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 small and statistically insignificant as in the Indonesian setting (Akresh et al., 2018). We do not find evidence that women match with spouses who have significantly higher levels of education and on average education differences between spouses go down. This is perhaps unsurprising given that the average education levels for males in regions with new females 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 so 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/family traditionally make marriage decisions for their offspring. 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 at this stage.

In short, I find that increasing education at the primary level may be insufficient to improve female participation in the labor market. Moreover increased schooling for females on average does not lead to better matches in the marriage market in terms of spousal education given the absence of corresponding changes to male education distribution.

C Intergenerational Impacts on Investments in Human Capital

Having observed effects of school construction program on a wide range of outcomes, including education, employment, and 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. Outcomes for children living in the parent's household are considered and age fixed effects are included

[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.

to ensure comparisons take place across children of the same age. By restricting analysis to children below age 16, I ensure that these children are unlikely to have left the parent household for work or marriage reasons. 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.

Table 6 (and Table 2 in appendix) show the reduced form results of mother exposure to school construction on 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 that there are indeed benefits of mother education that extend to the next generation- results consistent with (Andrabi et al., 2012) who find evidence of returns to even 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 the 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 got exposed more to new schools (See Appendix Tables 3-4). 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 on average has lower restrictions on women in the public space.

VI Threats to identification and robustness checks

In this section, I test for various threats to identification and provide robustness checks of my main specification.

A Pre-Trend test for Difference-in-Differences

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

I start by running a regression of the following form:

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

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 1960-89, scaled by population (in '000s) 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 (1) separately for males and females.

The assumption is that counterfactual trends should be uncorrelated with school construction and I test this by looking at the relationship between school construction and educational attainment of older cohorts. Figure 1 (a)-(d) in the appendix plot the coefficients obtained for rural male and female cohorts for regressions of completing any education and primary education on new schools of own gender. 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. Moreover, as more new schools get built, the more recent cohorts should benefit more from the total new school construction that happens between 1960-89 period. The effect size seems to get larger for more recent female cohorts whereas they remain insignificant for male cohorts. In short, the figures provide evidence that the new schools matter for rural female cohorts and that they are not correlated with trends in educational attainment of cohorts born prior to school construction. Therefore, they present evidence in support of school construction being a plausibly exogenous supply shock to access to schooling for the affected cohorts.

B Using an age-based Specification to check who benefits from School Construction

Since school construction follows a staggered design and using the fact that children in Pakistan typically attend primary school between ages 5-13, I also use an age-based specification to check if it is indeed the case that students of primary-school going age or younger are the ones who benefit from the school construction.

For this purpose, I use data from the DHS and the 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, $a_i \in \{0,20\}$

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

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. I use robust standard errors clustered at the district level.

This specification essentially 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 specification test that is more schools per 1000 kids should only matter for younger cohorts who are likely to benefit from school construction. Cohorts older than 12 are likely to be too old to benefit from more schools. This specification therefore serves as a test to validate my results for girls' education being driven by the younger cohorts who should benefit from the school construction.

Figure 4a and 4b 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 the age group younger than seven who benefit the most from the newly constructed schools. Girls in the 7-10 age group also benefit to a lesser extent as their 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.

C Placebo regressions to check the effect of school construction

One potential threat to my identification strategy is that effects on female education are driven by some omitted factor such as parental demand for education. 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.

Running these specifications for effect of male schools on female education and vice versa allows me to evaluate more credibly if any effects on female education are being driven by the supply of new schools and not some omitted factor such as parental demand for education.

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. Standard errors are clustered at the district of residence.

Table A.1 shows these results. I do not find any statistically significant effects of new male schools on female cohorts and vice versa. This implies that the effects of new schools are unlikely to be explained by demand side story or some omitted factor whereby motivated parents are able to request new schools.

D Using District of Birth instead of District of Residence

I also check if results could be driven by endogenous migration by focusing on a specification that uses birth district instead of the district of residence at the time of the survey and the results of the main specification hold (See Table 5 in Appendix). For this specification, I use data from an earlier Population Census (1973) since both birth and district of residence information is available for this census. I am also able to check migration rates between birth and current district and find that migration rates between districts are lower in Pakistan (around 5%) implying that endogenous migration is lesser of a concern in this context compared to Indonesia where district sizes are smaller and inter-district mobility is higher.

E Using Rural Males as unit of analysis to mitigate concerns related to endogenous migration of rural females

In order to mitigate concerns regarding endogenous migration by rural females as a response to school construction, I run specifications in which I use rural males as the unit of analysis. I run the following specification:

$$Y_{ijk} = \alpha + \delta NewSchools(Othergender)_{jk} + \beta_j + \gamma_k + \epsilon_{ijk} \tag{6}$$

where Y_{icjk} refers to educational outcomes for male individual's spouse i of birth cohort j, in district k. β_j and γ_k, refer to birth cohort, and district of residence fixed effects respectively. Standard errors are clustered at the male district of residence.

