

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Human Capital and Household Formation: Essays in Labor and Development Economics

Permalink

<https://escholarship.org/uc/item/5jm552xp>

ISBN

9798186382515

Author

Arora, Akanksha

Publication Date

2026-05-28

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Human Capital and Household Formation: Essays in Labor and Development Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Akanksha Arora

Committee in charge:

Professor Peter Kuhn, Co-Chair

Professor Kelly Bedard, Co-Chair

Professor Aashish Mehta

June 2026

The Dissertation of Akanksha Arora is approved.

Professor Aashish Mehta

Professor Kelly Bedard, Committee Co-Chair

Professor Peter Kuhn, Committee Co-Chair

May 2026

Human Capital and Household Formation:
Essays in Labor and Development Economics

Copyright © 2026

by

Akanksha Arora

For the teachers, friends, and family who shaped my education,
in ways large and small.

Acknowledgements

I am deeply grateful to my doctoral committee, Peter Kuhn, Kelly Bedard, and Aashish Mehta for their support, patience, and thoughtful feedback throughout this dissertation. I also thank Kelsey Jack and Shelly Lundberg for their guidance during earlier stages of my graduate training. Their insights substantially improved this work and informed the way I think about research and economics more broadly.

I am incredibly thankful for my friends from home, Kokil, Ganeev, and Akshay. Their encouragement, perspective, and good humour sustained me through the more difficult parts of graduate school. I also thank the friends I made at UCSB, especially Sonal, Kajari, and Reetika, for chai, conversations, and camaraderie.

Finally, I appreciate my family for the values they instilled in me. Many of the questions that motivate this dissertation, and much of how I understand the world, began with them.

Curriculum Vitæ

Akanksha Arora

Department of Economics
University of California, Santa Barbara
akanksha_arora@ucsb.edu
akanksha-arora.github.io

EDUCATION

2026	Ph.D. in Economics, University of California, Santa Barbara
2019	M.S. in Economics, University of Texas at Austin
2009	M.A. in Economics, Delhi School of Economics, University of Delhi
2007	B.A. (Honours) in Economics, Lady Shri Ram College, University of Delhi

WORKING PAPERS

1. “Educational Expansion and Assortative Matching: Evidence from the Marriage Market in India.” *Job market paper*.
2. “School Construction, Household Formation, and Intergenerational Human Capital in India.”
3. “Workfare, Marriage and Women’s Fertility: Evidence from India.”

TEACHING EXPERIENCE

Teaching Associate, University of California, Santa Barbara

2025	Global Economy and Development
------	--------------------------------

Teaching Assistant, University of California, Santa Barbara

2026	Introduction to Econometrics
2024–2025	Economic Development; Intermediate Macroeconomic Theory
2023–2024	Principles of Microeconomics; Intermediate Macroeconomic Theory; Economic Development
2022–2023	Economic Development; Principles of Microeconomics; Principles of Macroeconomics; Intermediate Microeconomic Theory

2021–2022	Intermediate Microeconomic Theory; Principles of Microeconomics; Intermediate Macroeconomic Theory
2020–2021	Economic Development; Intermediate Microeconomic Theory
2019–2020	Principles of Microeconomics

RESEARCH EXPERIENCE

University of California, Santa Barbara

2023	Research Assistant for Professor Kelsey Jack
------	--

University of Texas at Austin

2018–2019	Research Assistant for Professor Tom Vogl
-----------	---

Princeton University

2016–2017	Research Specialist for Dr. Jeremy Shapiro
-----------	--

J-PAL South Asia

2013–2015	Research Associate for Professors James Berry and Priya Mukherjee
2011–2012	Research Associate for Professors Abhijit Banerjee, Esther Duflo, Clément Imbert, and Rohini Pande

Indian Statistical Institute, Delhi

2010	Research Assistant for Professor E. Somanathan
------	--

Delhi School of Economics

2009	Research Assistant for Professor Rohini Somanathan
------	--

OTHER PROFESSIONAL EXPERIENCE

Evidence Action – Deworm the World

2012–2013	Data Coordinator
-----------	------------------

American Express

2010–2011	Business Analyst
-----------	------------------

ICICI Bank

2008	Analyst (Intern)
------	------------------

Abstract

Human Capital and Household Formation:
Essays in Labor and Development Economics

by

Akanksha Arora

My dissertation contains three chapters examining how changes in educational and labor market opportunities for women in India affect household formation, marriage, fertility, and outcomes for the next generation.

The first chapter “Educational Expansion and Assortative Matching: Evidence from the Marriage Market in India” studies how school construction reshaped educational attainment and marital outcomes over the long-run in India. Using staggered school construction across districts, I estimate the effects of increased school access on women’s education, marriage timing, and educational assortative matching between spouses. I find that greater access to schooling increased women’s years of education, delayed marriage, reduced child marriage, and narrowed spousal age gaps. Beyond individual outcomes, educational expansion also strengthened assortative matching, particularly among less educated groups, suggesting greater stratification at the lower end of the education distribution. However, I find limited evidence that increased schooling meaningfully changed women’s intra-household outcomes as measured by the incidence of domestic violence.

The second chapter “School Construction, Household Formation, and Intergenerational Human Capital in India” examines the effects of school construction on fertility and the formation of child human capital. Using the same staggered expansion in school access, I estimate the effects of increased schooling opportunities on women’s fertility outcomes and children’s education and health. I find that increased school access delayed

age at first birth and reduced completed fertility over the life cycle. Educational expansion also generated substantial gains for the next generation, increasing children's years of schooling, reducing the probability that children fall behind grade level, improving height-for-age, and increasing the likelihood of institutional delivery. While school construction increased parental education and changed the distribution of education across households, I find limited evidence that educational expansion substantially strengthened complementarities between parents in the production of child human capital.

The third chapter "Workfare, Marriage and Women's Fertility: Evidence from India" studies how increased employment opportunities for women affect marriage and fertility decisions. Using the staggered rollout of India's National Rural Employment Guarantee Act (NREGA), I estimate the effects of expanded work opportunities well-suited for women on marriage and childbearing. I find that the introduction of the program reduced marriage rates and shifted fertility toward later ages, with conceptions declining among younger women and increasing among older women. Though these are short-run impacts, the findings suggest that improvements in women's labor market opportunities and outside options can alter the timing of household formation and fertility behavior.

In summary, I find that large-scale expansions in women's educational and economic opportunities can generate persistent changes in marriage markets, fertility decisions, and household formation, with important implications for the distribution of human capital across households in developing economies.

Contents

Curriculum Vitae	vi
Abstract	viii
1 Educational Expansion and Assortative Matching	1
1.1 Introduction	3
1.2 Institutional Context on School Expansion and Marriage Markets in India	7
1.3 Conceptual Framework	9
1.4 Data	11
1.5 Empirical Strategy	14
1.6 Pre-trends analysis	17
1.7 Results	19
1.8 Conclusion	30
1.9 Figures and Tables	32
2 School Construction, Household Formation, and Intergenerational Human Capital	47
2.1 Introduction	49
2.2 Institutional Context on Educational Expansion and Intergenerational Outcomes in India	52
2.3 Conceptual Framework	54
2.4 Data	55
2.5 Empirical Strategy	58
2.6 Pre-trends Analysis	60
2.7 Results	61
2.8 Conclusion	71
2.9 Figures and Tables	72
3 Workfare, Marriage and Women’s Fertility	82
3.1 Background	84
3.2 Literature Review	86

3.3	Data	92
3.4	Empirical Strategy	93
3.5	Results	94
3.6	Figures and Tables	98
A	Appendix for Chapter 1	113
A.1	Data and Variable Construction	114
A.2	For Online Publication	117
B	Appendix for Chapter 2	124
B.1	Data and Variable Construction	125
B.2	Additional Tables	126

Chapter 1

Educational Expansion and Assortative Matching: Evidence from the Marriage Market in India

Abstract

Using staggered school construction across Indian districts, I estimate the effects of increased school access on women's educational attainment and marriage outcomes. An additional school per 1,000 school-age children, equivalent to about an 11 percent increase over the sample mean, raises women's years of schooling by nearly 2 percent and increases primary school completion by about 3 percent, with no significant effect on secondary education. Improved access to schooling also delays marriage, narrows the spousal age gap, and reduces child marriage: a comparable expansion raises the age at marriage by roughly 0.2 percent, decreases the spousal age gap by about 0.5 percent, and increases the probability of remaining unmarried by age 18 by approximately 0.5 percent. Beyond individual outcomes, school expansion reshaped marriage markets by strengthening educational sorting between spouses. Assortative matching increased among the least educated groups, while remaining comparatively stable or declining modestly at higher education levels, suggesting increased stratification at the lower end of the education distribution. Each additional school raises the assortative mating index by about 1.6 percent, and back-of-the-envelope calculations suggest that school expansion accounts for roughly 43 percent of the observed rise in assortative mating across marriage cohorts. However, I find little evidence that increased schooling improves women's intra-household outcomes: decision-making power remains unchanged, domestic violence exhibits strong intergenerational persistence, and there is no evidence of backlash driven by changes in relative spousal education. These findings indicate that large-scale investments in schooling improved women's human capital and delayed marriage while reshaping patterns of marital sorting, without substantially altering intra-household dynamics.

1.1 Introduction

Investments in education influence not only individual human capital, but also patterns of marital matching. By changing the distribution of educated men and women, educational expansion reshapes marriage markets, altering who marries whom and how inequality is transmitted across generations. These effects may be particularly important in settings where marriage is nearly universal and women's labor market opportunities remain limited. In such contexts, education generates returns not only through wages, but also through the marriage market. Understanding how educational expansion affects marital matching is therefore important not only for evaluating the private returns to schooling, but also for understanding broader processes of stratification and intergenerational mobility.

India provides a particularly useful setting in which to study these dynamics. Over the past several decades, India has experienced a dramatic expansion in schooling, especially for women, alongside persistently low female labor force participation. Figure 1.1 shows that women's educational attainment increased substantially across birth cohorts. The gender gap in schooling narrows sharply across cohorts, with women's average years of education converging toward those of their husbands among younger cohorts. At the same time, marriage remains nearly universal and continues to structure household formation, fertility, and intergenerational mobility. Moreover, because marriages in India remain heavily family-mediated¹, educational attainment functions as an especially salient matching characteristic within marriage markets. Figure 1.2 shows corresponding differences in marriage timing across education groups, including later marriage and lower rates of child marriage among more educated women. These simultaneous changes

¹Based on data from the India Human Development Survey, approximately 66 percent of ever-married women reported that their marriages were arranged primarily by parents or relatives, while an additional 29 percent reported that the match was chosen jointly by themselves and their families. Only about 5 percent reported choosing their spouse independently.

raise the possibility that educational expansion affected not only schooling outcomes directly, but also the equilibrium structure of marriage markets. As the distribution of education changes, educational sorting between spouses may strengthen or weaken, marriage timing may shift, and the distribution of household advantage may become more or less stratified. These changes raise a broader question of whether educational expansion reshapes marriage markets primarily through changes in the distribution of human capital, or whether it also alters patterns of matching and household formation more fundamentally.

A large literature studies the returns to education through labor market outcomes. However, education may also generate substantial returns through marriage markets. In particular, education may affect the probability and timing of marriage, partner quality, assortative matching, and intra-household allocations (Behrman et al., 1999; Peters, 2002; Desai and Andrist, 2010; Lafortune, 2013; Geruso and Royer, 2014; Ashraf et al., 2020; Anderberg et al., 2022; Agarwal, Bahure and Javadekar, 2023). Theoretical work in economics has long emphasized that educational sorting may arise through multiple channels. Individuals with similar educational attainment may derive gains from shared preferences and consumption complementarities (Becker, 1981; Lundberg and Pollak, 2012). Education may also increase household productivity through complementarities in investments in children and household production (Behrman and Rosenzweig, 2002; Chiappori, Costa Dias and Meghir, 2018). In addition, education may shape bargaining power and the distribution of surplus within households (Choo and Siow, 2006; Mourifié and Siow, 2017; Doorley, Dupuy and Weber, 2019). At the same time, rising educational attainment may mechanically increase educational homogamy even in the absence of stronger assortative preferences. Distinguishing between these forces is important because assortative matching has broader equilibrium implications for inequality and intergenerational persistence (Fernández, Guner and Knowles, 2005; Eika, Mogstad and

Zafar, 2019; Greenwood et al., 2014).

The paper also relates to recent work on school construction and marriage outcomes in India. Khanna (2023) studies the general equilibrium labor market effects of India's District Primary Education Program (DPEP), while Agarwal, Bahure and Javadekar (2023) examine the effects of school expansion on women's marriage outcomes using a regression discontinuity design. In contrast, I exploit continuous variation in school construction intensity across districts and cohorts over a substantially longer horizon and study not only marriage timing but also equilibrium changes in assortative matching and marriage-market stratification.

Building on this literature, this paper studies how large-scale school construction in India reshaped educational attainment, marriage timing, assortative matching, and intra-household outcomes. I combine administrative data on school construction with nationally representative Demographic and Health Survey (DHS) data and exploit variation across districts and birth cohorts in exposure to school construction during ages 6–15. The empirical design follows Duflo (2001), using district and cohort variation in school construction intensity over a substantially longer time horizon. The analysis examines both individual-level outcomes, such as schooling and age at marriage, and equilibrium marriage-market outcomes, including educational sorting and assortative mating.

I find that school construction substantially increased women's educational attainment, particularly at lower levels of schooling. An additional school per 1,000 school-age children, approximately an 11 percent increase relative to the sample mean, raises women's years of education by nearly 2 percent and significantly increases primary school completion, with little effect on secondary completion. Educational expansion also delayed marriage, reduced child marriage, and narrowed the age gap between spouses. Beyond these individual-level effects, I show that school construction reshaped the equilibrium structure of marriage markets. Educational sorting strengthened as school access

expanded, with each additional school increasing the assortative mating index by approximately 1.6 percent. Homogamy among the least educated increased substantially across marriage cohorts, even as average educational attainment rose. Back-of-the-envelope calculations suggest that school construction explains nearly 43 percent of the observed rise in educational assortative mating across marriage cohorts. These patterns imply that educational expansion may simultaneously increase average human capital accumulation while increasing educational segmentation among the least educated.

At the same time, I find little evidence that educational expansion substantially altered intra-household gender dynamics. Increased schooling does not meaningfully improve women's decision-making power or reduce domestic violence, and domestic violence remains strongly persistent across generations. Nor do I find evidence that increases in women's relative education generate backlash within marriage. Taken together, the results suggest that educational expansion reshaped who marries whom and when households form, while leaving deeply entrenched intra-household norms largely unchanged.

This paper contributes to several literatures. First, it contributes to work on the returns to education by showing that large-scale educational expansion affects not only labor market outcomes, but also marriage markets and household formation. While Duflo (2001) studies the labor market consequences of school construction in Indonesia, I examine how educational expansion reshapes marriage markets in a setting where women's labor force participation remains limited. Second, the paper contributes to the literature on assortative mating by providing new evidence from a developing-country setting undergoing rapid educational expansion. While prior work has documented rising assortative mating in high-income countries (Fernández, Guner and Knowles, 2005; Eika, Mogstad and Zafar, 2019), much less is known about how educational expansion reshapes sorting in marriage markets in low- and middle-income countries. Third, the paper contributes to a growing literature on educational expansion and inequality by

showing that schooling investments may increase stratification within marriage markets even while expanding access to education overall.

1.2 Institutional Context on School Expansion and Marriage Markets in India

School access expanded dramatically in India after Independence through a series of large-scale public investments in educational infrastructure. Beginning in the 1970s, the Indian government implemented successive programs aimed at universalizing elementary education and reducing geographic disparities in school access. The Minimum Needs Programme launched in the mid-1970s sought to ensure that every village had access to a primary school within a specified distance. Subsequent initiatives, including the District Primary Education Programme (DPEP) in the 1990s and the Sarva Shiksha Abhiyan (SSA) in the 2000s, substantially expanded school construction, teacher recruitment, and educational funding across districts. As a result of these policies, school availability increased sharply over time. Figure 1.3 shows that younger birth cohorts were exposed to substantially higher levels of school availability during their school-age years. Average exposure more than doubled between the 1965 and 1995 cohorts, generating substantial variation across cohorts in access to schooling. According to Census data, the share of villages with a primary school increased from approximately 53 percent in 1971 to over 70 percent by 1991, while access to middle and high schools also expanded substantially. By the 2010s, most rural households had access to a primary school within walking distance. In addition to public investments, India also experienced a large increase in low-cost private schooling during this period.

These educational expansions coincided with major changes in women's educational

attainment and marriage patterns. Figure 1.4 illustrates how the educational composition of women changed across cohorts over the same period. The share of women with no schooling declines sharply between the 1965 and 1995 birth cohorts, while the shares completing secondary and college education increase steadily over time. Average years of schooling also rose substantially across cohorts, accompanied by a narrowing gender gap in educational attainment. These descriptive patterns are consistent with educational expansion substantially altering the distribution of education within local marriage markets, potentially changing both the availability of similarly educated partners and the structure of marital matching.

Marriage continues to play a central role in social and economic life in India. Marriage is nearly universal, and most marriages remain arranged or family-mediated, with partner search often occurring within relatively narrow caste, linguistic, religious, and regional networks (Banerjee et al., 2013). Educational attainment has increasingly become an important dimension along which matches are formed. Unlike labor market outcomes—which remain highly gendered and are frequently observed only after marriage—education is observable prior to marriage, widely understood by families, and directly shaped by policy-driven investments in human capital. These features make education a particularly salient characteristic in the marriage market and a natural dimension along which assortative matching may occur.

To provide supporting evidence that education is an important attribute in partner selection, I construct a novel dataset of approximately 15,000 matrimonial advertisements published in a major Hindi-language newspaper between 2022 and 2025.² The advertisements reveal that education is among the most universally valued characteristics in partner search for both men and women. Approximately three-quarters of both male and female profiles explicitly mention education, indicating that schooling has become

²Appendix Table A.1 summarizes these patterns.

a near-universal characteristic in self-presentation within marriage markets. In contrast, labor market characteristics remain substantially more gendered. Nearly half of female profiles seeking grooms mention employment or occupation requirements, compared to fewer than one-fifth of male profiles seeking brides. These descriptive patterns suggest that education functions as a common dimension of matching for both genders, while labor market characteristics remain disproportionately important for men. This asymmetry is consistent with a setting in which educational expansion may reshape marital matching even in the presence of persistently low female labor force participation.

1.3 Conceptual Framework

Educational expansion may affect marriage markets through multiple channels. First, increased access to schooling raises educational attainment directly, altering the distribution of educated men and women within marriage markets. As the supply of educated individuals changes, patterns of matching may also change. If individuals prefer spouses with similar educational attainment, increases in schooling may strengthen educational assortative mating.

Positive assortative matching may arise for several reasons. Individuals with similar educational backgrounds may share preferences, lifestyles, or social environments, generating consumption complementarities within marriage (Becker, 1981; Lundberg and Pollak, 2012). Education may also increase household productivity through complementarities in investments in children, household decision-making, and specialization (Behrman and Rosenzweig, 2002; Chiappori, Costa Dias and Meghir, 2018). Alternatively, increases in educational homogamy may emerge mechanically as educational attainment rises across the population, even in the absence of stronger assortative preferences. Distinguishing between these mechanisms is important because they imply different conse-

quences for inequality and household formation (Almar et al., 2025).

Assortative matching has broader equilibrium implications beyond individual marriage outcomes. Stronger educational sorting can concentrate human capital within households and increase intergenerational persistence in socioeconomic status (Fernández, Guner and Knowles, 2005; Greenwood et al., 2014; Eika, Mogstad and Zafar, 2019). Conversely, educational expansion could weaken stratification if increased access broadens the set of feasible matches across educational groups. The distributional consequences therefore depend not only on whether schooling increases average attainment, but also on how it changes the structure of matching across households. Figure 1.5 illustrates these patterns by showing substantial increases in assortative matching among the least educated groups across marriage cohorts.

A key empirical challenge is distinguishing changes in assortative mating from changes in the marginal distribution of education. As educational attainment rises over time, educational homogamy may increase mechanically simply because more individuals attain similar levels of schooling. Following Eika, Mogstad and Zafar (2019), I therefore use a measure of assortative mating that accounts for changes in the marginal distribution of education and captures whether individuals with similar educational attainment marry each other more frequently than would occur under random matching.

Educational expansion may also affect the timing of marriage and intra-household dynamics. Increased schooling may delay marriage directly through enrollment or incapacitation effects if women remain in school longer. Education may additionally alter reservation quality in the marriage market, leading households to delay marriage in search of better matches. At the same time, increased education could improve women’s bargaining power within marriage by increasing outside options or productivity. However, if labor market opportunities remain constrained or gender norms remain persistent, educational expansion may reshape marriage markets without substantially changing

intra-household power relations. The empirical analysis examines each of these channels in turn.

1.4 Data

This paper combines administrative data on school construction with nationally representative household survey data and district-level population statistics to construct a long-run panel of educational exposure across Indian districts and cohorts.