Since rural males in this setting typically live in the parent household even after marriage, this specification allows me to evaluate the impact of female school construction in male individual's district of residence on the spouse education for the male while controlling for district and birth cohort fixed effect. This specification yields similar effects to my main specification and further mitigates endogeneity concerns related to migration of rural females as a function of school construction.

F Additional Robustness Tests

Given that I construct my measure of school construction at the district level for my main specifications, there might be 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 level. Using variation in school construction at the Tehsil (sub-district) level, I analyze impacts of school construction on educational attainment (See table 6 in Appendix). I find significant effects of female school construction on education and no effects for male schools as in my primary specification.

Moreover, in order to rule out effects being driven by higher than primary level schools, I also run a specification in which I control for the effect of new secondary school. My results remain robust to inclusion of controls for new secondary schools and are therefore unlikely to be driven by them.

VII Conclusion

This paper studies the long-term and inter-generational effects of large school construction in a context with significant gender disparity. I use a difference-in-differences estimation strategy exploiting the variation across birth cohorts and regions in the number of schools built. I combine this with provincially representative data from Punjab, Pakistan that contains information on a wide range of outcomes related to education, employment, marriage outcomes, and inter-generational investments in health and education. I find that females exposed to the program attain more education, with no significant effects on male education. As adults, females exposed to school construction are less likely to work in the agricultural sector which helps explain part of the low labor force participation rates for females in this setting. Exposure to school construction alters marriage market outcomes for males with their spouses being more educated but we do not find evidence of similar effects for females since the male education distribution has not changed as a function of school construction.

The benefits of increased education are transmitted to the next generation if the mother is exposed to the program, with additional benefits accruing to daughters consistent with past work (Thomas, 1994; Akresh et al., 2018). Moreover, (lack of) results on the marriage market, fertility and inter-generational health are consistent with the notion that focusing only on female-targeted interventions may miss some of the most productive investments for improving educational and later life opportunities for girls in absolute terms (Evans and Yuan, 2019; Evans et al., 2020).

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.

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

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.

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.

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.

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.

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.

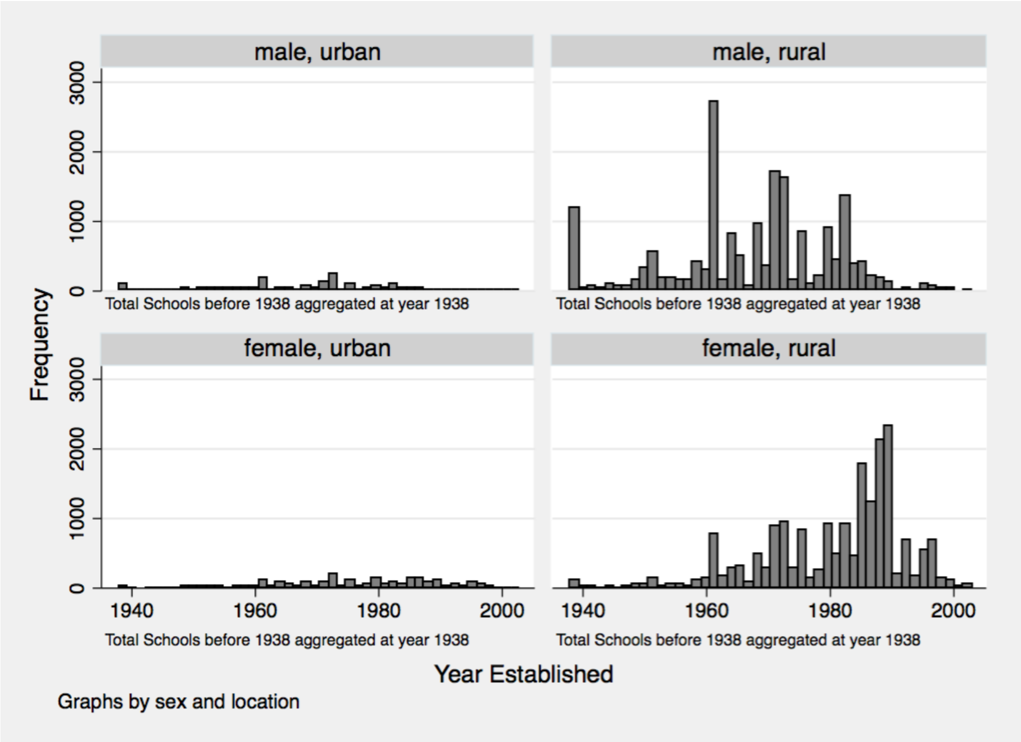
Watch, H. R. (2015). *World report 2015: Events of 2014*. Policy Press.

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.