Data on school construction and school locations come from the Unified District Information System for Education (UDISE), compiled by the Department of Education, Government of India. UDISE is the most comprehensive administrative database on schools in India and covers approximately 1.5 million public and private schools across the country. The database contains detailed school-level information including school location, management type, grade levels offered, infrastructure, enrollment, teacher characteristics, and year of establishment. For the purposes of this paper, I primarily use information on school location and year of establishment to construct district-level measures of cumulative school exposure over time. Although the UDISE data was collected in 2018 through school-level surveys, the database records establishment years for all schools that existed at the time of the survey, allowing the construction of a long-run historical panel of school expansion. I aggregate school establishment information to the district-year level and construct cumulative measures of schools per 1,000 school-age children. The treatment variable therefore captures the stock of schools available in a district during an individual's school-age years rather than contemporaneous school construction flows. One limitation of the UDISE data is that it only includes schools operating at the time of the 2018 survey. Schools that were constructed historically but shut down before the survey are not observed. This may generate measurement error in historical

school exposure, particularly for earlier cohorts whose exposure occurred further in the past. If school closures are unrelated to later marriage outcomes conditional on district and cohort fixed effects, the resulting measurement error would tend to attenuate estimated effects toward zero. However, because older schools are mechanically more likely to be missing, some non-classical measurement error remains possible. At the same time, available evidence suggests that the period under study was characterized primarily by rapid educational expansion rather than widespread school closures, particularly during earlier decades.

Information on education, marriage, and intra-household outcomes comes from the Indian Demographic and Health Survey (DHS), also known as the National Family Health Survey (NFHS) in India. The DHS is a nationally representative household survey that collects detailed demographic, educational, fertility, health, and domestic violence information. The primary analysis uses data from the women’s individual recode, which surveys women aged 15–49 and contains detailed information on educational attainment, age at marriage, fertility histories, household characteristics, and measures of women’s autonomy and domestic violence. A key advantage of the DHS is the availability of detailed retrospective marriage and fertility information, which allows outcomes to be linked to childhood exposure to school construction. Because spouse education variables in the women’s survey are available only for a restricted subsample of respondents, I supplement this data with information from the household roster. Husband characteristics, including years of schooling and age, are linked using the spouse line number reported in the women’s questionnaire, allowing the construction of measures of educational assortative matching and relative spousal education for co-resident couples. To examine intra-household outcomes, I additionally use the domestic violence module, which contains information on women’s participation in household decision-making, controlling behavior by husbands, and emotional, physical, and sexual violence. District identifiers

are only publicly available in the Indian DHS beginning with the 2015 round. As a result, the main analysis relies primarily on the 2019 survey round, which provides the most recent nationally representative data with district-level geographic identifiers. The analysis links women to district-level school construction exposure based on district of current residence. While the DHS does not observe district of childhood residence directly, migration during school-age years is relatively limited in India. Moreover, Census evidence suggests that most marriage migration occurs within relatively local geographic areas, implying that district of residence remains a reasonable proxy for childhood exposure for much of the population.

Population data are drawn from decadal Census of India population tables. Specifically, I use district-level population counts from the Census and decadal variation tables to construct measures of school-age population over time. Population figures are linearly interpolated across years within decades in order to construct annual district-level estimates. Importantly, the Census tables account for administrative boundary changes when estimating historical district populations, which is critical for constructing consistent exposure measures over time.³ A major challenge in conducting long-run district-level analysis in India is the large number of changes in district and state boundaries since the first post-Independence census in 1961. Over time, districts have been repeatedly split, merged, or reorganized, while new states have also been created. These changes complicate efforts to combine historical administrative data with modern survey datasets. A key contribution of this paper is the construction of a harmonized long-run district-level panel that reconciles these boundary changes across datasets and over time. Historical school construction data, population data, and DHS district identifiers are mapped into a common district framework based on harmonized 2011 district boundaries, allow-

³Additional details on the construction of the harmonized district panel are provided in Appendix A.1.1

ing consistent measurement of school exposure across cohorts and regions over multiple decades.

1.5 Empirical Strategy

The empirical analysis exploits variation across districts and birth cohorts in exposure to school construction during school-age years. School construction in India was governed largely by centralized population-and-distance norms under successive Five-Year Plans and later the Sarva Shiksha Abhiyan (SSA), with funding tied to predetermined infrastructure targets rather than contemporaneous local marriage-market conditions. Once constructed, schools represent long-lived infrastructure investments, making reverse causality from cohort-specific marriage outcomes implausible. In addition, treatment exposure is intended to capture school construction exposure during childhood, implying that variation in exposure is predetermined with respect to adulthood marriage outcomes.

The empirical design follows Duflo (2001) and exploits both spatial and temporal variation in school construction intensity. Specifically, I estimate regressions of the form:

$$Y_{idt} = \beta \text{SchoolExposure}_{dt} + \gamma_t + \delta_d + \varepsilon_{idt} \quad (1.1)$$

where Y_{idt} denotes the outcome of individual i born in cohort t and residing in district d , γ_t are cohort fixed effects, and δ_d are district fixed effects. The treatment variable, $\text{SchoolExposure}_{dt}$, measures cumulative schools constructed per 1,000 school-age children in district d during the years in which cohort t was between ages 6 and 15. Standard errors are clustered at the district level throughout.

The identifying assumption is that, absent school construction, outcomes across dis-

districts would have evolved similarly over cohorts. District fixed effects absorb time-invariant differences across districts, including baseline educational infrastructure and long-run socioeconomic characteristics, while cohort fixed effects absorb aggregate national trends affecting all districts similarly. Identification therefore comes from differences in the intensity of school expansion across districts experienced by different cohorts during school-age years.

I focus on several sets of outcomes. First, I estimate the effect of school construction on women's educational attainment, including years of schooling and completion of primary and secondary education. Second, I examine marriage-market outcomes, including age at marriage, the probability of remaining unmarried by age 18, and the spousal age gap. Third, I study equilibrium changes in matching patterns through measures of educational homogamy and assortative mating. Finally, I examine whether educational expansion altered intra-household outcomes such as women's decision-making power and domestic violence.

The primary treatment measure is continuous and captures cumulative school exposure during school-age years. I estimate the effect of adding one additional school per 1,000 school-age children to a district, which corresponds to approximately an 11 percent increase relative to the sample mean of 8.885 schools per thousand children. Following Duflo (2001), I additionally present specifications using a binary treatment measure equal to one if school exposure is above the cohort-specific mean and zero otherwise. The results are qualitatively similar across specifications.

A central outcome of interest is educational assortative mating. A key challenge in measuring assortative matching is distinguishing changes in matching behavior from changes in the marginal distribution of education. As educational attainment rises over time, educational homogamy may increase mechanically even in the absence of stronger assortative preferences. To address this issue, I follow Eika, Mogstad and Zafar (2019)

and construct an assortative mating index that measures whether individuals with a given education level marry each other more frequently than would occur under random matching.

Marital sorting between women with education level e_f and men with education level e_m is defined as:

$$s(e_f, e_m) = \frac{Pr(E_f = e_f, E_m = e_m)}{Pr(E_f = e_f)Pr(E_m = e_m)} \quad (1.2)$$

The index equals one under random matching, while values greater than one indicate positive assortative matching. I focus primarily on same-education matches and construct district-by-marriage-cohort measures of assortative mating using weighted joint and marginal education distributions from the DHS.⁴

Because assortative matching is measured at the district-by-marriage-cohort level rather than the individual level, I estimate separate specifications for the assortative mating analysis. Specifically, I estimate regressions of the form:

$$S_{dm}^{\text{agg}} = \beta \overline{\text{SchoolExposure}}_{dm} + \delta_d + \lambda_m + \varepsilon_{dm} \quad (1.3)$$

where S_{dm}^{agg} denotes the aggregate assortative mating index in district d and marriage cohort m , δ_d are district fixed effects, and λ_m are marriage-cohort fixed effects. The variable $\overline{\text{SchoolExposure}}_{dm}$ measures the average school construction exposure during ages 6–15 among women in district d and marriage cohort m . Regressions are weighted by the number of couples observed within each district–marriage-cohort cell, and standard errors are clustered at the district level.

One limitation of the empirical design is that the DHS does not observe childhood district directly. Women are therefore linked to school construction exposure using dis-

⁴Additional details on the construction of the index are provided in Appendix A.1.2.

trict of current residence. Migration during school-age years is relatively limited in India, and Census evidence suggests that most marriage migration occurs within relatively local geographic areas. Nevertheless, some measurement error in exposure remains possible. In addition, because the UDISE school census was collected in 2018, schools that shut down before the survey are not observed. To the extent that missing historical schools are unrelated to later marriage outcomes conditional on district and cohort fixed effects, this measurement error would tend to attenuate estimated effects toward zero.

1.6 Pre-trends analysis

A key identifying assumption in this paper is that, absent school construction, outcomes across districts would have evolved similarly over time. In standard difference-in-differences settings with discrete policy adoption, this assumption is often assessed using event-study specifications that test for flat pre-treatment trends. However, this approach is not well-suited to the present context. School construction in India expanded gradually over several decades rather than through a sharp, discrete policy change, and treatment intensity varies continuously across cohorts within districts. As a result, defining a single “event year” for each district is inherently imprecise, and event-study estimates may capture smooth development trends correlated with school expansion rather than causal effects.

To address this, I implement a cohort-based pre-trend test that aligns directly with the identification strategy. The empirical design exploits variation in cumulative school construction exposure across cohorts within districts. The relevant concern for identification is therefore whether districts that ultimately experience greater school expansion were already on different outcome trajectories prior to treatment. To test this, I construct a measure of eventual school exposure, defined as the average cumulative school

construction exposure among later cohorts within a district. I then estimate regressions on early cohorts only, interacting indicators for 5-year birth cohort bins with this measure of eventual exposure, while including district and cohort-bin fixed effects. Formally, the test examines whether the evolution of outcomes across early cohorts differs systematically across districts with higher versus lower eventual exposure.

Under the identifying assumption, these interaction terms should be jointly indistinguishable from zero: prior to meaningful exposure, cohorts in high-expansion and low-expansion districts should exhibit similar trends. A rejection would indicate that districts that expand more are already on different trajectories, undermining the causal interpretation of the main estimates.

For women's educational attainment (Table 1.1), the interaction coefficients between cohort bins and eventual school exposure are small in magnitude and statistically insignificant. Relative to the 1965 cohort, the coefficients for the 1970 and 1975 cohorts are -0.147 and -0.244, with standard errors of 0.291 and 0.299, respectively. These estimates are economically small and imprecisely estimated, and there is no systematic pattern across cohorts. The joint test of the interaction terms yields an F-statistic of 0.65 with a p-value of 0.524, failing to reject the null hypothesis that the interactions are jointly zero. This indicates that districts that eventually experience higher school construction exposure do not exhibit differential pre-treatment trends in women's educational attainment.

A similar pattern emerges for age at marriage (Table 1.2). The interaction coefficients for the 1970 and 1975 cohorts are 0.192 and 0.191, with relatively large standard errors (0.346 and 0.357), and are statistically indistinguishable from zero. The joint test yields an F-statistic of 0.16 ($p = 0.856$), again failing to reject the null of no differential trends. This suggests that, prior to school expansion, marriage timing did not evolve differently in districts that would later receive more schools.

For the probability of remaining unmarried by age 18 (Table 1.3), the interaction

coefficients are extremely small, 0.005 and 0.003 for the 1970 and 1975 cohorts, and precisely estimated around zero. The joint test produces an F-statistic of 0.04 with a p-value of 0.963, providing strong evidence against differential pre-trends. This further supports the conclusion that early cohort trends are similar across districts with different eventual exposure levels.

Overall, the evidence from these pre-trend tests supports the assumption that, in the absence of school construction, outcomes across districts would have followed similar trajectories, lending credibility to the causal interpretation of the estimated effects of school access on education and marriage outcomes.

1.7 Results

This section examines how school construction affected women’s educational attainment, marriage-market outcomes, assortative matching, and intra-household dynamics. Tables 1.4 through 1.11 present estimates of the effects of school exposure on education, marriage timing, educational sorting, and domestic violence. I first show that school construction substantially increased women’s educational attainment, particularly at lower levels of schooling. I then examine how these educational gains reshaped marriage markets, affecting marriage timing, spousal matching, and educational assortative mating. Finally, I examine whether these changes translated into improvements in women’s intra-household outcomes, including decision-making power and domestic violence. Across specifications, Panel A reports estimates using the continuous measure of cumulative schools per 1,000 school-age children, while Panel B uses a binary indicator for above-average exposure. All regressions include district and cohort fixed effects, with standard errors clustered at the district level.

1.7.1 Educational Attainment

If increased local availability of schools through school construction reduces the distance and cost of attending school, constraints that may be particularly salient for girls, then educational expansion may lead to higher levels of completed educational attainment among women later in life. Changes in women's educational attainment may in turn alter the composition of local marriage markets and affect subsequent marriage outcomes and assortative matching. Understanding the effect of school construction on educational attainment is therefore important for interpreting the broader marriage-market effects of educational expansion.

Across specifications, increased school construction leads to meaningful improvements in women's education, particularly at lower levels of schooling. As shown in Table 1.4, Panel A, an additional school per 1,000 school-age children increases women's schooling by 0.123 years ($p < 0.01$). This corresponds to an elasticity of 0.164, implying that a 10 percent increase in school exposure raises educational attainment by approximately 1.6 percent. School construction also increases the probability of completing primary education by 1.9 percentage points ($p < 0.01$), with a larger elasticity of 0.263, indicating substantial gains at the extensive margin of basic schooling. In contrast, there is no detectable effect on secondary school completion. The estimated coefficient is close to zero and statistically insignificant, with a slightly negative elasticity (-0.015), suggesting that school expansion during this period primarily affected access to and completion of lower levels of education rather than progression to higher levels. Panel B shows consistent patterns using the binary exposure measure. Women exposed to above-average school construction attain 0.428 additional years of schooling ($p < 0.01$) and are 4.5 percentage points more likely to complete primary school ($p < 0.01$). The implied elasticities are smaller in magnitude due to the discrete nature of the treatment, but the direction

and relative strength of effects mirror those in Panel A. As before, there is no statistically significant effect on secondary completion. These estimates are consistent with the broader descriptive expansion in women's schooling documented in Figure 1.4, though the regression results isolate variation in school exposure across districts and cohorts.

Taken together, these results indicate that school construction substantially expanded access to basic education for women, with the strongest impacts at the primary level. The absence of effects on secondary completion suggests that additional constraints—such as supply limitations at higher levels or demand-side factors—may continue to restrict progression beyond primary schooling.

1.7.2 Marriage Timing

Educational expansion may affect marriage timing through several channels. Increased access to schooling may delay marriage directly if women remain enrolled in school longer, reducing the likelihood of marriage during adolescence. At the same time, changes in educational attainment may alter reservation quality and matching opportunities within marriage markets, leading households to delay marriage in search of more desirable matches. Because marriage remains nearly universal in India, even modest shifts in marriage timing may have important implications for household formation and women's life trajectories.

Figure 1.2 shows that women with higher educational attainment consistently marry later and are substantially less likely to marry before age 18 across all cohorts. The gap between secondary-educated women and women with no schooling remains large throughout the period, suggesting that educational attainment is strongly associated with delayed marriage. Appendix Figure A.3 further shows that educational differences in marriage timing persist at later ages, particularly among women with secondary edu-

cation, although marriage remains nearly universal by age 25 for lower education groups.

Table 1.5 examines the effects of school construction on women’s age at marriage, the spousal age gap, and the probability of remaining unmarried by age 18. The results indicate modest but statistically significant effects across all outcomes. Panel A shows that an additional school per 1,000 school-age children increases the age at marriage by 0.036 years relative to a mean of 19.05 years. This corresponds to an elasticity of 0.016. At the same time, school construction reduces the age gap between husbands and wives by 0.026 years, implying an elasticity of -0.048 relative to the mean age gap of 4.84 years.⁵ Finally, the probability of remaining unmarried by age 18 increases by 0.3 percentage points, corresponding to an elasticity of 0.046. These findings indicate that improved access to schooling delayed marriage and modestly narrowed spousal age differences. Panel B presents results using the binary treatment. Moving from below-average to above-average school construction increases the age at marriage by 0.303 years, reduces the spousal age gap by 0.117 years, and raises the probability of remaining unmarried by age 18 by 2.4 percentage points. The corresponding elasticities are 0.005, -0.008 , and 0.013, respectively. Appendix Table A.2 shows that the estimated effects on remaining unmarried attenuate substantially by ages 21 and 25, suggesting that educational expansion primarily delayed the timing of marriage rather than generating persistent effects at later ages.

1.7.3 Educational Homogamy & Assortative Matching

The central question of this paper is whether educational expansion reshaped equilibrium patterns of marital sorting. While the previous section shows that school con-

⁵Husband’s age is constructed using information from both the women’s survey and the household roster of the India DHS. I use the husband’s age reported in the women’s survey when available and supplement missing values using the household roster for co-residing spouses. Household roster information is unavailable for husbands not residing in the household at the time of interview. For observations where husband’s age is reported in both sources, the measures are nearly perfectly correlated.

struction delayed marriage and increased women’s educational attainment, these changes may also alter who marries whom within local marriage markets. I therefore next examine whether educational expansion strengthened educational homogamy and assortative mating beyond what would be expected from mechanical changes in the distribution of education alone.

Table 1.6 presents reduced-form evidence on educational homogamy by examining whether women exposed to greater school construction during childhood marry more educated husbands.⁶ Across specifications, school construction exposure is associated with higher levels of husband educational attainment. In the continuous specification, an additional school per 1,000 school-age children increases husbands’ years of schooling by 0.081 years ($p < 0.01$), raises the probability of primary completion by 0.5 percentage points ($p < 0.01$), and increases the probability of secondary completion by 0.4 percentage points ($p < 0.05$). Relative to sample means, these correspond to elasticities of approximately 0.09, 0.06, and 0.09 respectively. The binary specification yields similar patterns. Exposure to above-average school construction increases husbands’ years of schooling by 0.268 years ($p < 0.01$) and raises the probability of secondary school completion by 2.4 percentage points ($p < 0.01$), while effects on primary completion remain small and statistically insignificant. Taken together, these results suggest that educational expansion altered the composition of local marriage markets and increased matching among more educated spouses.⁷

⁶Husband’s education is constructed using the household roster and linked to women through the spouse line number reported in the women’s questionnaire. As a result, husband characteristics are unavailable for spouses not residing in the household at the time of interview. In the analysis sample, approximately 92% of currently married women report co-residing with their husbands, suggesting that non-co-residence affects a relatively small share of observations. For observations where husband’s education is available both in the household roster and in the women’s survey, the two measures are highly correlated.

⁷Descriptively, marriages in which wives are more educated become more common among younger cohorts, while hypergamous marriages decline (Appendix Figure A.4). However, I find no evidence that school construction causally affects hypergamy or hypogamy patterns (Appendix Table A.3).

These regressions provide suggestive evidence of increased educational homogamy, but they do not by themselves distinguish between stronger assortative matching and mechanical changes in the marginal distribution of education. In particular, if school construction increases educational attainment for both men and women within local marriage markets, then observed increases in spousal education may partly reflect compositional changes rather than stronger sorting preferences. In addition, husbands' and wives' exposure to school construction may be correlated because couples often reside in the same district and age gaps between spouses have narrowed over time, implying that men and women in more recent cohorts may have experienced educational expansion simultaneously. As a result, the estimates in Table 1.6 should not be interpreted as causal effects of women's education on husbands' schooling. Instead, they capture equilibrium outcomes in local marriage markets undergoing educational expansion. To account explicitly for changes in the marginal distribution of education, I therefore turn to the assortative mating index, which measures sorting relative to random matching.⁸

Table 1.7 shows the assortative mating index and average school construction exposure for 5-year binned marriage cohorts. Figure 1.5 plots assortative mating separately by educational category across marriage cohorts. The figure reveals substantial increases in educational sorting at the lower end of the education distribution. Individuals with no schooling increasingly marry others with no schooling relative to what would be expected under random matching given the changing marginal distribution of education. The assortative mating index for the no-schooling group rises steadily across cohorts, increasing from roughly 1.4 in the earliest marriage cohorts to more than 7 in the most recent cohorts. This pattern is structurally similar to the evolution of assortative mating documented by Eika, Mogstad and Zafar (2019) for the United States and several European countries, though the increase in assortative mating among the least educated is sub-

⁸Construction of the index is described in Appendix A.1.2.

stantially larger in the Indian context. Individuals with primary education also exhibit positive assortative matching, though the degree of stratification is less pronounced than among individuals with no schooling. By contrast, assortative mating among individuals with upper primary education remains relatively stable across cohorts, while assortative matching among those with secondary education declines modestly over time. As educational attainment expands and secondary schooling becomes increasingly common, high school education may be becoming less distinctive within the marriage market, leading to weaker educational homogamy at higher education levels. Overall, these results suggest that educational expansion increased segmentation at the lower end of the education distribution even as average educational attainment rose substantially.

I next estimate the causal effect of school construction on the educational assortative mating index. Table 1.8 reports two-way fixed-effects estimates of the effect of district-level school construction on the assortative mating index by marriage cohort. The coefficient on school exposure is positive and statistically significant at the 1 percent level. Quantitatively, an additional school per 1,000 school-age children increases the assortative mating index by 0.033, corresponding to an increase of approximately 1.6 percent relative to the mean index value of 2.024. These estimates indicate that educational expansion modestly but significantly strengthened educational sorting within marriage markets. At the same time, the baseline level of the index indicates that India's marriage market is already strongly positively assortative even in the absence of additional school construction, with same-education matches occurring roughly twice as frequently as under random matching.

Figure 1.6 compares the observed rise in assortative mating across marriage cohorts with the increase predicted by school construction exposure using the estimated treatment effects. The predicted series tracks a substantial share of the upward trend in assortative mating, though observed assortative mating rises more sharply during later

cohorts. This suggests that educational expansion played an important role in strengthening educational sorting, while additional social and economic changes beyond school construction also contributed to the increase in assortative mating over time.