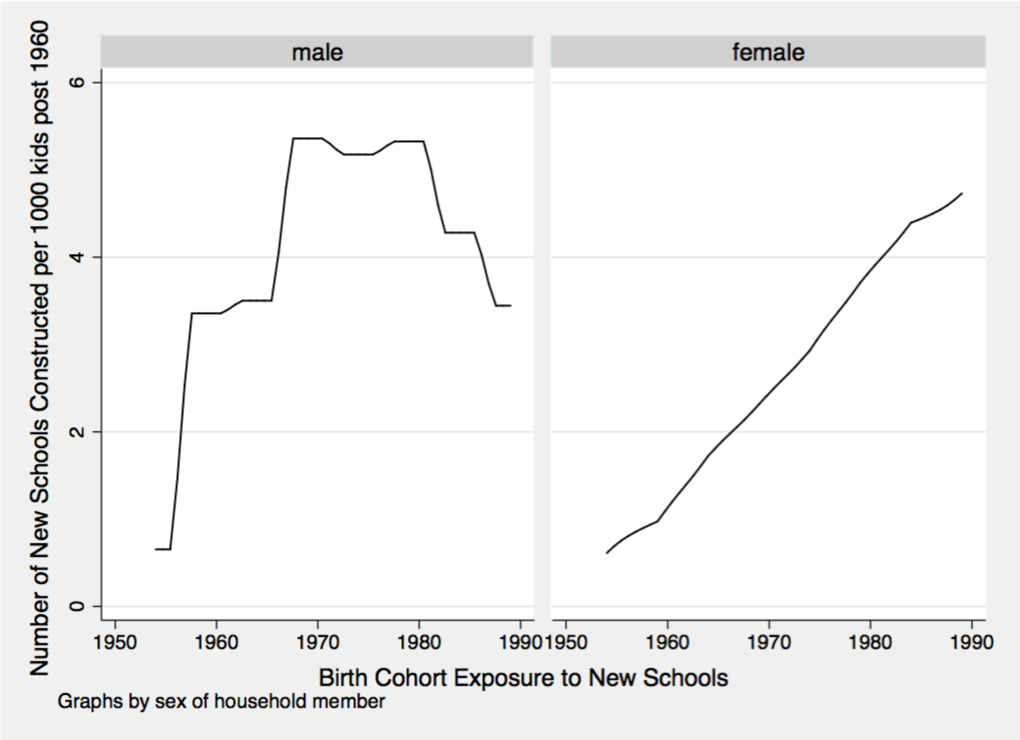
Figures

Figure 1a: **Number of New Schools by Gender and region for Punjab, Pakistan (1938-2000)**



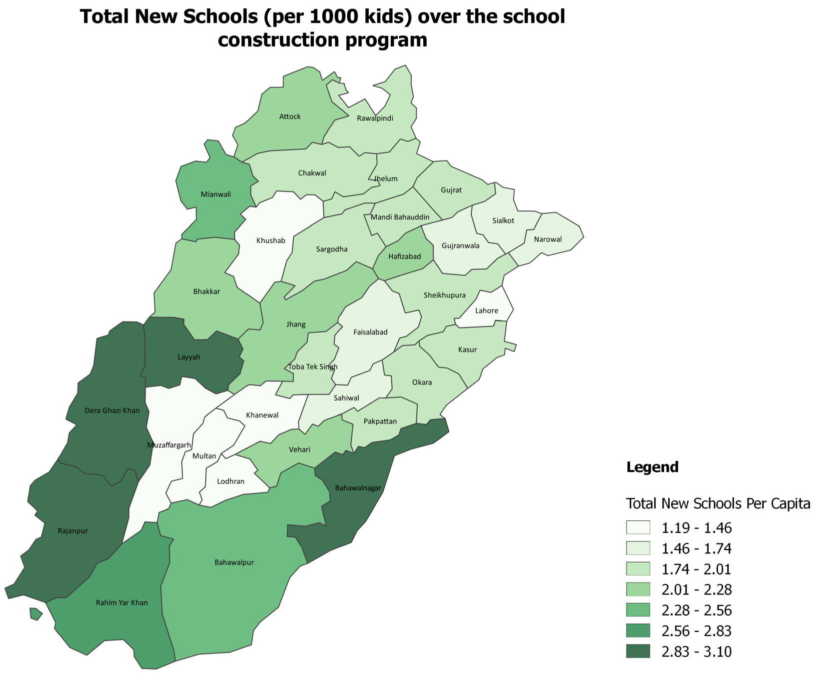
Note: Figure 1a plots the total number of new schools constructed between 1940-2000 for Punjab, Pakistan. Total Schools before 1940 are aggregated for at year 1938. Number of schools of own gender are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS).