To quantify the contribution of school construction to the observed rise in assortative mating, I compare changes in the assortative mating index across marriage cohorts with the changes predicted by the estimated treatment effects. Between the early marriage cohorts (1980–1985) and the late cohorts (2010–2015), the assortative mating index increases from approximately 1.65 to 2.11, corresponding to a rise of roughly 0.46. Over the same period, average school construction exposure increases from approximately 5.09 to 11.05 schools per 1,000 children, a change of nearly 5.96. Using the estimated coefficient from Table 1.8, this implies a predicted increase in the assortative mating index of approximately 0.19. Back-of-the-envelope calculations therefore suggest that school construction can account for roughly 43 percent of the observed increase in assortative mating across marriage cohorts. Overall, these findings indicate that educational expansion reshaped not only individual educational attainment, but also the structure of marital matching within local marriage markets.

1.7.4 Decision-making & Domestic Violence

An expansion in access to schooling can plausibly affect intra-household outcomes such as women’s decision-making power and the incidence of domestic violence through several channels. On the one hand, increased education may enhance women’s bargaining power within marriage by improving their outside options, increasing their productivity, and shifting preferences toward greater autonomy. This would predict improvements in decision-making and reductions in domestic violence. On the other hand, if labor market opportunities remain limited—as is the case in many parts of India—education may not

translate into meaningful outside options. In such settings, increases in women’s education may instead alter expectations within marriage without changing underlying gender norms, potentially leading to conflict or “backlash.” Finally, domestic violence may be driven by deeply entrenched norms and intergenerational transmission, in which case improvements in schooling alone may have limited effects. The empirical analysis therefore examines both average effects and potential mechanisms through which schooling could influence domestic violence outcomes.

Table 1.9 examines whether exposure to school construction affects women’s decision-making power and the incidence of domestic violence.⁹ The sample includes only women selected for the domestic violence module in the DHS data, and regressions use module weights accordingly. The results provide little evidence that school construction improved women’s intra-household outcomes. Across specifications, the coefficients on school construction are small in magnitude for decision-making and most measures of domestic violence. There is weak evidence of an increase in controlling behavior and emotional abuse under the continuous specification: an additional school per 1,000 children increases controlling behavior by 0.005 ($p < 0.10$) and emotional abuse by 0.004 ($p < 0.05$). However, these effects are not robust to the binary specification, where estimates are close to zero and statistically insignificant. There is no evidence of meaningful effects on physical or sexual violence. Overall, the estimates are small and imprecisely estimated, and do not suggest economically meaningful improvements in domestic violence or decision-making. These findings suggest that, despite increasing educational attainment, school construction did not translate into substantial changes in intra-household power or relationship quality.

Table 1.10 examines whether domestic violence is transmitted across generations and

⁹Additional details on the construction of domestic violence and decision-making measures are provided in Appendix A.1.3

whether increased schooling attenuates this transmission. Panel A estimates the baseline relationship between exposure to domestic violence in the parental household (measured by whether the respondent's father beat her mother) and her own experience of domestic violence in adulthood. Panel B augments the baseline specification by interacting parental exposure with school construction. Panel A reveals strong and highly significant intergenerational transmission of domestic violence. Women whose fathers beat their mothers are substantially more likely to experience emotional, physical, and sexual violence themselves. The magnitudes are large: emotional violence increases by 13.4 percentage points, physical violence by 26.8 percentage points, and sexual violence by 6.8 percentage points relative to mean outcomes of 0.12, 0.26, and 0.05, respectively. Panel B shows that increased access to schooling does not attenuate this intergenerational persistence. The interaction between school construction and parental exposure is close to zero and statistically insignificant for emotional and physical violence, and small and positive for sexual violence (0.002, $p < 0.05$). These results indicate that, while schooling expanded educational attainment and altered marriage outcomes, it did not weaken the transmission of domestic violence across generations. This points to the persistence of deeply rooted norms that are not easily shifted by improvements in schooling alone.

Table 1.11 tests whether increases in women's relative education generate "backlash" within marriage. Panel A estimates the relationship between relative spousal education and domestic violence, comparing couples in which the wife has more education than her husband or vice versa, with equal-education couples as the omitted category. Relative spousal education is constructed using husband education linked from the household roster through the spouse line number reported in the women's questionnaire, so the sample is restricted to women in the domestic violence module with successfully linked co-resident spouses. Panel B interacts relative education with school construction exposure to assess whether schooling-induced changes in women's relative education affect

domestic violence outcomes. Panel A provides little evidence that women who are more educated than their husbands experience higher levels of domestic violence. The coefficients on the “wife more educated” indicator are small in magnitude and statistically insignificant across emotional, physical, and sexual violence outcomes. Women whose husbands are more educated likewise do not experience systematically higher levels of violence. If anything, the estimates suggest modest reductions in violence relative to equal-education couples, though only the coefficient for sexual violence is statistically significant (-0.012 , $p < 0.05$). Panel B similarly provides little support for backlash effects arising from educational expansion. The interaction between school exposure and the “wife more educated” indicator is small and negative across specifications and statistically significant for physical and sexual violence. Rather than increasing violence, greater school exposure is associated with modest declines in these outcomes among couples in which wives are more educated than their husbands. Overall, the results provide little evidence that increases in women’s relative education induced by school construction generated backlash within marriage.

Taken together, the results in this section suggest that educational expansion substantially altered educational attainment and marriage-market outcomes while leaving intra-household gender dynamics largely unchanged. School construction does not meaningfully improve women’s decision-making power or reduce domestic violence, and domestic violence exhibits strong intergenerational persistence that is not attenuated by increased schooling. At the same time, there is no evidence that increases in women’s relative education generate backlash effects within marriage. These findings suggest that educational expansion reshaped who marries whom and when households form, while deeply entrenched norms governing intra-household dynamics remained persistent.

1.8 Conclusion

To summarize, I find that school construction increases women’s educational attainment and generates systematic changes in marriage timing and matching patterns—women marry later, are less likely to enter child marriage, and match with husbands closer to their own age. These shifts are consistent with education raising women’s value in the marriage market, thereby improving their outside options and allowing families to delay marriage in search of higher-quality matches. Taken together, these findings imply that increased access to schooling generates meaningful marital returns to education in equilibrium, even in settings where labor market returns to female education may be limited.

At the same time, the results show that these gains in the marriage market do not translate into corresponding improvements in intra-household outcomes. I find little evidence that school construction increases women’s decision-making power or reduces domestic violence. Domestic violence remains strongly persistent across generations, and increased schooling does not attenuate this transmission. Moreover, there is no evidence that schooling-induced changes in women’s relative education generate backlash within marriage. These findings suggest that while education reshapes who marries whom and when, it does not substantially alter the underlying norms governing intra-household dynamics.

The observed changes in assortative mating further indicate that schooling expansions reshape the equilibrium structure of marriage markets by altering the distribution of educated women. Educational sorting strengthens at the bottom of the distribution—where access remains limited—while weakening at higher levels where schooling becomes more widespread. This pattern points to a potential unintended consequence of educational expansion: increased stratification among the least educated, even as average attainment

risers.

Overall, the results highlight a divergence between returns to education in the marriage market and changes in intra-household bargaining power. Schooling expansions improve matching outcomes and delay marriage, but do not, on their own, shift deeply entrenched gender norms or reduce domestic violence. In future work, I plan to examine whether these changes in marriage market structure translate into intergenerational persistence in human capital, particularly through children's educational outcomes.

1.9 Figures and Tables

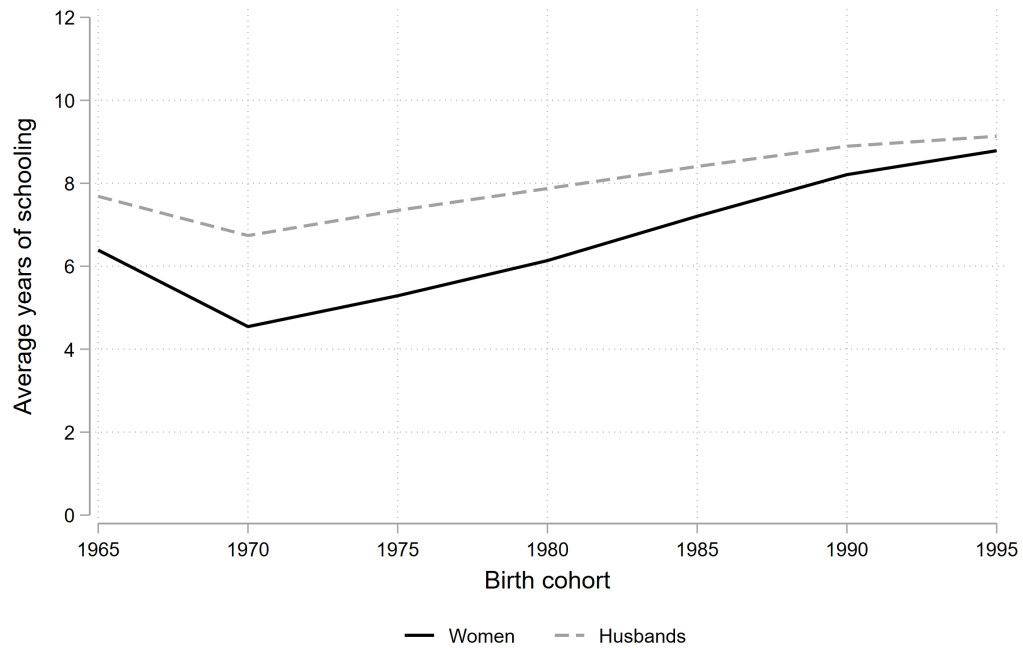


Figure 1.1: Average Years of Schooling among Married Couples by Birth Cohort

Notes: The figure plots average years of schooling for women and their co-resident husbands by women's birth cohort using DHS data. Women's educational attainment rises substantially across cohorts, narrowing the gender gap in schooling over time.

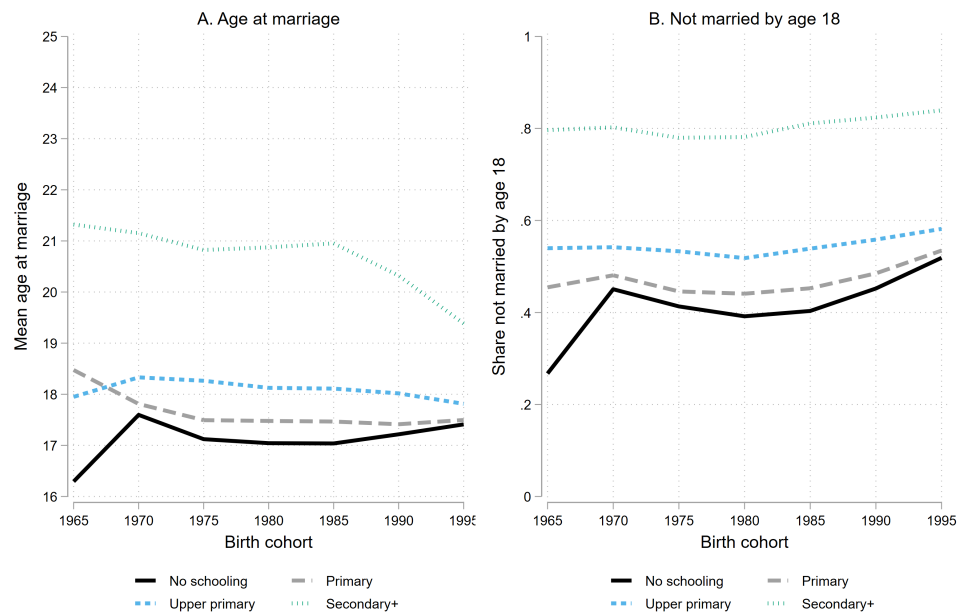


Figure 1.2: Marriage Timing by Education Level

Notes: Panel A plots mean age at marriage by women's birth cohort and education category. Panel B plots the share of women not married by age 18. Educational categories are defined as no schooling, primary (1–5 years), upper primary (6–8 years), and secondary or higher (9+ years).

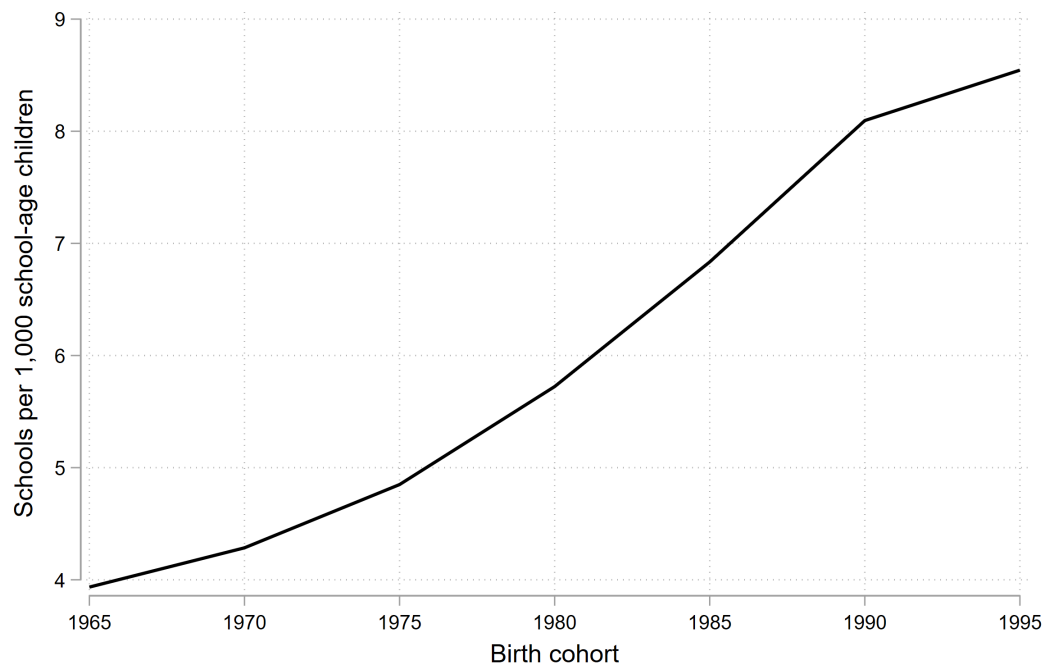


Figure 1.3: School Exposure by Birth Cohort

Notes: The figure plots average cumulative school exposure during ages 6–15 by women’s birth cohort. School exposure is measured as the cumulative number of schools constructed per 1,000 school-age children in a woman’s district of residence during her school-age years.

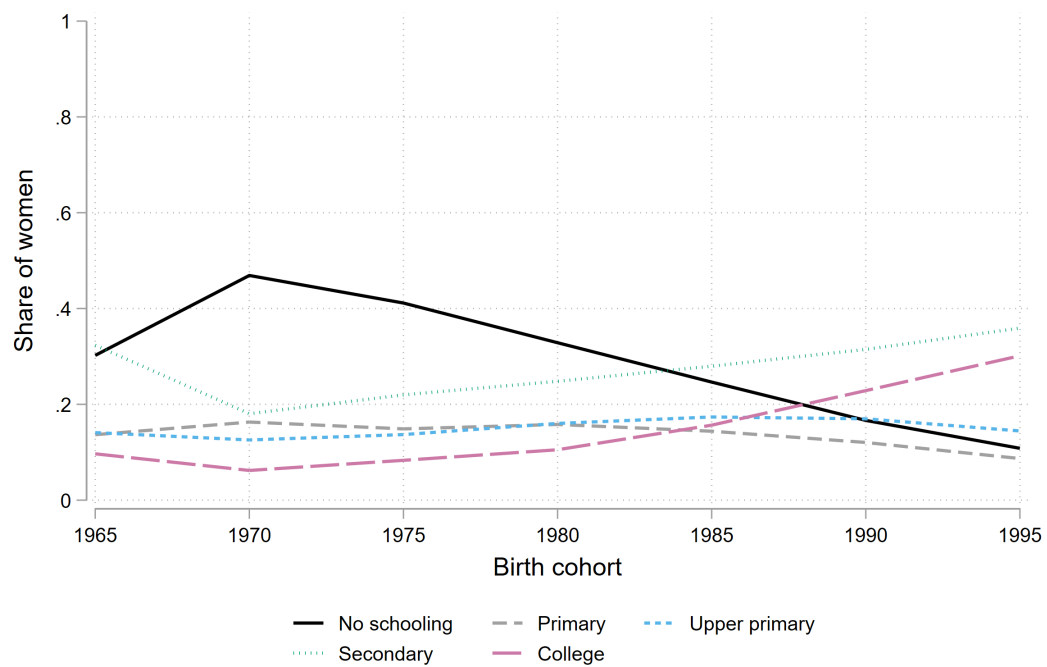


Figure 1.4: Distribution of Women's Education by Birth Cohort

Notes: The figure plots the share of women in each education category by five-year birth cohort. Education categories are defined as no schooling, primary, upper primary, secondary, and college.

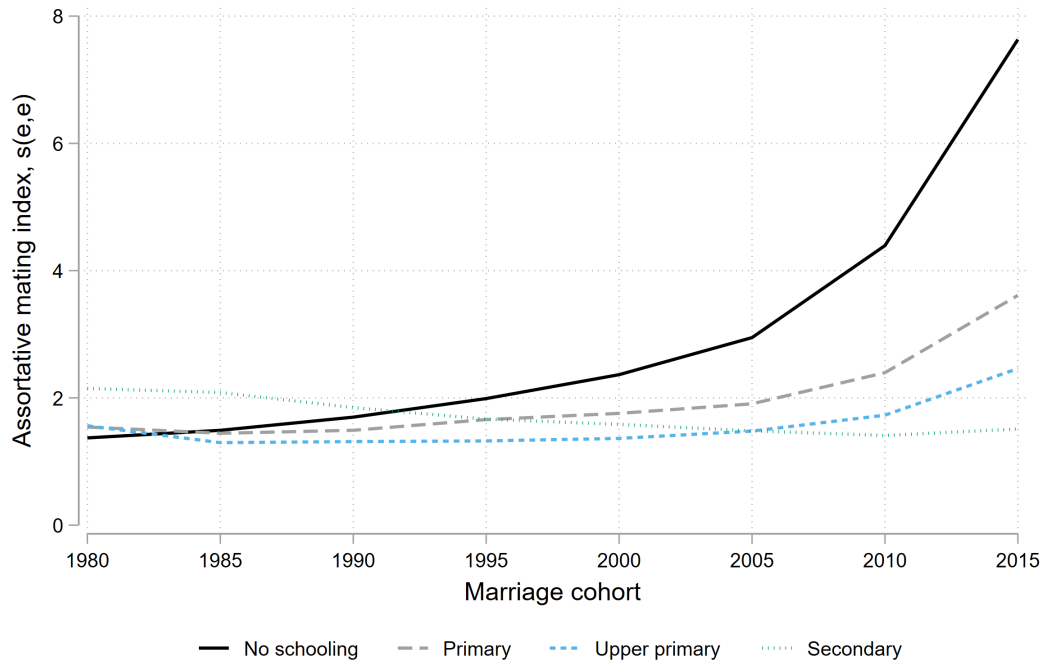


Figure 1.5: Educational Assortative Mating by Marriage Cohort

Notes: The figure plots the assortative mating index $s(e,e)$ for each education group, defined as the ratio of the observed share of same-education couples to the share expected under random matching. Cohorts are grouped into five-year marriage cohorts. The secondary category includes women with secondary or higher education. College is excluded due to small cell sizes.

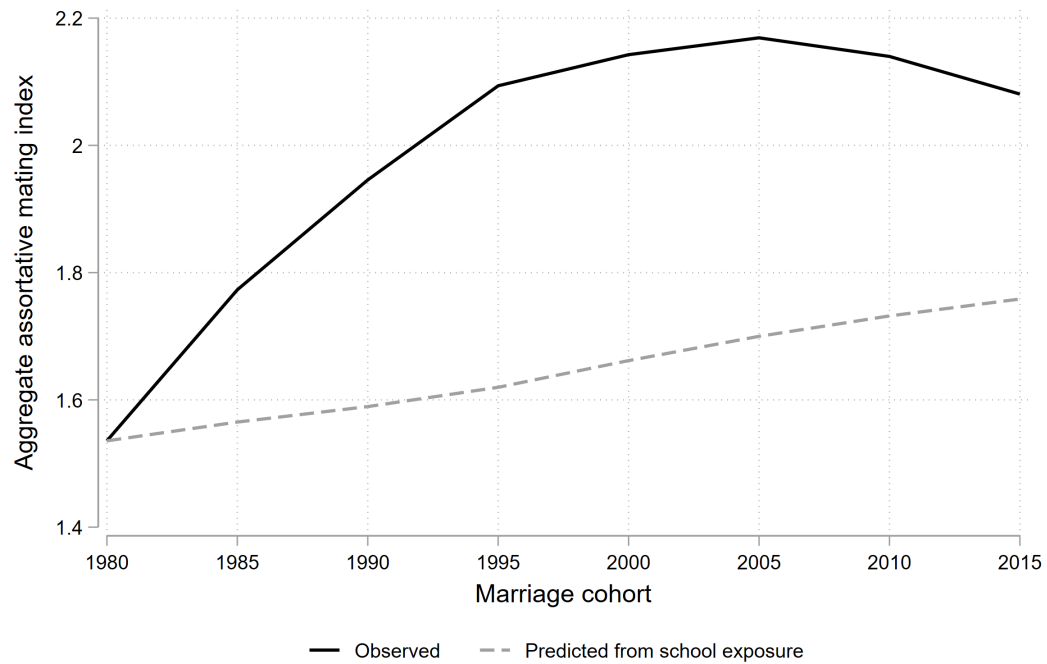


Figure 1.6: Observed and Predicted Assortative Mating

Notes: The figure compares the observed aggregate assortative mating index across marriage cohorts with the level predicted by changes in school exposure using the estimated regression coefficients. The predicted series is constructed by applying the estimated effect of school exposure on assortative mating to changes in average cohort-level school exposure over time.

Table 1.1: Pre-Trend Test for Women's Years of Education

	(1) Years of education
5-year birth cohort=1965 $\times$ Future school exposure	0.000 (.)
5-year birth cohort=1970 $\times$ Future school exposure	-0.147 (0.291)
5-year birth cohort=1975 $\times$ Future school exposure	-0.244 (0.299)
Mean of dependent variable	4.227
N	80,409

Notes: The table reports cohort-specific pre-trend tests using early cohorts only. The treatment variable is district-level eventual school construction exposure, measured as average cumulative school exposure among later cohorts and standardized to have mean zero and standard deviation one. The regression includes district and cohort-bin fixed effects. Standard errors are clustered at the district level. Joint test: $F = 0.65$, $p = 0.524$.

Table 1.2: Pre-Trend Test for Age at Marriage

	(1) Age at marriage
5-year birth cohort=1965 $\times$ Future school exposure	0.000 (.)
5-year birth cohort=1970 $\times$ Future school exposure	0.192 (0.346)
5-year birth cohort=1975 $\times$ Future school exposure	0.191 (0.357)
Mean of dependent variable	19.003
N	79,505

Notes: Sample restricted to ever-married women. The table reports cohort-specific pre-trend tests using early cohorts only. The treatment variable is district-level eventual school construction exposure, measured as average cumulative school exposure among later cohorts and standardized to have mean zero and standard deviation one. The regression includes district and cohort-bin fixed effects. Standard errors are clustered at the district level. Joint test: $F = 0.16$, $p = 0.856$.

Table 1.3: Pre-Trend Test for Child Marriage

	(1) Not married by 18
5-year birth cohort=1965 $\times$ Future school exposure	0.000 (.)
5-year birth cohort=1970 $\times$ Future school exposure	0.005 (0.026)
5-year birth cohort=1975 $\times$ Future school exposure	0.003 (0.027)
Mean of dependent variable	0.593
N	80,409

Notes: The table reports cohort-specific pre-trend tests using early cohorts only. The treatment variable is district-level eventual school construction exposure, measured as average cumulative school exposure among later cohorts and standardized to have mean zero and standard deviation one. The regression includes district and cohort-bin fixed effects. Standard errors are clustered at the district level. Joint test: $F = 0.04$, $p = 0.963$.

Table 1.4: Effect of School Construction on Women's Education Outcomes

	(1)	(2)	(3)
	Years of education	Primary completed	Secondary completed
<i>Panel A: Continuous treatment</i>			
Schools built	0.123*** (0.021)	0.019*** (0.002)	-0.001 (0.002)
Mean of treatment	8.904	8.904	8.904
Elasticity of outcome	0.164	0.263	-0.015
<i>Panel B: Binary treatment</i>			
Schools built > mean	0.428*** (0.110)	0.045*** (0.008)	0.005 (0.011)
Mean of treatment	0.342	0.342	0.342
Elasticity of outcome	0.022	0.024	0.005
Mean of outcome	6.704	0.642	0.338
N	462,963	462,963	462,963

Notes: Each column reports estimates from regressions of the dependent variable on school construction exposure, with district and cohort fixed effects. Panel A uses cumulative schools constructed per 1,000 population aged 6–15 years. Panel B uses an indicator for whether exposure is above the cohort-specific mean. Mean of treatment reports the estimation-sample mean of the treatment variable used in that panel. Elasticities are computed as $\beta \times \bar{X} / \bar{Y}$. Standard errors are clustered at the district level.

Table 1.5: Effect of School Construction on Women's Marriage Outcomes

	(1) Age at marriage	(2) Age gap	(3) Not married by age 18
<i>Panel A: Continuous treatment</i>			
Schools built	0.036** (0.015)	-0.026*** (0.009)	0.003** (0.002)
Mean of treatment	8.625	8.748	8.904
Elasticity of outcome	0.016	-0.048	0.046
<i>Panel B: Binary treatment</i>			
Schools built > mean	0.303*** (0.082)	-0.117*** (0.043)	0.024** (0.010)
Mean of treatment	0.333	0.343	0.342
Elasticity of outcome	0.005	-0.008	0.013
Mean of outcome	19.052	4.842	0.642
N	432,094	368,201	462,963

Notes: Each column reports estimates from regressions of the dependent variable on school construction exposure, with district and cohort fixed effects. In columns (1) and (2), the sample is restricted to ever-married women. Column (3) includes all women. Panel A uses cumulative schools constructed per 1,000 population aged 6–15 years. Panel B uses an indicator for whether exposure is above the cohort-specific mean. Mean of treatment reports the estimation-sample mean of the treatment variable used in that panel. Elasticities are computed as $\beta \times \bar{X}/\bar{Y}$. Standard errors are clustered at the district level.

Table 1.6: Effect of School Construction in Wife's District on Husband's Education

	(1)	(2)	(3)
	Years of education	Primary completed	Secondary completed
<i>Panel A: Continuous treatment</i>			
Schools built	0.081*** (0.017)	0.005*** (0.001)	0.004** (0.002)
Mean of treatment	8.759	8.759	8.759
Elasticity of outcome	0.089	0.058	0.089
<i>Panel B: Binary treatment</i>			
Schools built > mean	0.268*** (0.087)	0.003 (0.007)	0.024*** (0.008)
Mean of treatment	0.345	0.345	0.345
Elasticity of outcome	0.012	0.002	0.020
Mean of outcome	7.953	0.765	0.411
N	360,548	360,548	360,548

Notes: Each column reports estimates from regressions of the dependent variable on school construction exposure in the woman's district when young, with district and cohort fixed effects. Panel A uses cumulative schools constructed per 1,000 population aged 6–15 years. Panel B uses an indicator for whether exposure is above the cohort-specific mean. Mean of treatment reports the estimation-sample mean of the treatment variable used in that panel. Elasticities are computed as $\beta \times \bar{X}/\bar{Y}$. Standard errors are clustered at the district level.

Table 1.7: Trends in Assortative Mating and School Exposure

Binned marriage cohorts	Assortative mating index	Mean school construction exposure
1980	1.536	4.642
1985	1.773	5.544
1990	1.946	6.287
1995	2.094	7.213
2000	2.143	8.498
2005	2.169	9.663
2010	2.140	10.643
2015	2.081	11.456

Notes: The table reports the aggregate assortative mating index S^{agg} and mean school construction exposure for each five-year marriage cohort.

Table 1.8: Effect of School Construction on Educational Assortative Mating

	Assortative mating index
Schools built	0.033*** (0.007)
Mean of outcome	2.024
N	5200

Notes: District-by-marriage-cohort regression with district and cohort fixed effects. Cells are weighted by the number of couples. Mean cumulative school construction exposure in the estimation sample is 7.686.

Table 1.9: Effect of School Construction on Autonomy and Domestic Violence

	(1) Decision making	(2) Controlling partner	(3) Emotional abuse	(4) Physical violence	(5) Sexual violence
<i>Panel A: Continuous treatment</i>					
Schools built	0.002 (0.004)	0.005* (0.003)	0.004* (0.002)	0.001 (0.002)	-0.001 (0.001)
Mean of treatment	9.071	9.071	9.071	9.071	9.071
Elasticity of outcome	0.024	0.115	0.265	0.045	-0.114
<i>Panel B: Binary treatment</i>					
Schools built > mean	0.007 (0.016)	-0.010 (0.021)	0.001 (0.012)	-0.001 (0.017)	0.005 (0.007)
Mean of treatment	0.348	0.348	0.348	0.348	0.348
Elasticity of outcome	0.004	-0.008	0.002	-0.002	0.035
Mean of outcome	0.731	0.425	0.120	0.261	0.053
N	54,332	54,332	54,332	54,332	54,332

Notes: Each column reports estimates from regressions of the indicated outcome on school construction exposure, with district and cohort fixed effects. Panel A uses cumulative schools constructed per 1,000 population aged 6–15; Panel B uses an indicator for whether exposure is above the cohort-specific mean. Outcomes are measured for women in the DHS domestic violence module using DV weights. Decision-making is the average of indicators for participation in key household decisions; other outcomes are binary indicators of controlling behavior, emotional abuse, physical violence, and sexual violence. Elasticities are computed as $\beta \times \bar{X}/\bar{Y}$. Standard errors are clustered at the district level.

Table 1.10: Effect of School Construction on Intergenerational Transmission of Domestic Violence

	(1) Emotional	(2) Physical	(3) Sexual
<i>Panel A: Intergenerational transmission</i>			
Father beat mother	0.134*** (0.015)	0.268*** (0.014)	0.068*** (0.010)
<i>Panel B: Intergenerational transmission $\times$ school exposure</i>			
Schools built	0.003* (0.002)	0.000 (0.002)	-0.001 (0.001)
Father beat mother	0.133*** (0.022)	0.259*** (0.023)	0.056*** (0.015)
Schools $\times$ Father beat mother	0.000 (0.002)	0.002 (0.002)	0.002* (0.001)
Mean of outcome	0.120	0.261	0.053
N	54,332	54,332	54,332

Notes: Each column reports estimates from regressions of the indicated domestic violence outcome on indicators for exposure to domestic violence in the parental household and school construction exposure, with district and birth-cohort fixed effects. Panel A estimates the baseline relationship between a respondent's exposure to parental domestic violence and her own experience of domestic violence in adulthood. Panel B interacts this indicator with school construction exposure in the woman's district during her school-age years. The sample is restricted to women selected into the DHS domestic violence module; regressions use domestic violence weights. Standard errors are clustered at the district level.

Table 1.11: Testing for Backlash: Relative Spousal Education, School Construction, and Domestic Violence

	(1) Emotional	(2) Physical	(3) Sexual
<i>Panel A: Relative education</i>			
Wife more educated	-0.010 (0.008)	-0.009 (0.011)	-0.005 (0.006)
Husband more educated	-0.011 (0.007)	-0.009 (0.010)	-0.012** (0.005)
<i>Panel B: Relative education $\times$ school exposure</i>			
Schools built	0.003 (0.002)	0.001 (0.002)	0.000 (0.001)
Wife more educated	-0.007 (0.012)	0.005 (0.014)	0.003 (0.008)
Schools $\times$ Wife more	-0.001 (0.001)	-0.002** (0.001)	-0.001** (0.001)
Mean of outcome	0.119	0.269	0.052
N	43,320	43,320	43,320

Notes: Each column reports estimates from regressions of the indicated domestic violence outcome on measures of relative spousal education, with district and birth-cohort fixed effects. Panel A includes indicators for whether the wife has more years of education than her husband and whether the husband has more education than the wife, with equal-education couples as the omitted category. Panel B interacts the “Wife more educated” indicator with school construction exposure in the woman’s district during her school-age years. The sample is restricted to women selected into the DHS domestic violence module with non-missing information on both spouses’ education; regressions use domestic violence weights. Standard errors are clustered at the district level.

Chapter 2

School Construction, Household Formation, and Intergenerational Human Capital in India

Abstract

Using staggered school construction across Indian districts, I estimate the effects of increased school access on women's fertility behavior and the educational and health outcomes of the next generation. An additional school per 1,000 school-age children, equivalent to roughly an 11 percent increase relative to the sample mean exposure, delays age at first birth for women by approximately 0.03 years and reduces completed fertility by nearly 6 percent over the life cycle. Increased school access also generated significant intergenerational improvements in children's educational attainment, increasing years of schooling by 8.6 percent and reducing the probability of falling behind grade level by nearly 19 percent. School construction additionally improves child height-for-age and increases the likelihood of institutional delivery. Overall, the findings suggest that large-scale investments in schooling generated persistent intergenerational benefits extending beyond directly exposed cohorts. School construction increased parental educational attainment and altered the distribution of education across households, but I find limited evidence that educational expansion substantially strengthened complementarities between parents in the production of child human capital.

2.1 Introduction

Large-scale investments in schooling are typically evaluated by their effects on the directly exposed generation, including educational attainment, labor market outcomes, and demographic behavior. Yet educational expansion may also have consequences that extend across generations. By changing who becomes educated, patterns of marital sorting, and the timing and quantity of fertility, schooling policies may reshape the distribution of human capital across households and thereby affect the conditions into which children are born and the investments they receive (Becker, 1981; Fernández, Guner and Knowles, 2005; Greenwood et al., 2014; Eika, Mogstad and Zafar, 2019). These intergenerational effects are central to understanding the long-run returns to education policy, particularly in settings where marriage and fertility remain closely linked to household formation and intergenerational mobility.

India provides a particularly useful setting in which to study these dynamics. Over the past several decades, India has experienced a dramatic expansion in schooling access alongside substantial declines in fertility, delays in marriage, and large increases in women’s educational attainment. At the same time, marriage remains nearly universal and continues to structure household formation, fertility, and the transmission of socioeconomic advantage across generations (Becker and Tomes, 1979; Black and Devereux, 2011). Moreover, because marriages in India remain heavily family-mediated¹, educational attainment functions as an especially salient matching characteristic within marriage markets. Educational assortative matching has also strengthened over time, suggesting that educational expansion may have altered not only average human capital accumulation, but also the distribution of parental human capital across households. As

¹Based on data from the India Human Development Survey, approximately 66 percent of ever-married women reported that their marriages were arranged primarily by parents or relatives, while an additional 29 percent reported that the match was chosen jointly by themselves and their families. Only about 5 percent reported choosing their spouse independently.

educational attainment rises, changes in assortative matching and fertility behavior may concentrate parental human capital within some households more than others, potentially reshaping the transmission of advantage across generations.

This paper studies the long-run intergenerational effects of educational expansion in India. I combine administrative data on school construction with nationally representative Demographic and Health Survey (DHS) data to examine how women’s childhood exposure to school construction affects fertility behavior and the educational and health outcomes of their children. The empirical strategy exploits variation across districts and birth cohorts in cumulative exposure to school construction during ages 6–15. This design allows me to study cumulative educational expansion over multiple decades rather than the effects of a single policy episode or short-run schooling reform.

The paper contributes to several literatures. First, it contributes to work on educational expansion and human capital accumulation. A large literature shows that increased access to schooling through school construction and educational expansion policies raises educational attainment and affects later-life demographic and labor market outcomes (Duflo, 2001; Breierova and Duflo, 2004; Osili and Long, 2008; Abanokova, Campos and Maitra, 2020; Khanna, 2023). In India, much of the existing evidence focuses on specific schooling reforms, particularly the District Primary Education Programme (DPEP), or on relatively shorter-run effects of educational expansion (Azam and Saing, 2016; Khanna, 2023; Sunder, 2019). I instead construct a long-run district-cohort panel of school construction exposure, allowing me to trace how cumulative educational expansion affected fertility behavior and the next generation’s outcomes over multiple decades.

Second, the paper connects educational expansion to fertility and household formation. Education may affect the next generation not only by increasing parental schooling directly, but also by changing marriage timing, fertility behavior, and the allocation of human capital across households through assortative matching. This link is especially im-

portant in India, where marriage and childbearing remain closely connected and where educational sorting within marriage markets has strengthened over time. The paper therefore relates to work on fertility, marriage markets, and household sorting (Becker, 1981; Desai and Andrist, 2010; Vogl, 2013; Fernández, Guner and Knowles, 2005; Greenwood et al., 2014; Eika, Mogstad and Zafar, 2019; Schwartz, 2013). I show that school construction delayed fertility and reduced completed fertility, suggesting that educational expansion altered both the timing and quantity of childbearing. Together with the child education results, these patterns are consistent with a quantity–quality framework in which lower fertility is accompanied by greater investments in children’s human capital (Becker, 1960; Becker and Lewis, 1973; Becker and Tomes, 1976; Doepke, 2015).

Third, the paper contributes to the literature on intergenerational transmission of human capital. Prior work shows that parental education is strongly associated with children’s schooling and health, though identifying causal effects is difficult because parental education is correlated with family background, preferences, and resources (Behrman et al., 1999; Behrman and Rosenzweig, 2002; Currie and Moretti, 2003; Black, Devereux and Salvanes, 2005; Oreopoulos, Page and Stevens, 2006; Azam, 2015; Sunder, 2019). I use school construction as a source of variation in parental schooling and household formation to examine whether educational expansion generated persistent intergenerational improvements in children’s outcomes. The estimates indicate significant positive effects on children’s educational attainment and grade progression, with children of exposed women attaining more schooling and exhibiting smaller schooling deficits relative to age-appropriate progression. Effects on child health are more selective. School construction increases height-for-age and institutional delivery, but has limited effects on stunting and vaccination, suggesting that improvements in child health may depend more heavily on complementary health infrastructure and service provision.

I also examine how parental education and household educational composition shape

these intergenerational effects. School construction increased women's educational attainment and altered the joint distribution of parental education through changes in marital sorting. The evidence suggests that improvements in children's schooling are closely linked to parental human capital and household educational composition. However, I find little evidence that school expansion strengthened complementarities between mothers' and fathers' education in producing child human capital. Instead, the relationship between maternal and paternal education exhibits diminishing marginal returns within households.

Overall, the paper shows that educational expansion generated persistent benefits extending beyond the directly exposed generation through changes in fertility behavior, parental human capital, and household formation. These findings broaden the interpretation of schooling investments: their effects are not limited to the individuals directly exposed to new schools, but extend to the next generation through the structure and resources of the households they form.

2.2 Institutional Context on Educational Expansion and Intergenerational Outcomes in India

School access expanded dramatically in India after Independence through a series of large-scale public investments in educational infrastructure. Beginning in the 1970s, the Indian government implemented successive programs aimed at universalizing elementary education and reducing geographic barriers to schooling. Initiatives such as the Minimum Needs Programme in the 1970s, the District Primary Education Programme (DPEP) in the 1990s, and the Sarva Shiksha Abhiyan (SSA) in the 2000s substantially increased school construction, teacher recruitment, and educational funding across the country.

These expansions generated large increases in the availability of primary and middle schools over time, particularly in rural areas where initial access had been limited.

As a result of these policies, younger cohorts were exposed to substantially greater school availability during their childhood years. Figure 2.1 shows that cumulative school exposure rose sharply across birth cohorts between the 1960s and 1990s. Average exposure more than doubled over this period, generating substantial variation across districts and cohorts in access to schooling during childhood. Census statistics similarly document major increases in educational infrastructure during this period. The share of villages with a primary school increased substantially between the 1970s and 1990s, while access to middle and secondary schools also expanded rapidly. By the 2010s, most rural households had access to a primary school within walking distance.

These educational expansions coincided with major changes in women's educational attainment, marriage timing, fertility behavior, and household formation. Figure 2.2 shows substantial increases in educational attainment across cohorts for both women and their husbands, implying large changes in the distribution of parental human capital across households over time. Women's average years of schooling rose sharply across cohorts, while the gender gap in education narrowed substantially among younger cohorts. At the same time, child marriage declined and fertility rates fell sharply over the same period. Educational assortative matching within marriage markets also strengthened, increasing the concentration of human capital within households. These patterns suggest that educational expansion may have affected children not only through increases in maternal schooling, but also through broader changes in household educational composition and parental matching patterns.

Several mechanisms may link parental educational expansion to children's outcomes. More educated parents may invest more in children's schooling, possess greater educational and health knowledge, delay fertility, or allocate household resources differently.

Educational expansion may also alter the types of households children are born into by changing marriage market matching patterns and the joint distribution of parental human capital within families. These channels imply that large-scale schooling expansions may generate substantial intergenerational effects extending beyond the directly exposed cohorts themselves.

India provides a particularly useful setting for studying these intergenerational effects because educational expansion occurred over multiple decades and varied substantially across cohorts and districts. This allows the construction of cohort-specific measures of childhood exposure to school construction that can be linked to later fertility behavior and children's educational outcomes observed in nationally representative household surveys.

2.3 Conceptual Framework

Educational expansion may affect children's outcomes through several interconnected household channels. First, increased access to schooling raises parental educational attainment. More educated parents may invest more in children's human capital, possess greater educational and health knowledge, or allocate household resources differently. These effects may improve both children's educational attainment and health outcomes. Second, educational expansion may alter fertility and household formation decisions. Related work using the same school construction setting finds that increased school access delayed marriage, while the empirical results presented later in this paper show corresponding changes in fertility timing. Lower fertility or delayed childbearing may increase investments per child through quantity-quality tradeoffs and changes in parental resource allocation over the life cycle. Third, educational expansion may reshape marriage market matching patterns and the joint distribution of parental education within households. If school expansion increases educational assortative matching, children may increasingly

be raised in households where both parents have relatively high levels of human capital. Changes in household educational composition may therefore affect intergenerational outcomes beyond the direct effect of maternal education alone.

These channels are jointly determined and difficult to separately identify empirically. Parental education, fertility behavior, and marital sorting are all endogenous responses to educational expansion and may affect one another simultaneously. The empirical analysis therefore focuses primarily on reduced-form intergenerational effects of school construction exposure. The parental education and household composition specifications should be interpreted as evidence on potential mechanisms and household correlates of intergenerational transmission rather than as a formal causal mediation decomposition.

2.4 Data

This paper combines administrative data on school construction with nationally representative household survey data and district-level population statistics to construct measures of educational exposure across districts and cohorts in India.

Data on school construction and school locations come from the Unified District Information System for Education (UDISE), compiled by the Department of Education, Government of India. UDISE is the most comprehensive administrative database on schools in India and covers approximately 1.5 million public and private schools nationwide. The database contains detailed school-level information including school location, grade levels offered, management type, enrollment, infrastructure, teacher characteristics, and year of establishment. I use information on school location and establishment year to construct district-level measures of cumulative school availability over time. Although the UDISE data were collected in 2018, the database records establishment years for all schools operating at the time of the survey, allowing the construction of a long-run

historical panel of school expansion. School establishment information is aggregated to the district-year level and merged with district-level population estimates to construct measures of schools per 1,000 school-age children. The main treatment variable is the cumulative number of schools constructed per 1,000 school-age children in a woman's district of residence during the years in which she was between ages 6 and 15. The treatment therefore captures exposure to the stock of schools available during childhood rather than contemporaneous construction flows. One limitation of the UDISE data is that it only includes schools operating at the time of the 2018 survey. Schools constructed historically but closed prior to the survey are not observed, potentially generating measurement error in historical exposure measures, particularly for earlier cohorts. If school closures are unrelated to later outcomes conditional on district and cohort fixed effects, this measurement error would tend to attenuate estimates toward zero. However, because older schools are mechanically more likely to be missing, some non-classical measurement error remains possible.