Figure 1b: **Number of New Schools per 1000 children Over time (RURAL)**



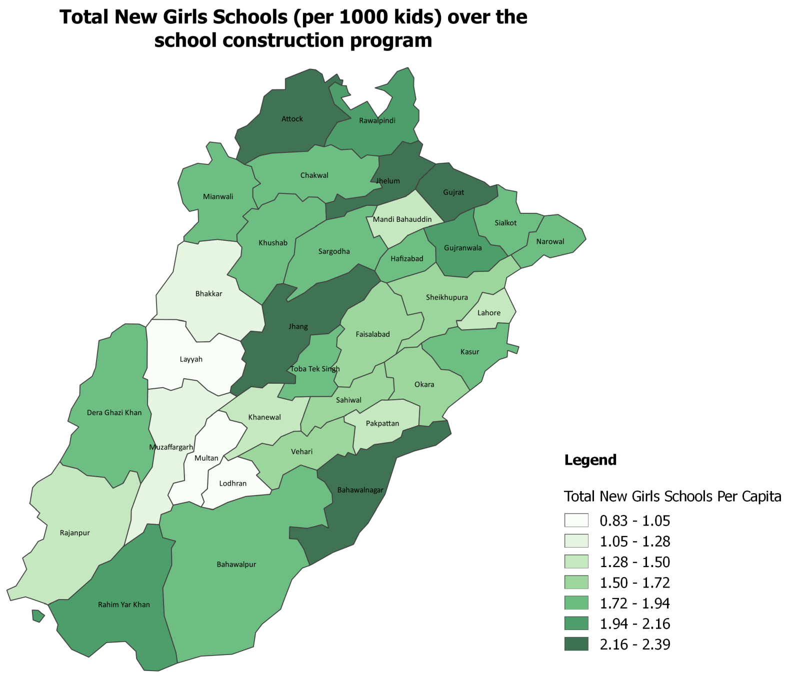
Note: Figure 1b plots the total number of new schools per 1000 children constructed during the 1960-1989 for Punjab, Pakistan. Number of schools of own gender are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per capita measure of school construction is obtained by dividing new schools at the district level by the population of primary school going children in thousands.

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



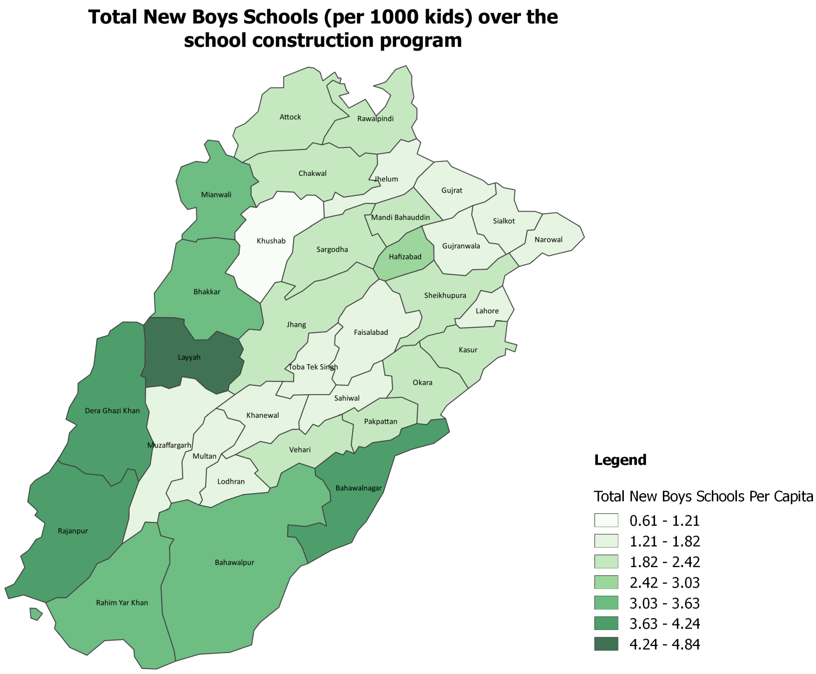
Note: Figure 2a plots the total number of new schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per capita 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**



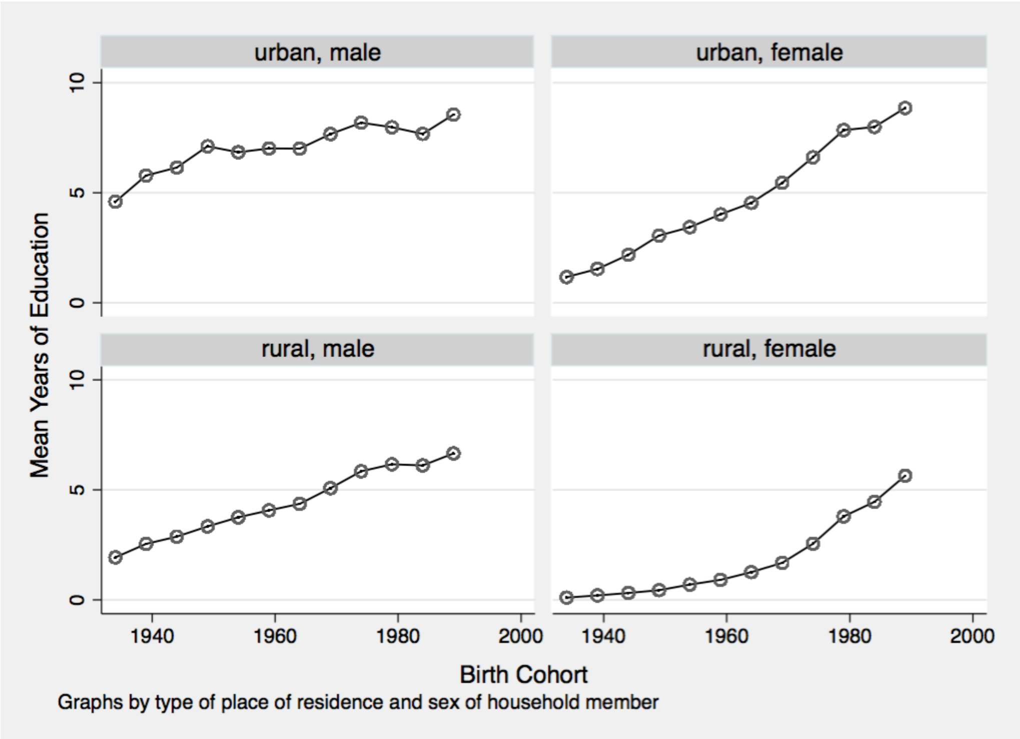
Note: Figure 2b plots the total number of new girls' schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per 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**