Information on education, fertility, and child outcomes comes from the Indian Demographic and Health Survey (DHS), also known as the National Family Health Survey (NFHS). The DHS is a nationally representative household survey collecting detailed information on demographic behavior, educational attainment, fertility histories, health outcomes, and household characteristics. The primary analysis uses data from the women's individual recode (IR), the household roster (PR), and the child recode (KR). The women's survey contains information on women aged 15–49, including years of schooling, fertility histories, age at first birth, marriage timing, and household characteristics. The household roster is used to construct education outcomes for school-age children and to recover husband characteristics using spouse line numbers reported in the women's questionnaire. The child recode contains detailed information on children under age five, including anthropometric measures, vaccination histories, and birth character-

istics. The main fertility outcomes include age at first birth, children ever born, age at last birth, and the interval between marriage and first birth. Child education outcomes include years of schooling, a schooling gap measure defined relative to age-appropriate grade progression, and an indicator for being behind grade level. Child health outcomes include height-for-age z-scores (HAZ), stunting (an indicator for HAZ below -2), a vaccination index based on receipt of basic childhood vaccines, and an indicator for institutional delivery, defined as birth occurring in a public or private health facility rather than at home. District identifiers are publicly available in the Indian DHS beginning with the 2015 round. The analysis therefore relies primarily on the 2015 and 2019 DHS rounds, which provide nationally representative data with district-level geographic identifiers. The analysis links individuals to school construction exposure based on district of current residence. While the DHS does not directly observe childhood district of residence, migration during school-age years is relatively limited in India. Moreover, Census evidence suggests that most marriage migration occurs over relatively short geographic distances, implying that district of residence is a reasonable proxy for childhood exposure for much of the population.

Population data are drawn from decadal Census of India population tables and district variation tables. These data are used to construct annual estimates of district-level school-age population through linear interpolation across census years. Combining these population estimates with school establishment information allows the construction of cohort-specific measures of school availability. A major challenge in long-run district-level analysis in India is the large number of administrative boundary changes since the first post-Independence census. Districts have been repeatedly split, merged, and reorganized, while several new states have also been created. To address this issue, I construct a harmonized district panel that maps historical school construction data, Census population estimates, and DHS district identifiers into a common geography

based on harmonized 2011 district boundaries.² This harmonization allows consistent measurement of school exposure across districts and cohorts over multiple decades.

Table 2.1 reports summary statistics for the main analysis samples. Women in the mother-level sample have an average of 4.6 years of schooling and approximately 3 children ever born. Among children aged 6–14, average schooling is 3.6 years, while 35 percent are behind age-appropriate grade level. Parents in the child education sample have substantially higher average schooling than women in the overall mother-level sample, reflecting selection into the sample of households with school-age children observed in the household roster. Child health outcomes indicate substantial deprivation: average height-for-age is more than one standard deviation below the WHO reference population, and roughly 35 percent of children are stunted.

2.5 Empirical Strategy

The empirical analysis exploits variation across districts and birth cohorts in exposure to school construction during childhood. School expansion in India occurred gradually over several decades through large-scale public investments in educational infrastructure under successive Five-Year Plans and later programs such as the District Primary Education Programme (DPEP) and Sarva Shiksha Abhiyan (SSA). These programs substantially expanded school availability across the country, generating variation across districts and cohorts in exposure to schooling during the school-age years.

The empirical design follows Duflo (2001) and exploits both spatial and temporal variation in school construction intensity. Specifically, I estimate regressions of the form:

$$Y_{ijdt} = \beta \text{SchoolExposure}_{dt} + \gamma_t + \delta_d + \varepsilon_{ijdt} \quad (2.1)$$

²Additional details on the construction of the harmonized district panel are provided in Appendix B.1.1.

where Y_{ijdt} denotes either an outcome for woman i or an outcome for child j born to woman i , where the mother belongs to birth cohort t and resides in district d . The treatment variable, $\text{SchoolExposure}_{dt}$, measures the cumulative number of schools constructed per 1,000 school-age children in district d during the years in which woman i was between ages 6 and 15.

The identifying assumption is that, absent school construction, outcomes across districts would have evolved similarly across cohorts. District fixed effects absorb time-invariant differences across districts, including baseline educational infrastructure and long-run socioeconomic characteristics, while cohort fixed effects absorb aggregate national trends affecting all districts similarly. Identification therefore comes from differences in the intensity of school expansion across districts experienced by different cohorts during childhood.

The primary treatment measure is continuous and captures cumulative school exposure during school-age years. I estimate the effect of adding one additional school per 1,000 school-age children to a district, which corresponds to approximately an 11 percent increase relative to the sample mean exposure. Following Duflo (2001), I additionally present specifications using a binary treatment measure equal to one if exposure is above the cohort-specific mean and zero otherwise. The results are qualitatively similar across specifications.

The analysis examines several categories of outcomes. First, I estimate the effect of school construction exposure on fertility and household formation outcomes, including age at first birth, completed fertility, and fertility timing over the life cycle. Second, I examine children's educational outcomes, including years of schooling, schooling gaps relative to age-appropriate progression, and the probability of falling behind grade level. Third, I examine child health outcomes including height-for-age, stunting, vaccination, and institutional delivery. Finally, I explore the role of parental educational composition

and assortative matching in shaping intergenerational transmission.

One limitation of the empirical design is that the DHS does not directly observe childhood district of residence. Individuals are therefore linked to school construction exposure using district of current residence. Migration during school-age years is relatively limited in India, and Census evidence suggests that most marriage migration occurs within relatively local geographic areas. Nevertheless, some measurement error in exposure remains possible. In addition, because the UDISE school census was collected in 2018, schools that shut down before the survey are not observed. To the extent that missing historical schools are unrelated to later fertility or child outcomes conditional on district and cohort fixed effects, this measurement error would tend to attenuate estimated effects toward zero.

2.6 Pre-trends Analysis

A key identifying assumption underlying the empirical strategy is that, absent school construction, outcomes across districts would have evolved similarly over time. Standard event-study approaches are less well suited to this setting because school expansion occurred gradually over several decades rather than through a discrete policy shock, and treatment intensity varies continuously across cohorts within districts. The relevant concern is therefore whether districts that ultimately experienced larger school expansion were already on different outcome trajectories prior to substantial exposure. To assess this, I implement cohort-based pre-trend tests using early maternal cohorts only. I first construct a measure of eventual school exposure, defined as the district-level average cumulative school construction exposure among later cohorts. I then interact this measure with indicators for 5-year maternal birth cohort bins while including district and cohort-bin fixed effects. Under the identifying assumption, cohorts born prior to major school

expansion should not exhibit systematically different trends across districts with higher versus lower eventual exposure.

Appendix Table B.1 reports the results for age at first birth, completed fertility, and child years of schooling. Across all outcomes, the interaction coefficients between cohort indicators and eventual exposure are small in magnitude and statistically insignificant. For age at first birth, the coefficients for the 1970 and 1975 cohorts are -0.125 and -0.145 , respectively, with large standard errors relative to the point estimates. Similarly, for completed fertility, the estimated interactions are close to zero and imprecisely estimated. Joint tests fail to reject the null hypothesis that the interaction terms are jointly equal to zero, with p-values of 0.867 and 0.423 for age at first birth and completed fertility, respectively. The final column presents an analogous exercise for children's years of schooling. The interaction coefficients are again statistically indistinguishable from zero, and the joint test yields a p-value of 0.810. This specification should be interpreted more cautiously because children observed among early maternal cohorts constitute a selected sample. Nonetheless, the absence of differential trends is broadly consistent with the identifying assumptions of the empirical design.

Overall, the evidence from these cohort-based pre-trend tests suggests that districts with greater eventual school expansion were not already experiencing systematically different trends in fertility or intergenerational outcomes prior to substantial exposure, supporting the interpretation of the main estimates as reflecting the causal effects of school construction.

2.7 Results

For the main outcome tables in the paper, I use two measures of school construction exposure. The first is a continuous measure capturing the cumulative number of schools

constructed per 1,000 school-age children in a woman's district during ages 6–15. The second is a binary indicator equal to one if exposure is above the cohort-specific mean. For the continuous measure, one additional school per 1,000 children corresponds to roughly an 11 percent increase relative to the sample mean exposure.

2.7.1 Effects on Women's Fertility Outcomes

Educational expansion may affect fertility behavior through several channels. Increased access to schooling may delay marriage formation, while additional education may also alter fertility preferences, bargaining power within households, and investments in child quality. In settings such as India, where marriage and childbearing remain closely linked, changes in marriage timing may translate directly into changes in fertility behavior over the life cycle. Earlier results in the paper show that school construction increased women's educational attainment and delayed marriage formation. This raises the question of whether educational expansion also affected subsequent fertility timing and completed fertility.

Figure 2.3 provides a descriptive view of fertility patterns across maternal birth cohorts by plotting age-specific fertility rates over the life cycle. Later cohorts exhibit systematically lower fertility throughout the prime childbearing ages, particularly during the early twenties when fertility is most concentrated among older cohorts. Fertility profiles also become somewhat flatter across cohorts, with fertility shifting modestly away from the youngest childbearing ages. At older maternal ages, fertility rates for younger cohorts do not fully converge to those of older cohorts, suggesting that delayed fertility does not completely offset reductions in total fertility. These descriptive patterns are consistent with both declining fertility levels and modest delays in fertility timing across cohorts.

Table 2.2 reports the estimated effects of school construction exposure on fertility outcomes. Increased school access significantly delays the timing of childbearing. A one-unit increase in school construction exposure increases age at first birth by 0.034 years. At the same time, school construction substantially reduces fertility levels. A one-unit increase in exposure reduces the number of children ever born by 0.068 children, corresponding to roughly a 19 percent decline relative to the sample mean. The results further suggest that these effects reflect persistent changes in lifetime fertility rather than temporary delays alone. Completed fertility among women aged 40 and above also declines significantly with school exposure, with an estimated reduction of approximately 6 percent relative to the sample mean.³ The estimates also indicate changes in fertility timing over the life cycle. Although school exposure increases age at first birth, age at last birth among women aged 40 and above declines modestly. In contrast, the interval between marriage and first birth remains small and statistically insignificant. Together, these patterns suggest that delayed fertility operates primarily through later marriage formation rather than through postponement of births within marriage. The decline in age at last birth is similarly consistent with lower completed fertility among exposed women.

Overall, the fertility results indicate that educational expansion affected not only women's schooling and marriage outcomes, but also subsequent household formation and demographic behavior over the life cycle. These fertility responses may themselves contribute to intergenerational effects if lower fertility and delayed childbearing increase investments in children's human capital.

³Appendix Table B.2 shows that the negative effects on completed fertility are broadly similar when restricting the sample to women aged 35 and above or 45 and above, although estimates become less precise at older age thresholds due to smaller sample sizes.

2.7.2 Intergenerational Effects on Children's Education

The effects of educational expansion may extend beyond women's own schooling, marriage, and fertility outcomes to influence investments in the next generation. Increased maternal education may improve children's educational attainment through several channels, including higher household resources, greater parental investments in child quality, changes in fertility behavior, and improved educational matching within households. The fertility results presented earlier are particularly suggestive of a quantity–quality trade-off, as school construction exposure reduced completed fertility while simultaneously delaying childbearing. This raises the question of whether these demographic changes translated into improved educational outcomes for children.

To examine these intergenerational effects, I estimate variants of equation (2.1), where exposure is defined based on the mother's childhood district and cohort. Table 2.3 reports the effects of school construction exposure on children's education outcomes among children aged 6–14. Outcome variables include years of schooling, a schooling gap measure defined as completed schooling minus age-appropriate grade-for-age, and an indicator for being behind the age-appropriate grade level. All specifications include district and maternal birth cohort fixed effects. Regressions for years of schooling additionally include child age fixed effects, while all specifications include a gender indicator for the child. The results indicate significant improvements in children's educational attainment. A one-unit increase in school construction exposure increases children's years of schooling by 0.034 years. Exposure also improves grade progression outcomes. Children of exposed mothers are 0.7 percentage points less likely to fall behind the age-appropriate grade level, corresponding to an elasticity of roughly -0.19.

Taken together with the fertility results, the estimates suggest that educational expansion affected not only women's own human capital accumulation, but also investments

in the next generation. The combination of lower fertility and improved child educational outcomes is consistent with a quantity–quality trade-off in which reductions in fertility are accompanied by greater investments in children’s schooling.⁴

2.7.3 Heterogeneity by Child Gender

While the baseline results show positive intergenerational effects on children’s education, an important question is whether these gains are distributed similarly across boys and girls. Gender disparities in schooling remain substantial in many developing country contexts, including India, and educational investments within households may differ systematically by child gender. Examining heterogeneity by gender therefore provides insight into whether the intergenerational effects of educational expansion reinforce or reduce existing gender inequalities in child human capital accumulation.

Table 2.4 examines heterogeneity by child gender by interacting school construction exposure with an indicator for female children. The coefficient on school exposure captures the effect for boys, while the interaction term measures the differential effect for girls relative to boys. The estimates indicate that school construction exposure improves educational outcomes for boys across all measures. Boys experience increases in years of schooling, reductions in schooling gaps, and lower probabilities of being behind the age-appropriate grade level. The interaction terms are statistically significant but small in magnitude. Relative to boys, girls experience slightly smaller gains in years of schooling and schooling progression outcomes. Nevertheless, the overall effects for girls remain positive across outcomes.

These results suggest that educational expansion generated broad intergenerational improvements in children’s education for both boys and girls, although gender gaps in

⁴Appendix Table B.3 shows that the results are qualitatively similar when expanding the sample to include children aged 6–18, although estimates become modestly larger in magnitude for some outcomes.

educational attainment and progression persist. The findings are therefore more consistent with unequal gains across gender rather than complete convergence in educational outcomes.

2.7.4 Parental Educational Composition and Children’s Education

The baseline results in Table 2.3 show that school construction exposure improves children’s educational outcomes, increasing years of schooling and reducing both schooling gaps and the probability of being behind grade level. The most direct mechanism through which these intergenerational effects are likely to operate is parental human capital. School construction increased women’s educational attainment and altered marriage market matching patterns, changing the educational composition of households. More educated parents may invest more in children’s schooling, possess greater educational and health knowledge, or provide home environments more conducive to learning. At the same time, changes in assortative matching may alter the joint distribution of parental education within households, potentially affecting child outcomes through household educational composition rather than maternal education alone. In practice, however, these channels are difficult to cleanly separate because parental education and marital sorting are themselves endogenous outcomes of school construction exposure. The analysis in this section therefore does not attempt a formal causal decomposition of mechanisms. Instead, it examines how the estimated relationship between school construction and child schooling changes once parental education and household educational composition are incorporated into the specification. This provides evidence on the extent to which the reduced-form intergenerational effects are closely associated with parental human capital and sorting patterns within households.

All specifications in Table 2.5 use children's years of schooling as the outcome and include district and maternal birth cohort fixed effects, along with child age fixed effects and an indicator for child gender. Mother's and father's years of education are standardized to have mean zero and standard deviation one. Standardization allows the interaction coefficients to be interpreted in comparable units and aligns naturally with the assortative matching framework, since relative educational position within the maternal and paternal education distributions may matter more for household formation and child outcomes than absolute years of schooling alone. In the estimation sample, one standard deviation corresponds to approximately five years of schooling for both mothers and fathers.

Column (1) includes additive measures of maternal and paternal education. Both are strongly positively associated with child schooling. A one standard deviation increase in maternal education is associated with approximately 0.04 additional years of child schooling, while the corresponding association for paternal education is substantially larger at approximately 0.13 years. Relative to the baseline specification in Table 2.3, the coefficient on school construction exposure declines meaningfully once parental education is included, although it remains positive and statistically significant. This pattern suggests that an important share of the reduced-form intergenerational relationship operates through parental educational attainment and household composition. Column (2) additionally includes an interaction between maternal and paternal education. The interaction term is negative and statistically significant, indicating that the marginal association between one parent's education and child schooling is smaller when the other parent is already highly educated. This pattern is inconsistent with strong complementarities or supermodularity in parental education. Instead, it suggests diminishing marginal returns to parental human capital within households. Put differently, increases in maternal education are most strongly associated with child schooling in households where

paternal education is relatively low, while the marginal association becomes weaker when fathers are already highly educated. Column (3) further allows these parental education interactions to vary with school construction exposure through a three-way interaction. The estimated triple interaction is positive and statistically significant, but economically small in magnitude. Similarly, the interactions between parental education and school construction exposure are quantitatively modest. Taken together, these estimates provide little evidence that school expansion substantially changes the relationship between maternal and paternal education in the production of child human capital. Because parental education is itself affected by school construction, these specifications should not be interpreted as a formal causal mediation analysis. Rather, the results indicate that much of the intergenerational relationship between school construction and child schooling operates through parental educational attainment and household educational composition, with limited evidence of additional changes in complementarities between parents.

Figure 2.4 provides a complementary visualization of these patterns by plotting predicted child years of schooling across different levels of maternal education separately for low, average, and high levels of paternal education. The figure illustrates the negative interaction estimated in Table 2.5. When paternal education is relatively low, increases in maternal education are strongly associated with higher levels of child schooling. In contrast, when paternal education is already high, the relationship between maternal education and child outcomes becomes substantially flatter and may even turn slightly negative at the upper end of the distribution. This pattern is more consistent with diminishing marginal returns to parental human capital within households than with strong complementarities between parents. At the same time, households with highly educated fathers maintain consistently higher predicted levels of child schooling overall, indicating persistent advantages associated with higher parental human capital.

2.7.5 Intergenerational Effects on Children's Health

Educational expansion may affect not only children's educational attainment, but also broader dimensions of child well-being. Increased parental schooling can improve health knowledge, healthcare utilization, nutrition practices, and investments in children, potentially generating intergenerational improvements in child health. In addition, the fertility results presented earlier suggest that school construction reduced fertility and delayed childbearing, both of which may alter parental investments per child. Examining child health outcomes therefore provides a complementary perspective on the broader intergenerational consequences of educational expansion.

Table 2.6 reports the effect of school construction exposure on child health outcomes using the same empirical specification as equation (2.1), applied to a sample of children under age five from the DHS KR data. The outcomes examined include height-for-age z-scores (HAZ), an indicator for stunting, a vaccination index, and institutional delivery. Height-for-age is a widely used measure of long-run nutritional status and chronic childhood health because it captures cumulative health and nutritional deficits experienced during early childhood. Unlike short-run measures such as weight-for-age, height-for-age reflects persistent investments in child health and is therefore particularly useful for studying longer-term intergenerational effects. Stunting is defined as an indicator equal to one if height-for-age falls below -2 standard deviations relative to WHO reference standards, capturing moderate or severe chronic undernutrition. The remaining outcomes capture healthcare utilization and access to formal health services. The vaccination index is constructed as the average of indicators for receipt of basic childhood vaccines, including BCG, DPT, polio, and measles vaccines. Institutional delivery is an indicator equal to one if the child was delivered in a public or private health facility rather than at home or in a non-institutional setting. Relative to anthropometric outcomes, these measures are

likely to depend more directly on access to healthcare infrastructure and public health service provision. Specifications for vaccination outcomes include child age fixed effects to account for the age-specific timing of immunization schedules, while gender indicators are included throughout.

The results suggest that school construction had modest but meaningful effects on some dimensions of child health. Exposure to school construction significantly increases children's height-for-age scores, indicating improvements in long-run nutritional status and early-life health investments. In contrast, the estimated effects on stunting are negative but small and statistically insignificant. This pattern suggests that educational expansion may have improved average nutritional outcomes without substantially reducing the prevalence of severe chronic undernutrition at the lower tail of the distribution. The results also indicate positive effects on institutional delivery. A one-unit increase in school construction exposure increases the probability of institutional delivery by approximately 0.009, corresponding to an elasticity of roughly 0.10 relative to the sample mean. This finding is consistent with increased healthcare utilization and greater engagement with formal medical services among exposed cohorts. By contrast, the estimated effects on vaccination outcomes are close to zero and statistically insignificant. One possible explanation is that vaccination coverage depends more heavily on the local supply and outreach of public health systems, limiting the extent to which increased maternal education alone translates into higher vaccination rates.

Taken together, the results suggest that the intergenerational effects of educational expansion extend beyond schooling outcomes to include some improvements in child health and healthcare utilization. However, the health effects appear more modest and selective than the corresponding effects on children's educational attainment, indicating that educational expansion alone may be insufficient to substantially improve all dimensions of child health in the absence of complementary health infrastructure and service

provision.

2.8 Conclusion

The evidence presented in this paper suggests that large-scale educational expansion in India generated significant intergenerational effects on education and health outcomes through changes in fertility behavior, parental human capital, and household formation. An additional school per 1,000 school-age children delayed fertility and reduced completed fertility over the life cycle, while also improving children's educational attainment and grade progression. In contrast, the effects on child health were more modest, with positive effects on height-for-age and institutional delivery but weaker effects on vaccination and stunting. School construction increased women's educational attainment and altered patterns of marital sorting, changing the distribution of parental human capital across households. At the same time, I find little evidence that educational expansion strengthened complementarities between mothers' and fathers' education in the production of child human capital. Instead, the relationship between maternal and paternal education exhibits diminishing marginal returns within households.

More broadly, the paper highlights the importance of studying educational expansion in general equilibrium household settings rather than focusing solely on individual-level returns to schooling. Changes in who becomes educated, who marries whom, and when households form may all influence the long-run distribution of human capital across generations. In future work, I hope to further examine how educational expansion interacts with labor markets and household formation to shape intergenerational mobility and household stratification over time.

2.9 Figures and Tables

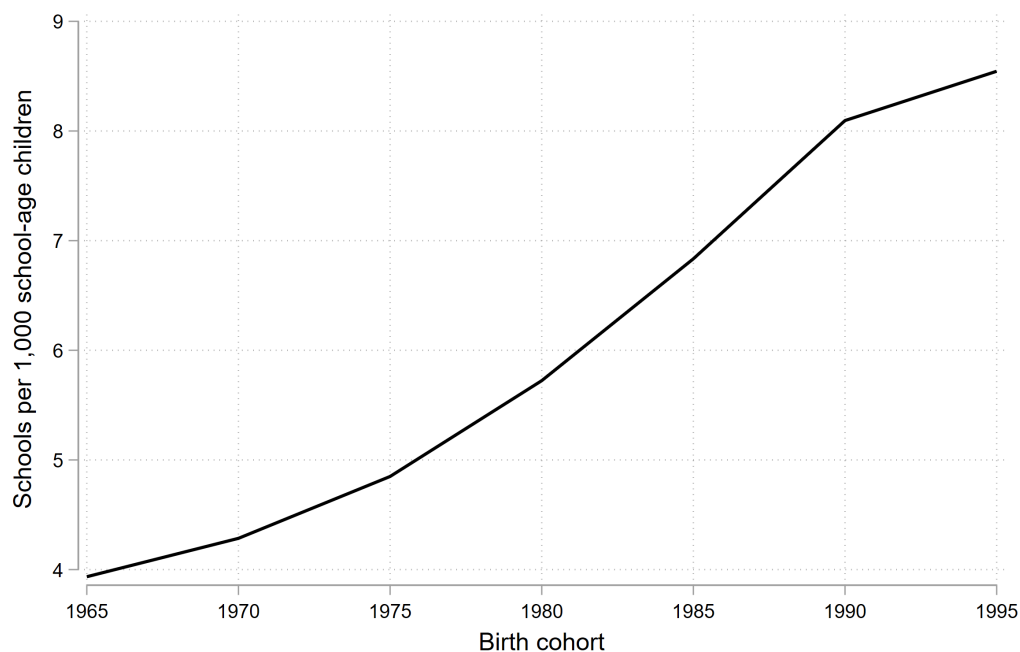


Figure 2.1: School Exposure by Women's Birth Cohort

Notes: The figure plots average cumulative school exposure during ages 6–15 by women's birth cohort. School exposure is measured as the cumulative number of schools constructed per 1,000 school-age children in a woman's district of residence during her school-age years.

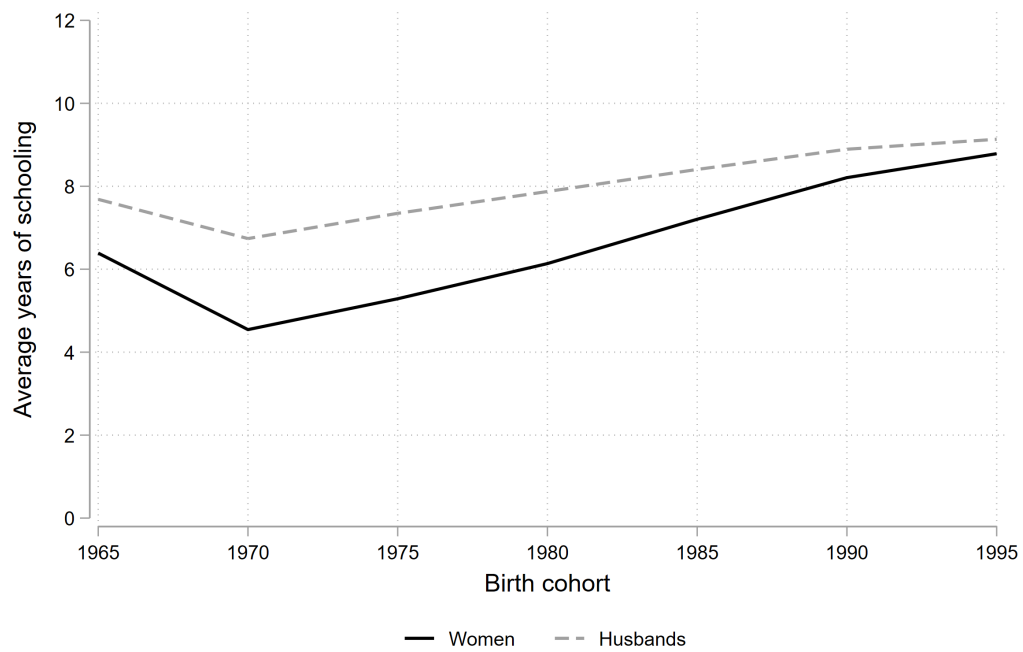


Figure 2.2: Average Years of Schooling among Married Couples by Women's Birth Cohort

Notes: The figure plots average years of schooling for women and their co-resident husbands by women's birth cohort. Women's educational attainment rises substantially across cohorts, narrowing the gender gap in schooling over time.

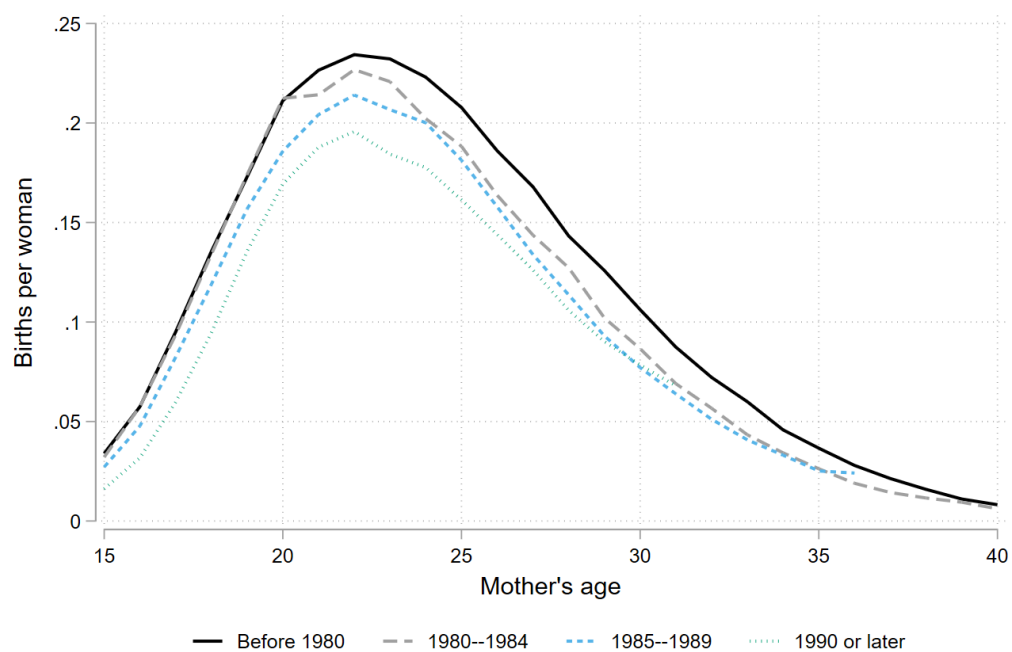


Figure 2.3: Age-Specific Fertility by Mother's Birth Cohort

Notes: Fertility rates are calculated as the weighted number of births occurring at each maternal age divided by the weighted number of women observed at that age. Cohorts are grouped by mother's year of birth. The denominator includes all women observed at a given age, including women with no births. Observations at later ages for younger cohorts are based on progressively smaller samples.

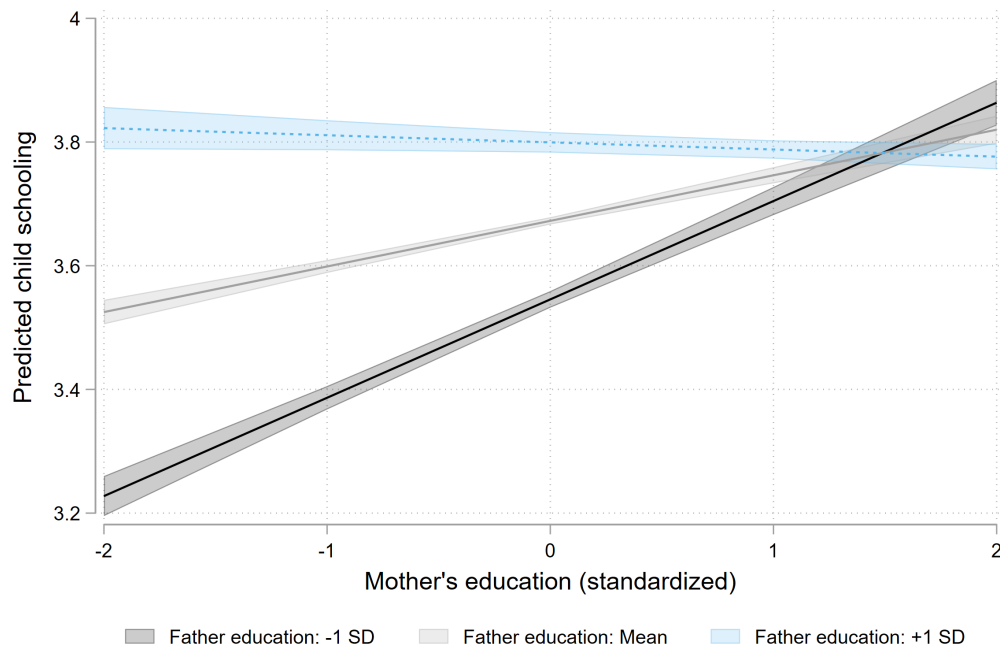


Figure 2.4: Predicted Child Schooling by Maternal and Paternal Education

Notes: The figure plots predicted years of child schooling from the specification including interactions between maternal and paternal education reported in Column 2 of Table 2.5. Mother's and father's education are standardized to have mean zero and standard deviation one. In the estimation sample, one standard deviation corresponds to approximately five years of schooling for both mothers and fathers. The horizontal axis plots maternal education continuously from -2 to $+2$ standard deviations, while separate lines correspond to paternal education evaluated at -1 standard deviation, the mean, and $+1$ standard deviation. Shaded areas represent 95 percent confidence intervals. Predictions are evaluated holding other covariates at their observed values.

Table 2.1: Summary Statistics

Variable	Mean	SD	N
<i>Panel A: Mother fertility outcomes sample</i>			
School exposure	6.638	5.611	146,388
Years of schooling	4.565	4.993	146,388
Age at first birth	21.113	4.341	141,074
Children ever born	3.073	1.657	146,388
Age at last birth	27.922	5.359	141,074
<i>Panel B: Child education outcomes sample</i>			
Child years of schooling	3.556	2.628	372,586
Schooling gap	-0.344	1.328	372,586
Behind grade level	0.354	0.478	372,586
Mother years of schooling	5.438	5.002	372,671
Father years of schooling	7.345	4.813	322,053
<i>Panel C: Child health outcomes sample</i>			
Height-for-age z-score	-1.312	1.816	151,541
Stunted	0.353	0.478	151,541
Vaccination index	0.817	0.312	89,367
Institutional delivery	0.861	0.346	170,241

Notes: Panel A reports summary statistics for the mother-level fertility sample restricted to women over 40 with completed fertility. Panel B reports summary statistics for the child education sample of children aged 6–14. Panel C reports summary statistics for the child health sample of children under age five. School exposure is the cumulative number of schools constructed per 1,000 school-age children in the mother’s district of residence during ages 6–15.

Table 2.2: Effect of School Construction on Fertility Outcomes

	(1) Marriage to first birth	(2) Age at first birth	(3) Age at last birth	(4) Children ever born	(5) Completed fertility
<i>Panel A: Continuous treatment</i>					
Schools built	-0.010 (0.007)	0.034** (0.014)	-0.276*** (0.071)	-0.068*** (0.008)	-0.037** (0.017)
Mean of treatment	6.258	6.255	4.626	6.383	4.627
Elasticity of outcome	-0.027	0.010	-0.047	-0.185	-0.058
<i>Panel B: Binary treatment</i>					
Schools built > mean	-0.021 (0.032)	0.248*** (0.066)	0.066 (0.124)	-0.176*** (0.031)	-0.069 (0.051)
Mean of treatment	0.164	0.164	0.148	0.166	0.148
Elasticity of outcome	-0.001	0.002	0.000	-0.012	-0.003
Mean of outcome	2.327	20.717	27.378	2.347	2.993
N	397,202	407,881	141,074	462,963	146,388

Notes: The table reports estimates of the effect of school construction exposure on fertility outcomes. The treatment variable in Panel A is the cumulative number of schools built per 1,000 school-age children in a woman's district of residence during ages 6–15. Panel B uses a binary indicator equal to one if exposure is above the sample mean. Age at first birth, the interval between marriage and first birth, and age at last birth are measured in years, while children ever born and completed fertility are measured as counts. Age at last birth and completed fertility are measured among women aged 40 and above. Regressions for children ever born additionally include mother age fixed effects to account for mechanical lifecycle variation in fertility. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment to the mean of the outcome. Standard errors are clustered at the district level.

Table 2.3: Intergenerational Effect of School Construction on Child Education Outcomes

	(1) Years of schooling	(2) Schooling gap	(3) Behind grade level
<i>Panel A: Continuous treatment</i>			
Schools built	0.034*** (0.006)	0.030*** (0.006)	-0.007*** (0.002)
Mean of treatment	8.993	8.993	8.993
Elasticity of outcome	0.086	-0.791	-0.188
<i>Panel B: Binary treatment</i>			
Schools built > mean	0.011 (0.023)	0.009 (0.023)	-0.000 (0.010)
Mean of treatment	0.336	0.336	0.336
Elasticity of outcome	0.001	-0.009	-0.000
Mean of outcome	3.556	-0.344	0.354
N	372,586	372,586	372,586
Child age FE	Yes	No	No
Child gender FE	Yes	Yes	Yes

Notes: The table reports estimates of the intergenerational effect of school construction exposure on children's education outcomes. The treatment variable in Panel A is the cumulative number of schools built per 1,000 school-age children in the mother's district of residence during ages 6–15. Panel B uses a binary indicator equal to one if exposure is above the cohort-specific mean. The sample includes children aged 6–14 at the time of the survey. Outcome variables include years of schooling, schooling gap, and an indicator for being behind the age-appropriate grade level. Regressions for years of schooling include child age fixed effects, while all specifications include a gender indicator for the child. All specifications include district and maternal birth cohort fixed effects. Standard errors are clustered at the district level. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment to the mean of the outcome.

Table 2.4: Intergenerational Effect of School Construction on Child Education: Heterogeneity by Gender

	(1) Years of schooling	(2) Schooling gap	(3) Behind grade level
Schools built	0.035*** (0.006)	0.032*** (0.006)	-0.008*** (0.002)
Female child	0.119*** (0.012)	0.119*** (0.012)	-0.042*** (0.004)
Schools $\times$ Female child	-0.003** (0.001)	-0.003** (0.001)	0.001*** (0.000)
Mean of outcome	3.556	-0.344	0.354
N	372,586	372,586	372,586
Child age FE	Yes	No	No

Notes: The table reports heterogeneous effects of school construction exposure on children's education outcomes by child gender. The treatment variable is the cumulative number of schools built per 1,000 school-age children in the mother's district of residence during ages 6–15. The sample includes children aged 6–14. Regressions include an interaction between school construction exposure and an indicator for female child. The coefficient on “Schools built” represents the effect for boys, while the interaction term captures the differential effect for girls. Child age fixed effects are included only for the years of schooling outcome. For schooling gap and behind-grade indicators, age is accounted for in the construction of the dependent variable and is therefore not included as a control. All specifications include district and maternal birth cohort fixed effects. Standard errors are clustered at the district level.

Table 2.5: Intergenerational Effect of School Construction on Child's Years of Schooling: Parental Education Channel

	Child years of schooling		
	(1) Additive parental education	(2) Parental education interaction	(3) Interaction with school exposure
Schools built	0.026*** (0.006)	0.026*** (0.005)	0.026*** (0.005)
Mother education	0.036*** (0.005)	0.074*** (0.005)	0.076*** (0.010)
Father education	0.134*** (0.008)	0.127*** (0.007)	0.156*** (0.012)
Mother education $\times$ Father education		-0.085*** (0.005)	-0.101*** (0.009)
Mother education $\times$ Schools built			-0.000 (0.001)
Father education $\times$ Schools built			-0.004*** (0.001)
Mother $\times$ Father $\times$ Schools built			0.002** (0.001)
Mean of outcome	3.581	3.581	3.581
N	321,998	321,998	321,998
Child age FE	Yes	Yes	Yes
Child gender FE	Yes	Yes	Yes

Notes: The table reports estimates of the intergenerational effect of school construction exposure on children's years of schooling. The treatment variable is the cumulative number of schools built per 1,000 school-age children in the mother's district of residence during ages 6–15. The sample includes children aged 6–14 at the time of the survey. Mother's and father's education are standardized to have mean zero and standard deviation one. Column (1) includes additive measures of parental education. Column (2) additionally includes an interaction between mother's and father's education. Column (3) further allows the parental education interaction to vary with school construction exposure. All specifications include district and maternal birth cohort fixed effects, child age fixed effects, and an indicator for child gender. Standard errors are clustered at the district level.

Table 2.6: Intergenerational Effect of School Construction on Child Health Outcomes

	(1) Height-for-age (z-score)	(2) Stunted (z-score < -2)	(3) Vaccination index	(4) Institutional delivery
<i>Panel A: Continuous treatment</i>				
Schools built	0.018** (0.009)	-0.002 (0.002)	-0.001 (0.002)	0.009*** (0.002)
Mean of treatment	10.556	10.556	10.568	10.479
Elasticity of outcome	-0.147	-0.057	-0.007	0.104
<i>Panel B: Binary treatment</i>				
Schools built > mean	0.058 (0.043)	-0.014 (0.011)	-0.001 (0.009)	0.015* (0.009)
Mean of treatment	0.356	0.356	0.353	0.352
Elasticity of outcome	-0.016	-0.014	-0.000	0.006
Mean of outcome	-1.312	0.353	0.817	0.863
N	151,540	151,540	89,367	164,195
Child age FE	No	No	Yes	No
Child gender FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates of the intergenerational effect of school construction exposure on child health outcomes. The treatment variable in Panel A is the cumulative number of schools built per 1,000 school-age children in the mother's district of residence during ages 6–15. Panel B uses a binary indicator equal to one if exposure is above the sample mean. The sample consists of children under age five. Height-for-age is constructed using WHO-standardized z-scores, rescaled by dividing by 100. Stunting is an indicator equal to one if height-for-age is below -2. The vaccination index is the average of indicators for receipt of basic vaccines. Institutional delivery is an indicator equal to one if the child was delivered in a public or private health facility. Child age fixed effects are included for vaccination outcomes. A gender indicator is included in all specifications. All regressions include district and maternal birth cohort fixed effects. Standard errors are clustered at the district level. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment to the mean of the outcome.

Chapter 3

Workfare, Marriage and Women's Fertility: Evidence from India

Abstract

I estimate the effect of an expansion in employment opportunities available to women in India on their marriage and fertility choices. While overall female labor force participation in India has stagnated at very low levels, women's participation in the workfare program, the National Rural Employment Guarantee Act, is fairly high. Using the staggered adoption of this workfare program across Indian districts for identification, I estimate whether the program had an effect on rates of marriage, fertility and sex ratio at birth. In my analysis, conditional on district covariates interacted with time in an age-adjusted model, the introduction of the workfare program led to 7.6 fewer marriages per 1000 women (3.6 fewer marriages based on the heterogeneity-robust estimator). Conceptions ending in live births per 1000 women fell by 19.1 in the 20-24 age group but rose by 15.5 in the 25-29 age group, indicating a shift towards later childbearing which may potentially reduce overall fertility in the long-run.

3.1 Background

3.1.1 Female labor Force Participation in India

Over the past few decades, India has experienced rapid economic growth, averaging 6.2 percent between 1990 and 2019. But despite this growth, the female labor force participation rate, defined as the percentage of working age (15-64 years old) women engaged in paid employment outside the household, has remained low and even fallen below 30 percent in recent years. Despite rising education levels among women, many women opt out of the paid labor force.

There have been a number of reasons posited for why Indian women are conspicuously absent from the labor force. Low labor force participation may reflect women's individual preferences or household preferences for prioritizing domestic work over outside employment. Low participation may also be a rational response to wage discrimination in the labor market. Patriarchal social norms also restrict women's mobility and these may be internalized by women. The lack of suitable outside options further weakens the intra-household bargaining position of women.

There may also be a mismatch between the type of new employment opportunities being created by economic growth and the actual skills possessed by working age women. Indian economic growth in recent decades is largely attributable to the growth of the services sector so there is a lack of suitable jobs for those transitioning out of agriculture. National Sample Survey (NSS) data indicates that there is also a skills gap between men and women, so the latter are more likely to be constrained by the unavailability of unskilled work opportunities. However, according to NSS data, over 30 percent of women who are engaged primarily in domestic activities and not currently counted in the labor force expressed a strong willingness to work if suitable local jobs were available.