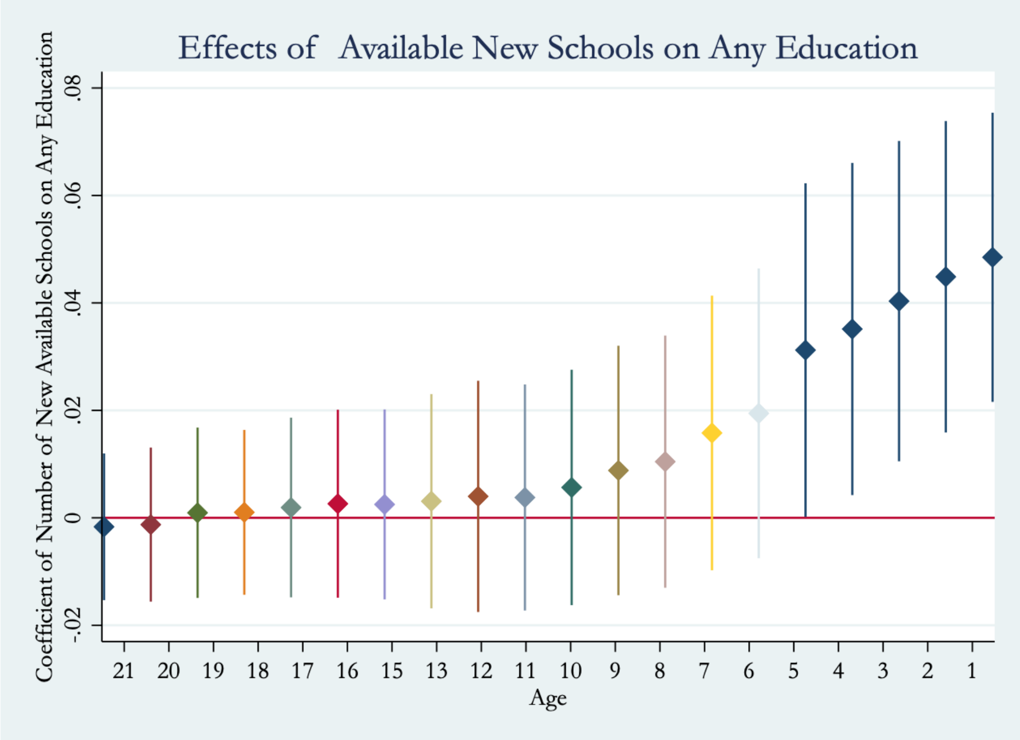
Note: Figure 2c plots the total number of new boy's schools per 1000 children constructed during the 1960-1989 for rural Punjab, Pakistan. Number of total schools at the district level are constructed using administrative data on schools from Punjab Education Management Information Systems (EMIS). Per 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)**