The introduction of the National Rural Employment Guarantee Act (NREGA) pro-

vides exactly such feasible employment opportunities for women. The NREGA is an Indian workfare program that guarantees 100 days of employment per year to any household that demands it. The nature of the work consists of unskilled manual labor building durable assets within 5 kilometers of the worker's residence. The program also has a wage parity policy which may further induce participation by women who typically face large gender wage gaps in the private sector. Under NREGA, a third of employment is reserved for women, but in many states, actual employment outstrips this which is unsurprising given the latent demand by women for locally available, casual work. In contrast to the national labor market, which is comprised of only 22 percent women overall, 52 percent of NREGA workers were female in 2013-14.

3.1.2 Female Employment and Fertility Choices

Not only does the low female labor force participation rate translate to a loss of productive efficiency due to the misallocation of talent, it may also have important consequences for household-level decision-making including women's marriage and fertility choices. Demographic transition is regarded as an essential component of economic development. The quality-quantity tradeoff entails that as fertility falls, investment in children rises. This increase in the human capital stock of future generations fosters economic development. In turn, fall in fertility is itself theorized to be an outcome of economic growth. But it's useful to unpack the mechanisms of this latter process. In particular, I plan to focus on the potential impact of an increase in labor market opportunities for women on marriage and fertility choices.

An improvement in women's outside options may result in delayed age of marriage and lowered fertility. These changes could be mediated by increased female autonomy, particularly over contraception; a higher opportunity cost of childcare; and less need to

get married and have children, particularly sons, to contribute to the household income. The last factor might be especially pertinent for understanding fertility choices in this context. A preference for sons over daughters is well-documented in India. The worsening sex ratio is indicative of this latent preference as parents selectively abort females or continue to have children in order to have the desired number of males.

While there are many socio-cultural reasons for this phenomenon of son preference, a proximate economic cause is the fact that sons can work and contribute to the household's income while daughters marry and migrate away due to norms of patrilocality. Reinforcing this mechanism are gender-discriminatory cultural norms that restrict young women's mobility and labor force participation. An expansion in women's work opportunities may alleviate the economic need to have sons, but desire for a given family size and increased household income could result in sex-selective abortion if son-biased preferences are invariant. An open empirical question is how an increase in women's labor force participation would affect these fertility patterns and preferences.

3.2 Literature Review

There is a large literature drawing on long-run historical data from developed countries on the steady rise of women's labor force participation and effect on fertility patterns.

Voigtländer and Voth (2013) document that in the 14th century, the Black Death created land abundance which led to a shift towards animal husbandry where women had a comparative advantage and were often required to remain celibate as part of the terms of employment. In response to better labor market prospects, women delayed marriage which subsequently lowered fertility. The regions west of the "Hajnal line" which stretched from St. Petersburg to Trieste were characterized by this "European Marriage Pattern". For instance, the average female age at marriage in England and

Germany in the 17th century was 25-26, with a significant fraction of women never marrying. An estimated 25 to 40 percent of all possible births were avoided in Europe during this period. According to Voigtländer and Voth, it was fertility control which predated the growth in per-capita incomes in Europe and allowed the transition from “Malthus to Solow” to occur.

Greenwood, Seshadri and Vandenbroucke (2005) posit that among US women, the secular decline in fertility over the last 200 years is due to the relentless rise in real wages that increased the opportunity cost of having children, and that the mid-20th century baby boom was due to household technological improvements that lowered the direct costs of having children. Doepke, Hazan and Maoz (2013) suggest a different mechanism for the US baby boom where the increase in women's labor force participation during the Second World War was a major factor. The cohort of women who participated in the labor market and accumulated job experience during the war were able to crowd out younger women in the labor market post-war who lacked that experience. The authors argue that many of these younger women were crowded out of the labor force and decided to start having children earlier instead.

Overall, the relationship between women's labor force participation and economic development is described by Goldin (1994) to be U-shaped. At low levels of development, women have high employment rates, primarily in agriculture or self-employment. As incomes rise, women leave the labor market, in part because of the stigma attached to women engaging in manual labor in industrial jobs. But with greater development, and increases in women's human capital, white-collar service sector opportunities become available and more women enter the labor force.

It's possible that in developing countries, women's agricultural or self-employment work is under-measured in official survey data unlike paid wage labor outside the household. Time use surveys could help rectify this to some extent but data availability is

lacking. It is notable that in India, over half the workers in the rural workfare program are women. More recent studies from developing countries look at the effects of the growth in specific manufacturing sector jobs in the garment sector and service sector jobs in business process outsourcing that may be especially amenable to women.

The most relevant papers on the relationship between female labor force participation and fertility in the South Asian context are Heath and Mobarak (2015), Sivasankaran (2014), and Jensen (2012).

Heath and Mobarak (2015) study the effects of the growth of the garment industry in Bangladesh (where nearly 80 percent of workers are young women) on fertility, age at marriage, and girls' educational attainment. The authors estimate a difference-in-differences specification using spatial variation in the proximity of girls' village to a garment factory and temporal variation in years of lifetime exposure. For the effect on education, a comparison of girls with their brothers is also used in a triple difference. Girls that had factory exposure during the critical ages of 10 to 23 are 13 percentage points more likely to work outside the home than girls of the same age in control villages. Discrete time hazard models indicate that the average exposure of 6.4 years to the garment industry lowers probability of marriage in that year by 0.3 percentage points (28 percent lower) and lowers probability of childbearing by 0.23 percentage points (29 percent lower). The effects are driven by the younger age group who are unlikely to be directly employed by the sector but are more likely to be induced to stay in school in expectation of higher returns to cognitive skills in the garment sector.

Sivasankaran (2014) finds similar effects of textile mill employment in South India. Using variation in tenure created by unanticipated changes to fixed term contracts, this study focuses on the intensive margin of duration of employment. Women who worked longer delayed marriage, without any adverse effect on eventual spousal quality. This also led to fall in women's stated desired fertility. Survey data suggests these effects may

be attributed to an increase in female empowerment and autonomy.

Jensen (2012) conducts a randomized control trial in India which allows for clean identification of these effects. Young women in treatment villages are provided with three years of recruiting services for business process outsourcing jobs in rural areas near Delhi, India. Women aged 18 to 24 in treated villages were 4.6 percentage points more likely to have a BPO job, compared to a control mean of 0. The treatment also increased women's stated desire to work before and after marriage, but not after childbirth. Women aged 15 to 21 at baseline from treatment villages were 5.1 percentage points less likely to get married and 5.7 percentage points less likely to have given birth during the three-year period of the study. This suggests that women are more likely to delay childbearing in order to keep working. Primary survey data allows Jensen to conclude that the delay in marriage and childbirth is driven by women's own participation in the labor market due to higher returns. It is not due to other mechanisms such as young women delaying marriage whose working mothers have more bargaining power or need daughters to stay at home to help with household chores. It should be noted that the low take-up rate of the intervention (due to the higher education requirements for BPO job eligibility) raises some concerns over external validity. However, it does lend support to Goldin's contention that increasing girls' education may be key to increasing their later participation in the labor force.

There are also a number of studies in developing countries with strong gender norms that examine the impact of plausibly exogenous changes in the value of women's work on fertility patterns, especially the child sex ratio.

Qian (2008) finds that an exogenous increase in relative female income as a share of household income due to agricultural reforms in China improved the survival rate for girls. The reforms increased the returns to cultivation of tea, in which women had a comparative advantage, and returns to orchards, where men had the advantage. Qian uses a difference-

in-differences strategy comparing cohorts exposed to the reforms to earlier cohorts and comparing tea-planting counties proxying for an increase in female income (holding male income fixed) to non-tea counties. A similar difference-in-differences model is estimated for orchard counties with an increase in relative male income. She finds that during the early 1980s, increasing annual adult female income by USD 7.70 (10 percent of average rural annual household income) while holding adult male income constant increased the fraction of surviving girls by one percentage point. Conversely, increasing male income while holding female income constant decreased survival rates. Qian interprets the results in a non-unitary model with intrahousehold bargaining where women are more likely than men to favor equal treatment of boys and girls resulting in more girls surviving when their bargaining position improves. As there is no effect of agricultural planting or harvesting seasons on fertility, Qian concludes that the results are not likely to be driven by changes in the opportunity cost of women's time.

Carranza (2014) examines the same relationship in India using the fact that deep tillage of soil, which reduces the need for female labor in transplanting, fertilizing and weeding, is only possible in loamy and not in clayey soils. The identification strategy relies on the assumption that the exogenous soil texture fractions affect the rural child sex ratio only through the impact of deep tillage on relative female employment, conditional on state fixed effects and other controls. Both reduced form and instrumental variable estimates (using the exogenous fractions of loamy and clayey soils as instruments) indicate that loamy soils lower women's labor force participation as well as female-to-male sex ratios. Based on the IV estimates, removing the rural deficit of girls in India would require a 5.8 percentage point increase in the average share of female agricultural laborers in the rural labor force. Carranza suggests that the results can be explained by more household resources being allocated to female children in response to a higher return when female earnings and market opportunities improve. While a bargaining model where women

have more egalitarian preferences as in Qian would also fit Carranza's empirical findings, survey data from India indicates that mothers exhibit even stronger son preference than fathers. Also, while Qian uses data from an earlier time period before the advent of ultrasound technology, Carranza uses more recent data and in a related paper, shows that sex-selective abortion is a proximate determinant of lower child sex ratios in loamier soils.

Another mechanism that could strengthen women's relative bargaining power is through increasing their control over financial resources, especially their own earnings. This is examined by Field et al. (2021) in a randomized control trial where women's wages from the Indian NREGA workfare program are deposited into their own bank account instead of their husband's. The authors find that this not only increased women's participation in private and public sector work, but also changed social norms, especially among husbands who were most strongly opposed to women working. While this study focuses on already married women's labor supply, it's possible that a similar intervention could also change the labor supply as well as marriage and fertility choices of single women.

Empirical studies on the impact of NREGA such as Azam (2012), Imbert and Papp (2015), Muralidharan, Niehaus and Sukhtankar (2018), Imbert and Papp (2020), Zimmermann (2020) and Cook and Shah (2022) primarily focus on the effects of the program on output, employment, wages, and rural-urban migration. Specifically, Azam (2012) estimates that the introduction of NREGA led to a 2.4% increase in female labor force participation and 8.3% increase in women's wages. Other papers such as Afridi, Mukhopadhyay and Sahoo (2016) and Shah and Steinberg (2019) have focused on impacts of the program rollout on human capital investment in children. The effects of the program on women's marriage and fertility choices have not been previously examined.

3.3 Data

The NREGA was rolled out across districts in phases as shown in Figure 3.1. The program was implemented in 200 districts around April 2006. In April 2007, a further 130 districts were added, and in April 2008, the program became available in all remaining districts. Data on NREGA program implementation dates across 624 districts is publicly available through India's Ministry of Rural Development. I use this data to create a binary indicator for a district being treated (i.e. NREGA being operational) during a given time period.

I use marriage and fertility data from India's Demographic and Health Survey (DHS), a large-scale nationally representative survey of women aged 15-49 years. While using earlier rounds of the DHS conducted during the NREGA rollout from 2006-08 would be preferable, only DHS 2015-16 includes district identifiers which are required for matching with the NREGA district rollout data.

After dropping urban women, the dataset consists of interviews with 494,861 rural women of reproductive age (15-49 years) about their marriage, birth and health outcomes. I construct retrospective marriage and fertility histories using date of first cohabitation (equivalent to marriage in rural India) and dates of child birth. Dates of child birth are used to determine dates of conception, assuming a gestation period of nine months before birth. A limitation is that this only includes conceptions that end in live births, whereas the treatment could affect miscarriages and abortions. However, this is a common issue in demographic studies using such data.

As I am specifically interested in the effect of NREGA introduction, I consider the time window April 2003 to March 2011, approximately three years before the implementation of the program in any districts and three years post implementation in all districts. I drop women who were too young to be of reproductive age during the time window I

consider. For the marriage outcome, I drop women already married before April 2003. I construct a woman-year panel that has one observation per woman for each year of her observed reproductive period, defined as starting at age 15, within the 8-year time window. I drop observations where the woman was younger than 15 in that year. This gives me a sample of 408,414 women and 2,770,312 women-years. I create dummy variables indicating marriage or conception in a given district-year. I also create a dummy for number of males conceived. These are used for my outcome variables – marriage rate (number of marriages in given district-year per 1000 women), conception rate (number of live births conceived in given district-year per 1000 women), and sex ratio at birth (number of males conceived in given district-year per 100 females conceived). I also use data on mother's own date of birth to construct age-specific rates. Figure 3.2 shows the mean rates of marriage, conceptions and sex ratio at birth during the time window considered.

3.4 Empirical Strategy

I leverage the staggered adoption of the NREGA to estimate a difference-in-differences specification. Identification relies on the assumption that in the absence of NREGA, the districts that received the program in a specific phase did not have systematically different time patterns in the outcome variables.

A concern for identification is that allocation of the program to districts was based on the government's "backwardness" rankings which were based on pre-program levels of agricultural wages, percent of scheduled caste/scheduled tribe, and agricultural productivity. Imbert and Papp (2015), Shah and Steinberg (2019) and Cook and Shah (2022) rely on the assumption that the program was rolled out based on static characteristics of the districts, rather than underlying trends in fertility. I follow these papers and also

show results which include district-level controls interacted with time fixed effects. Since data on the variables used for constructing backwardness rankings is not available for all districts, the sample size is reduced when controls are included.

3.5 Results

3.5.1 Two-Way Fixed Effects

First, I estimate an event study with district and year fixed effects.

$$y_{idt} = \alpha + \sum_{j=2}^J \beta_j (\text{Lag } j)_{dt} + \sum_{k=1}^K \gamma_k (\text{Lead } k)_{dt} + \mu_d + \lambda_t + \varepsilon_{idt} \quad (3.1)$$

$(\text{Lag } J)_{dt} : 1[t \leq \text{Event}_d - J]$

$(\text{Lag } j)_{dt} : 1[t = \text{Event}_d - J] \text{ for } j \in \{1, \dots, J - 1\}$

$(\text{Lead } K)_{dt} : 1[t \geq \text{Event}_d + K]$

$(\text{Lead } k)_{dt} : 1[t = \text{Event}_d + k] \text{ for } k \in \{1, \dots, K - 1\}$

μ_d : District fixed effects

λ_t : Time (Year) fixed effects

y_{idt} : Outcomes (Marriage rate; conception rate; sex ratio at birth)

Table 3.1 shows estimates from the above two-way fixed effects model and Figure 3.3 shows estimates from the event study. The introduction of NREGA in a district increases the number of marriages per 1000 women by 0.215, the conception rate per 1000 women by 1.636 and the sex ratio (number of male births per 100 female births) by 1.4. None of these effects are statistically significant.

To get a better match on pre-trends, I also estimate this event study regression with age-adjusted marriage and fertility rates, which take into account whether a woman in

district d in year t married or conceived a live-born child when she turned age a .

$$y_{idta} = \alpha + \sum_{j=2}^J \beta_j (\text{Lag } j)_{dt} + \sum_{k=1}^K \gamma_k (\text{Lead } k)_{dt} + \mu_{da} + \lambda_{ta} + \varepsilon_{idta} \quad (3.2)$$

Table 3.2 and Figure 3.4 show estimates with age-adjusted rates. The introduction of NREGA in a district lowers the number of marriages per 1000 women by 1.279, raises the conception rate per 1000 women by 1.355 and lowers the sex ratio (number of male births per 100 female births) by 0.991. However, none of these effects are statistically significant.

Figure 3.5 shows estimates (and 95% confidence intervals) by 5-year binned age groups but the effects are not significant.

I also estimate the age-adjusted model with district controls interacted with time fixed effects in Table 3.3 and Figure 3.6. The introduction of NREGA in a district lowers the number of marriages per 1000 women by 7.557, raises the conception rate per 1000 women by 1.641 and raises the sex ratio (number of male births per 100 female births) by 7.846. The effect on the marriage rate is statistically significant at the 5% level.

Splitting the data into 5-year binned age groups, I observe declining marriage rates upto age 35, though only the 30-34 group shows a statistically significant decline of 25.629 (table 3.4 & Figure 3.7). Conceptions decline by 19.142 in the 20-24 age group but increase in the 25-29 age group, indicating later childbearing (table 3.5 & Figure 3.7). This may potentially result in many births being averted in the long-run.

I also run another sub-group analysis by estimating effects on women in the bottom 25% (table 3.6 & Figure 3.8) and top 25% of the household wealth distribution (table 3.7 & Figure 3.9). The effects are not statistically significant. But it should be noted that the wealth index is post-treatment data from the 2015-16 DHS survey, so it may be endogenous.

3.5.2 Heterogeneous Treatment Effects

I check for robustness to treatment effect heterogeneity across districts and over years by examining negative weights in the two-way fixed effects regressions. I find that for the marriage outcome regression, 990 ATTs have negative weights out of 2402, summing to -0.52 whereas for births, 990 ATTs have negative weights out of 2402, summing to -0.44. The large number of negative weights implies that the two-way fixed effects estimator is biased.

Hence, I use the estimator by de Chaisemartin and D'Haultfœuille (2020) which is robust to heterogeneous treatment effects and allows for estimation of dynamic effects. Estimates from regression with district and year fixed effects are shown in Table 3.8, event study estimates Figure 3.10. The average effect on the marriage rate is -3.620 and significant at the 10% level. The effect on the conception rate is 1.729 and sex ratio is -1.947 but not statistically significant. It is not possible to compute age-specific rates with this estimator, which restricts my model to the inclusion of only district and time fixed effects.

3.5.3 Conclusion

In the models without district controls, I find imprecisely estimated null effects. Null effects may be reasonable in this context since only a small percentage of the population is employed through NREGA, even though the entire adult rural population is eligible. Implementation problems and lack of awareness also reduced take-up in the initial stages of the program. General equilibrium effects through spillover on private sector employment and wages may also take time to emerge.

However, conditional on the inclusion of district covariates interacted with time fixed effects in an age-adjusted model, I find that the introduction of the workfare program

led to 7.6 fewer marriages per 1000 women. With the de Chaisemartin & D'Haultfœuille estimator, the average effect is 3.6 fewer marriages per 1000 women. Conceptions ending in live births per 1000 women fell by 19.1 in the 20-24 age group but rose by 15.5 in the 25-29 age group, indicating later childbearing which may potentially reduce overall fertility in the long-run. A longer term analysis of the effect of female labor force participation on fertility may be able to confirm this.