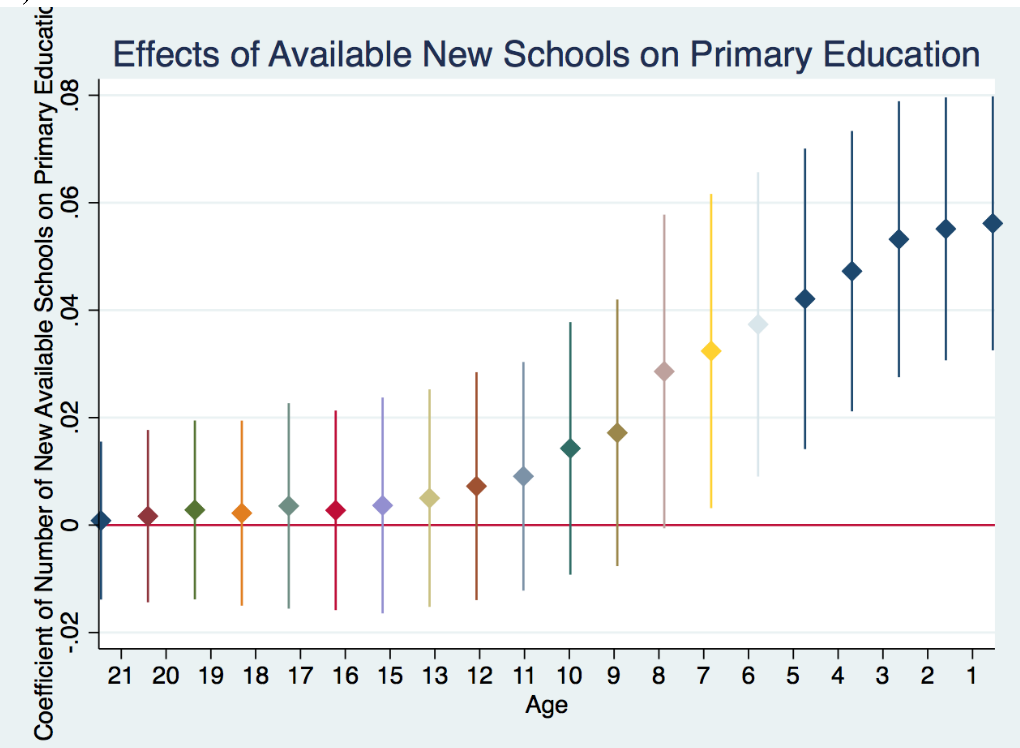
Note: Figure 3 plots the mean years of education for rural and urban cohorts by gender in Punjab, Pakistan over the 1934-1989 period. It uses data on educational attainment from multiple waves of DHS. Rural and Urban regions are defined as per the DHS definition.

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 capita by Gender over time for Rural Punjab, Pakistan**

Schools (percapita)	1934-38	1939-43	1944-48	1949-53	1954-58	1959-63	1964-68	1969-73	1974-78	1979-83	1984-88
Male	2.81	2.91	2.49	**3.87**	3.92	6.83	6.17	7.74	7.14	6.93	**5.49**
Female	0.21	0.28	0.32	**0.59**	0.70	1.81	1.96	3.00	3.10	3.77	**5.18**

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 capita measure, I divide schooling measure by total children 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 Demographic Health Survey

	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)	0.0434	0.00855	0.0485	0.0497	0.000435	0.0503	0.522	-0.0261	0.480
	(0.00924)***	(0.0113)	(0.0250)*	(0.00994)***	(0.0103)	(0.0255)*	(0.100)***	(0.108)	(0.267)*
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	92543	40363	52180	92543	40363	52180	92543	40363	52180

Panel B: Results using Population Census 1998

	Any Education			**Primary Education**			**Secondary Education**		
Dependent variable	Full Sample	Male	Female	Full Sample	Male	Female	Full Sample	Male	Female
New Schools Available per 1000 kids (Own Gender)	0.0550	0.00971	0.0495	0.0492	0.00571	0.0373	0.0105	0.00294	0.0172
	(0.00870)***	(0.0123)	(0.0229)**	(0.00670)***	(0.00913)	(0.0169)**	(0.00173)***	(0.00208)	(0.00582)***
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3424563	1749614	1674949	3147059	1605926	1541133	2394937	1210386	1184551

Note: Table 3 presents the results from Specification (1). 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. Clustered Robust Standard errors at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

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. Clustered Robust Standard errors at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

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
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
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
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. Clustered Robust Standard errors at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

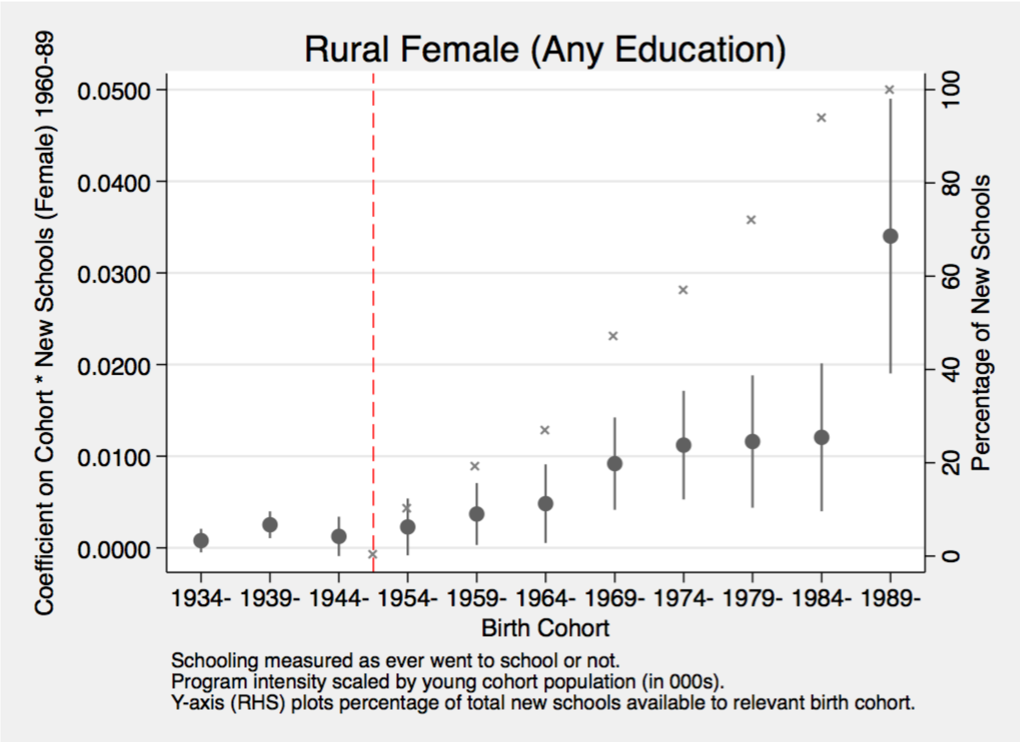
Table 6: **Effect on Inter-Generational Education**

	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)	0.0253	0.00551	0.0372	0.0170	0.00326	0.0253
	(0.00792)***	(0.00746)	(0.0145)**	(0.00843)*	(0.00999)	(0.0128)*
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	556271	396213	160058	556271	396213	160058

Note: Sample size is children aged 10-30 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. Since daughters move out of the household upon marriage, sample size for males is bigger. Clustered Robust Standard errors at the district level included. *** p<0.01, ** p<0.05, * p<0.1

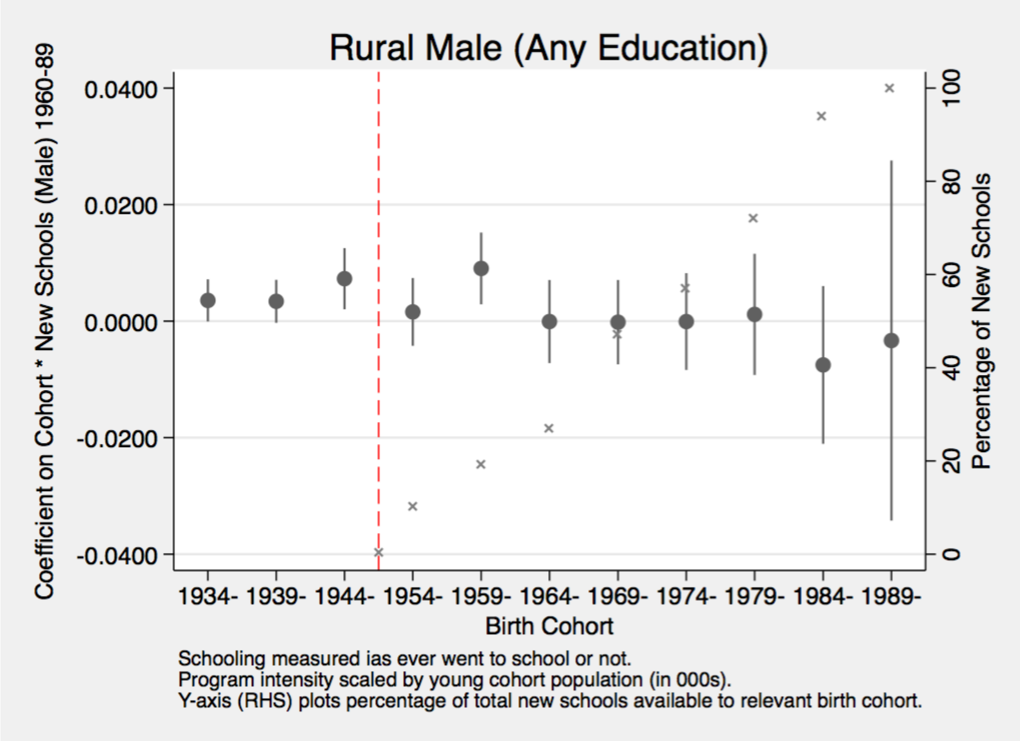
Appendix

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



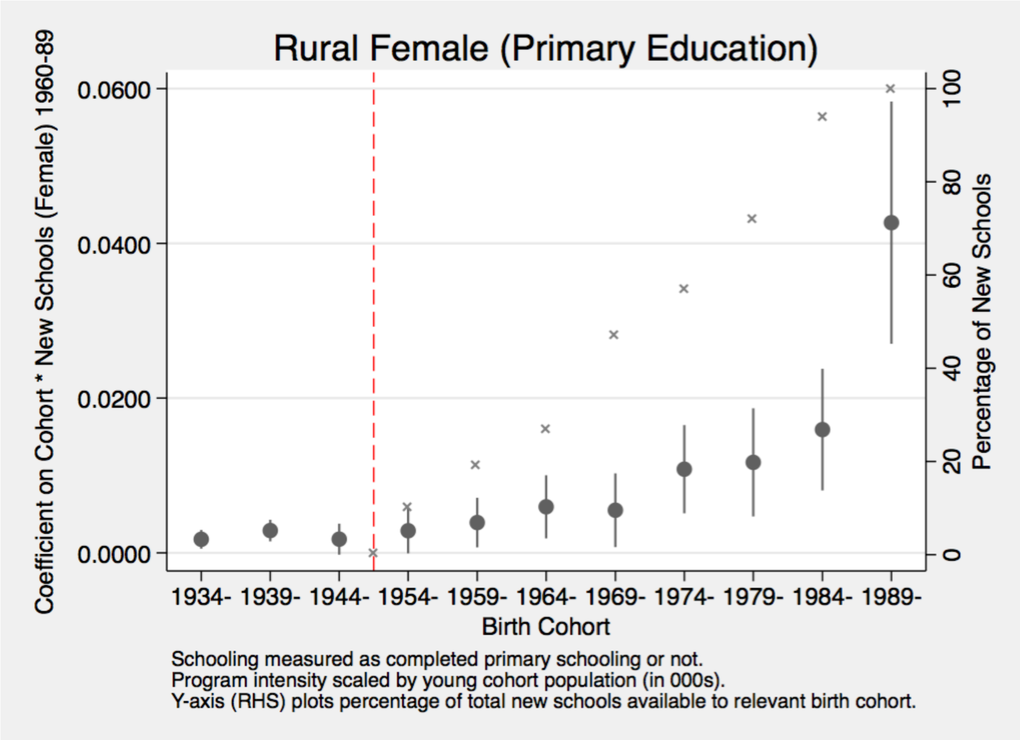
Note: Estimates plotted correspond to equation 4 for Rural Females with Any Schooling measure. Any Schooling measured as a binary variable equal to 1 if received at least 1 year of education. Birth cohorts split into 5-year bins. RHS axis plots the percentage of new female schools that have been constructed 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)**



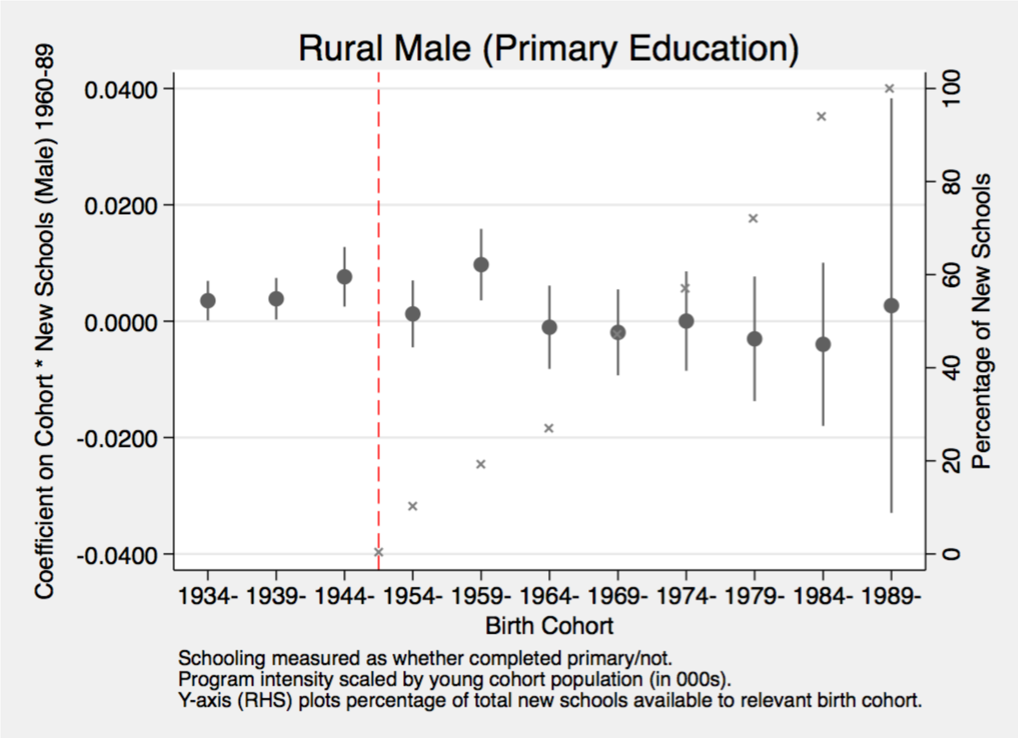
Note: Estimates plotted correspond to equation 4 for Rural Males with Any Schooling measure. Any Schooling measured as a binary variable equal to 1 if received at least 1 year of education. Birth cohorts split into 5-year bins. RHS axis plots the percentage of new male schools that have been constructed 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. Clustered Robust Standard errors at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: **Effect on Inter-Generational 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. Clustered Robust Standard errors at the district level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: **Effect on Inter-Generational 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. Clustered Robust Standard errors at the district level included. *** p<0.01, ** p<0.05, * p<0.1

Table 4: **Effect on Inter-Generational 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. Clustered Robust Standard errors at the mother district level included. *** 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. Clustered Robust Standard errors at the mother district level included. *** 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. Clustered Robust Standard errors at the tehsil level included. *** $p<0.01$, ** $p<0.05$, * $p<0.1$