3.6 Figures and Tables

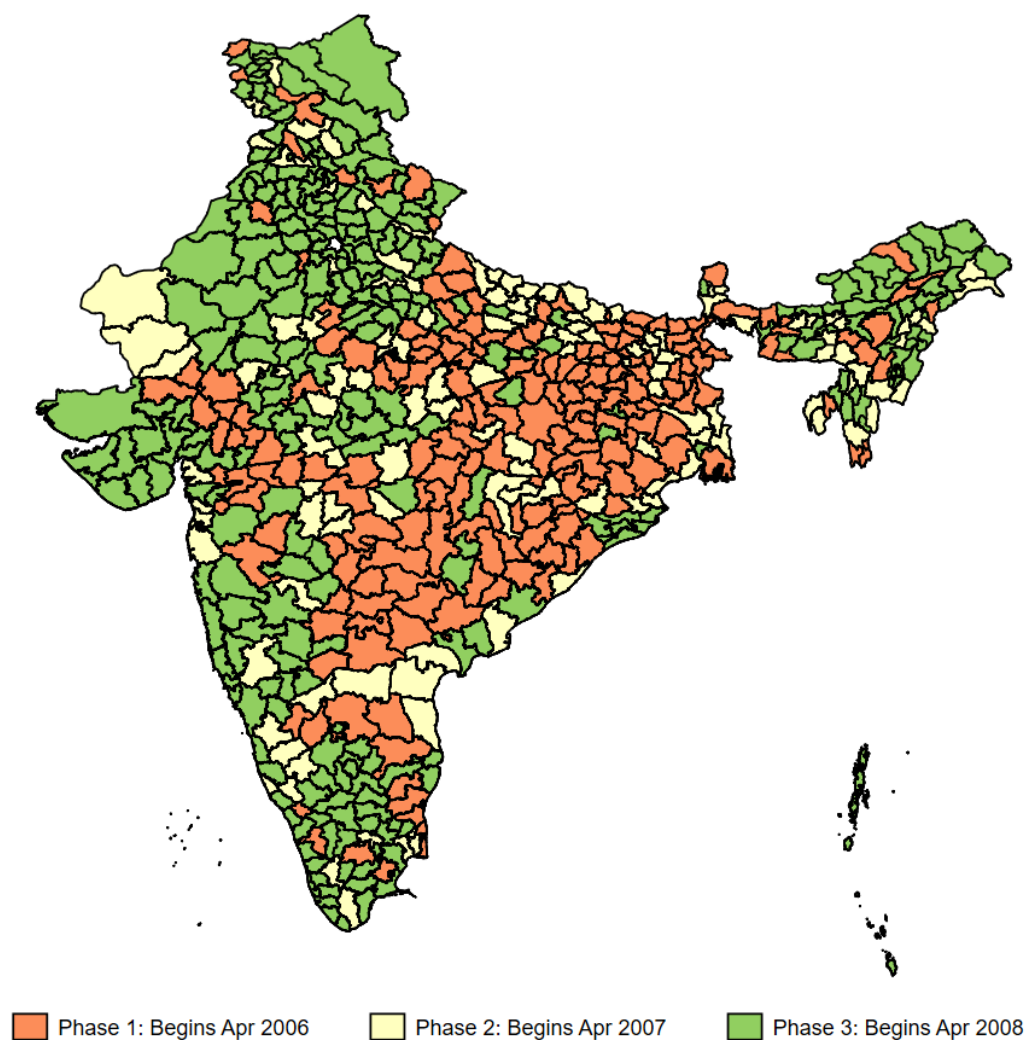


Figure 3.1: NREGA Implementation Phases

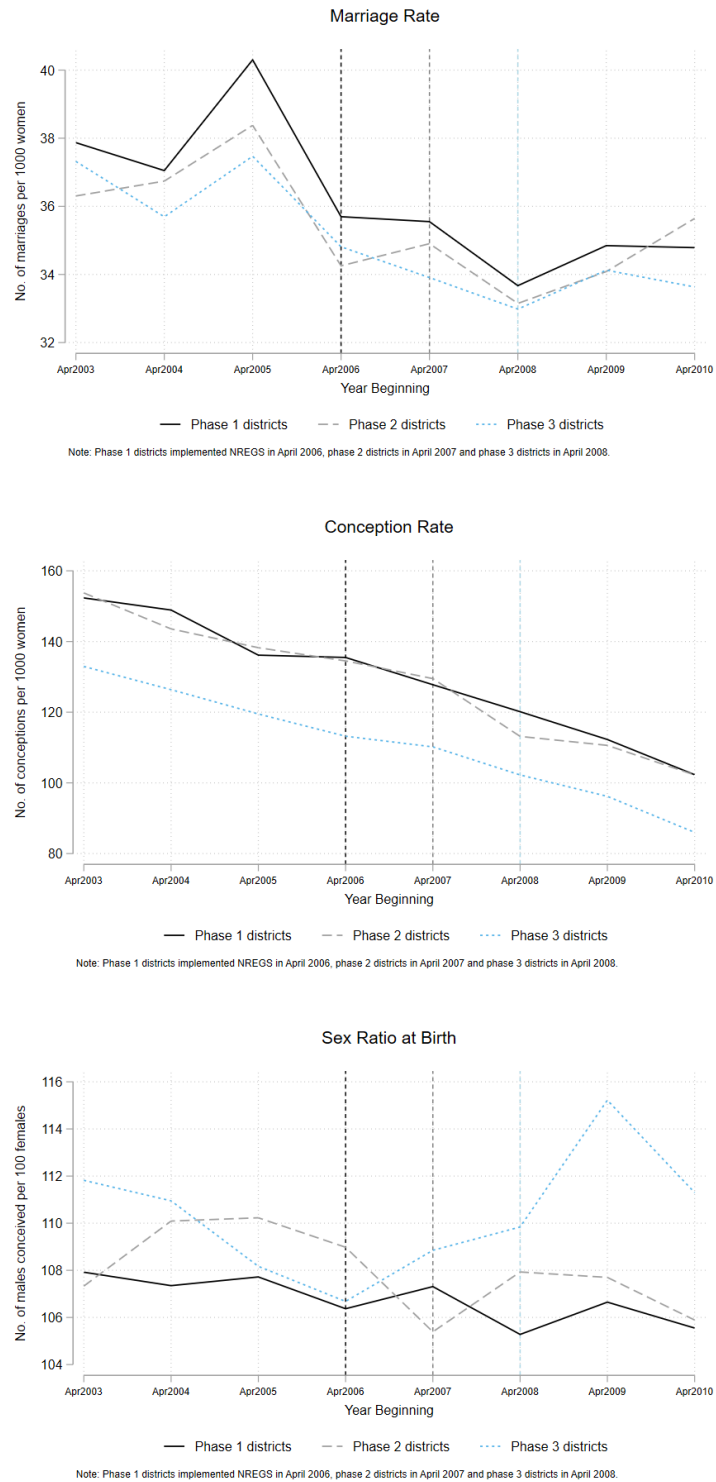


Figure 3.2: Mean Rates of Marriage, Conception, and Sex Ratio at Birth

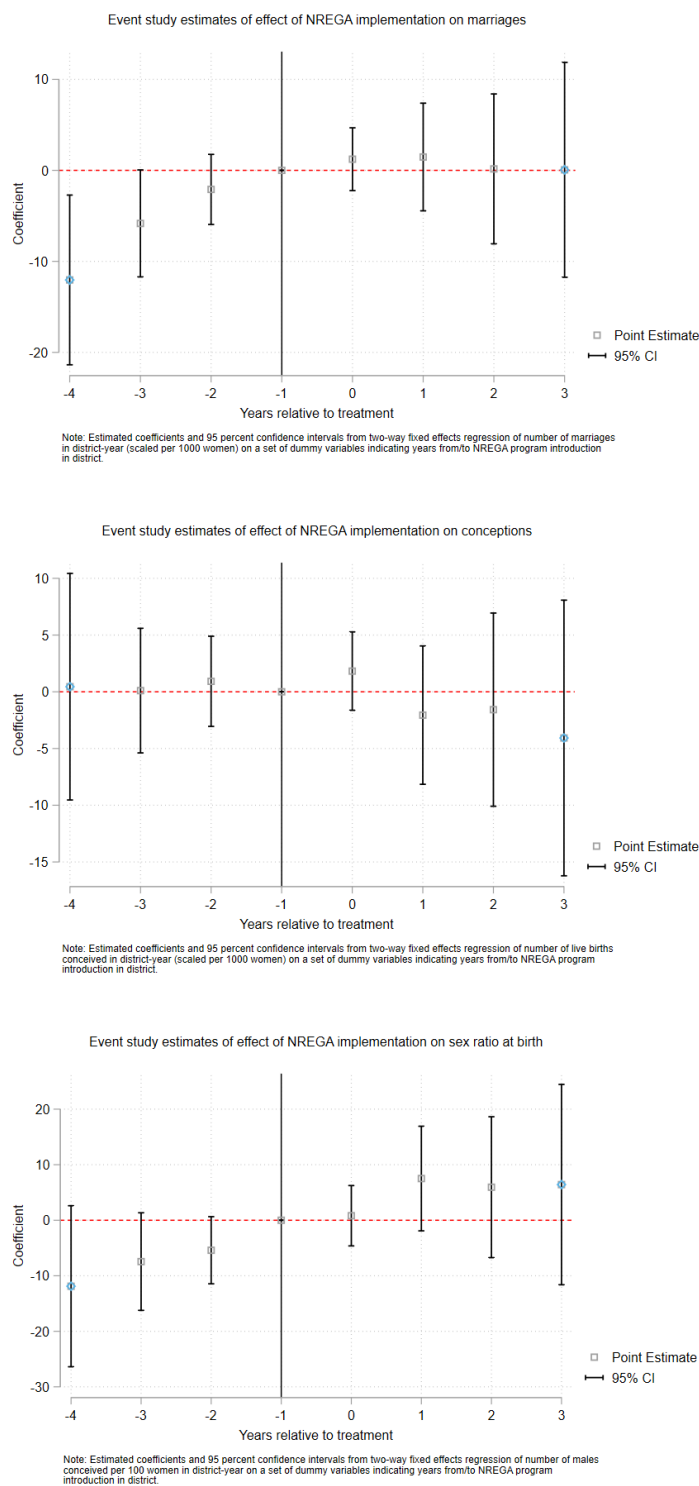


Figure 3.3: Event Study

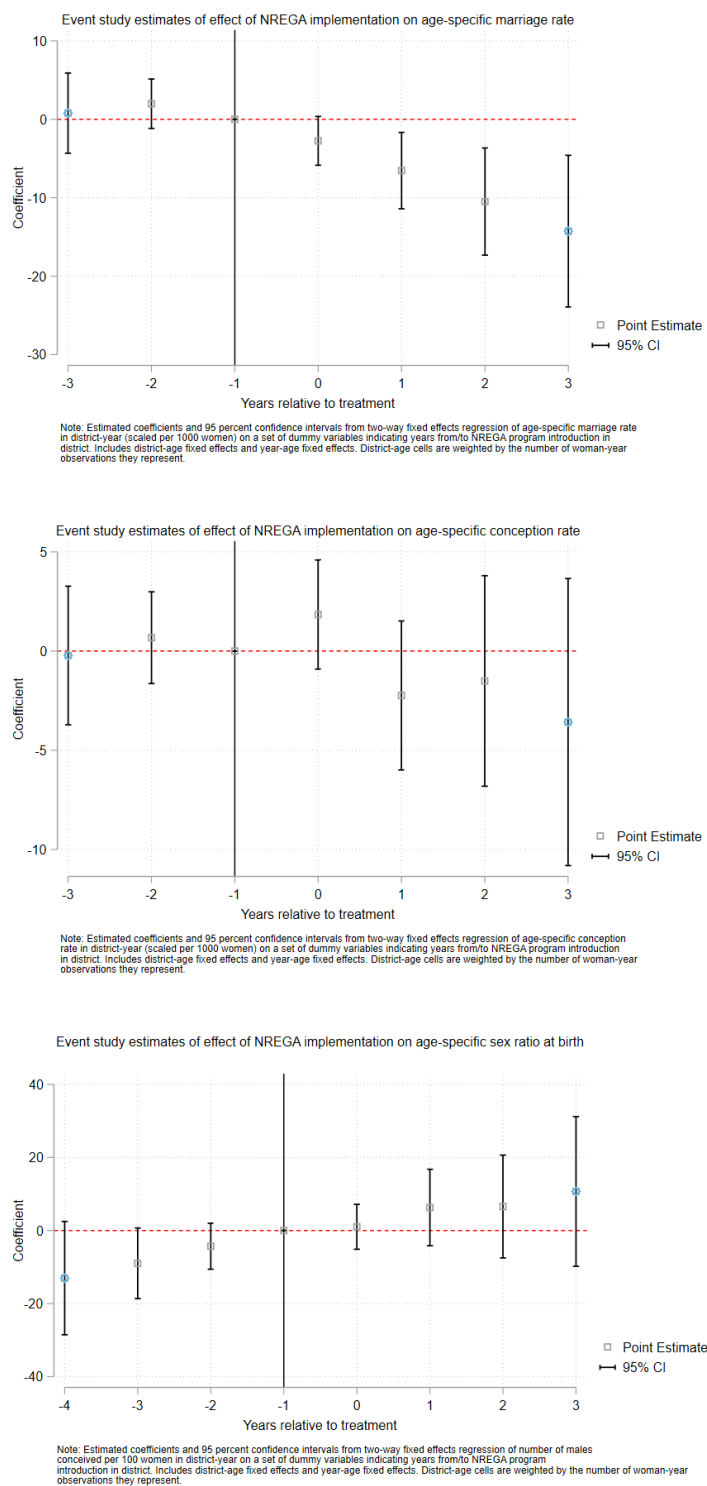


Figure 3.4: Event Study: Age-Adjusted Estimates

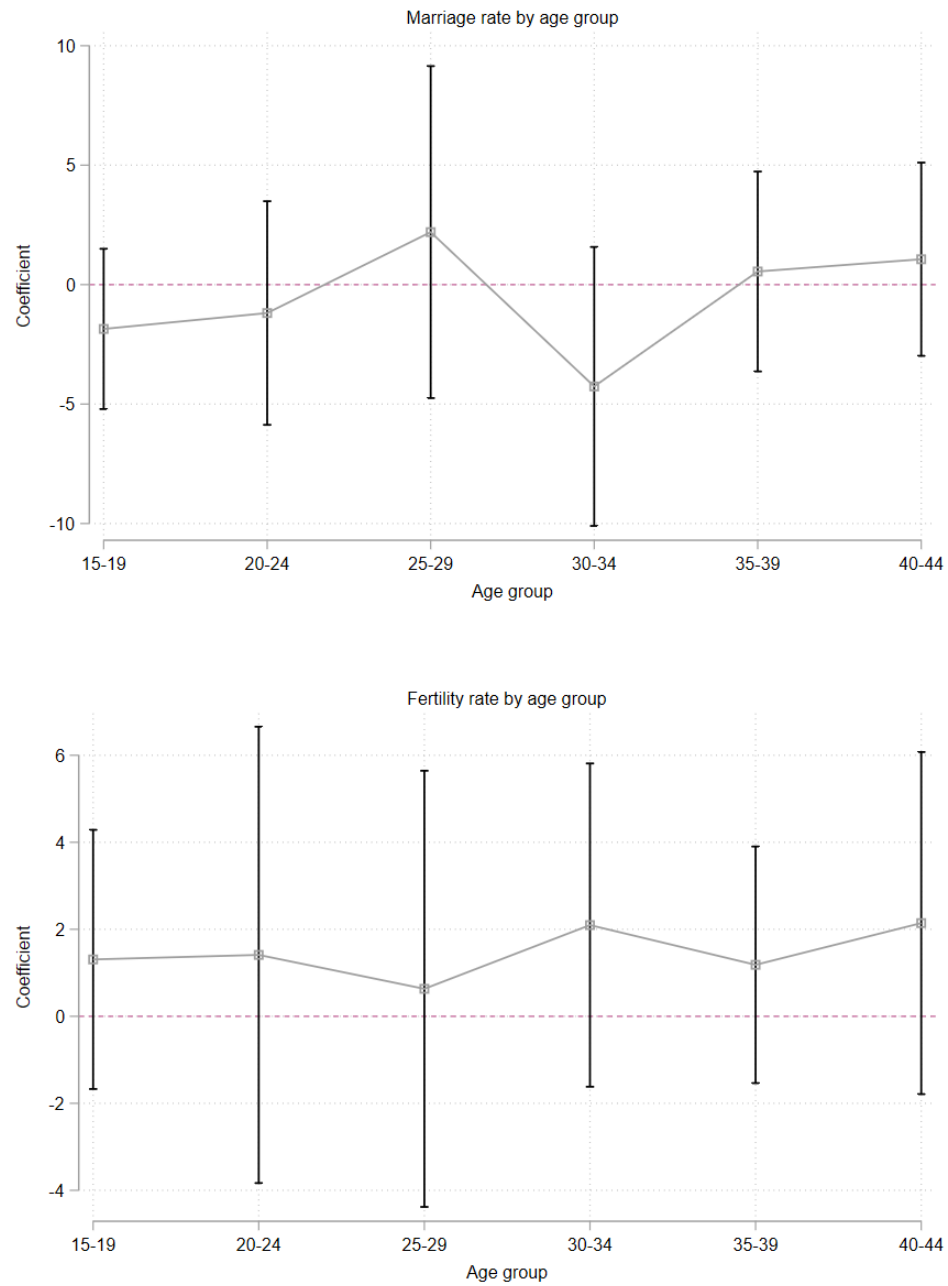


Figure 3.5: Age-Adjusted Two-Way Fixed Effects Estimates by Age Group

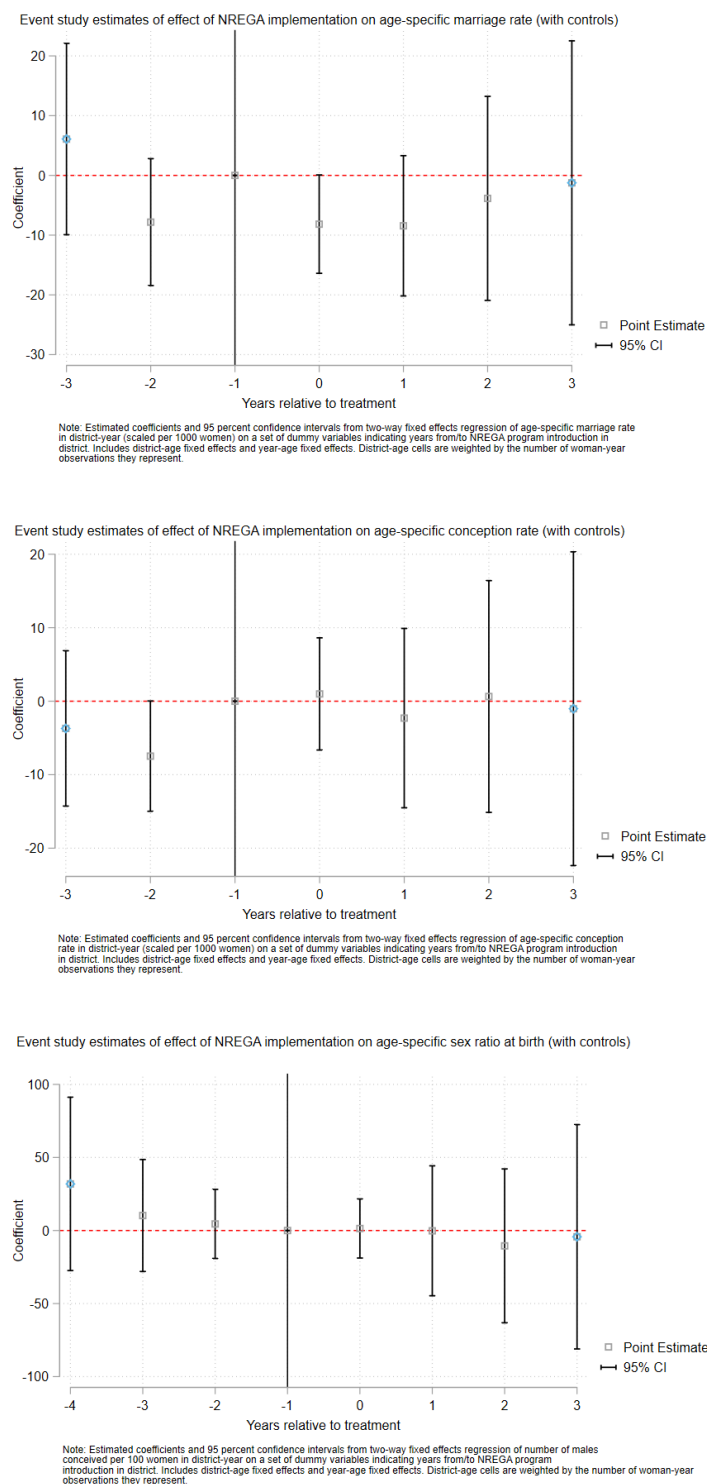


Figure 3.6: Event Study: Age-Adjusted Estimates with Controls

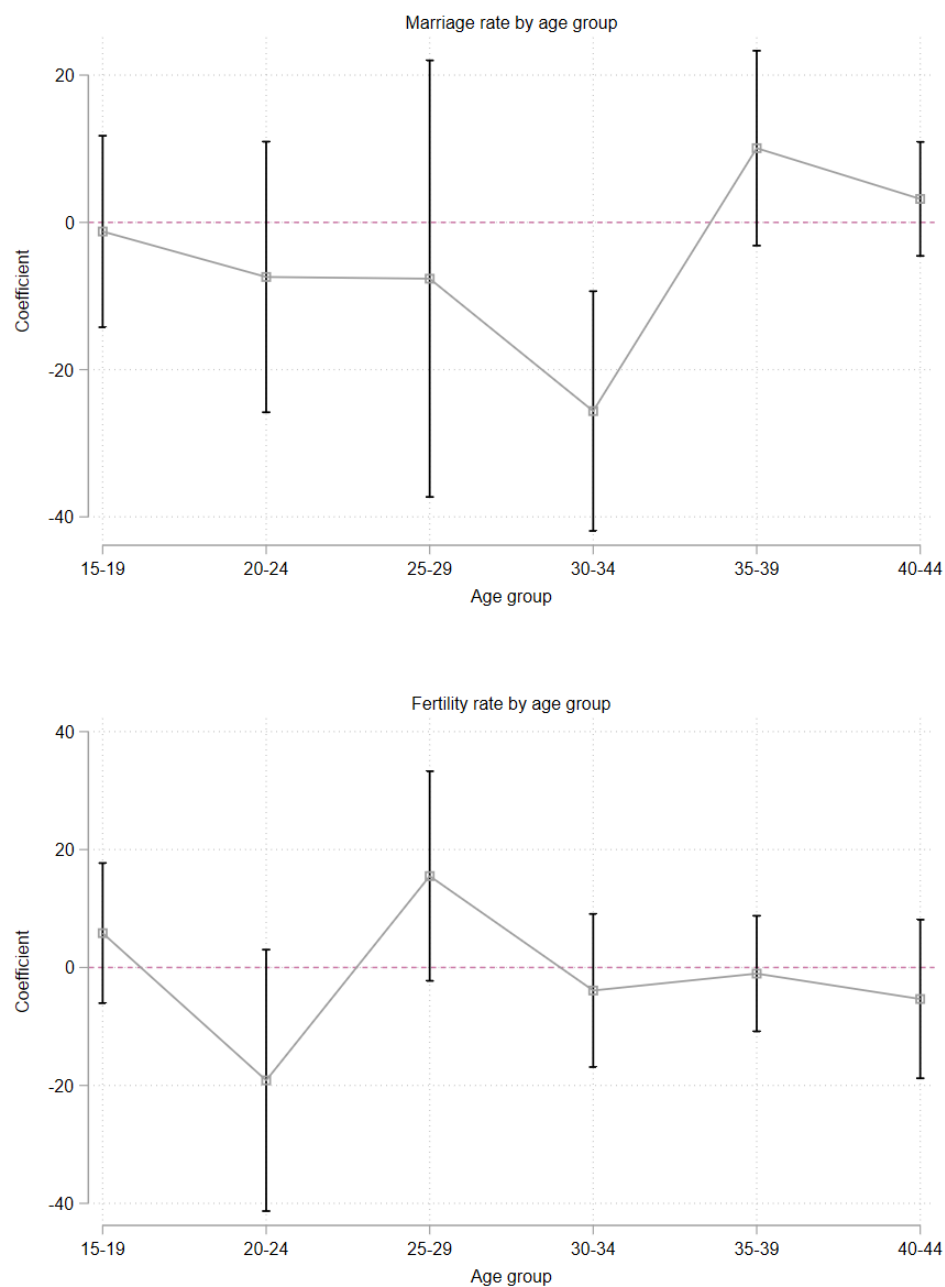


Figure 3.7: Age-Adjusted Two-Way Fixed Effects Estimates with Controls by Age Group

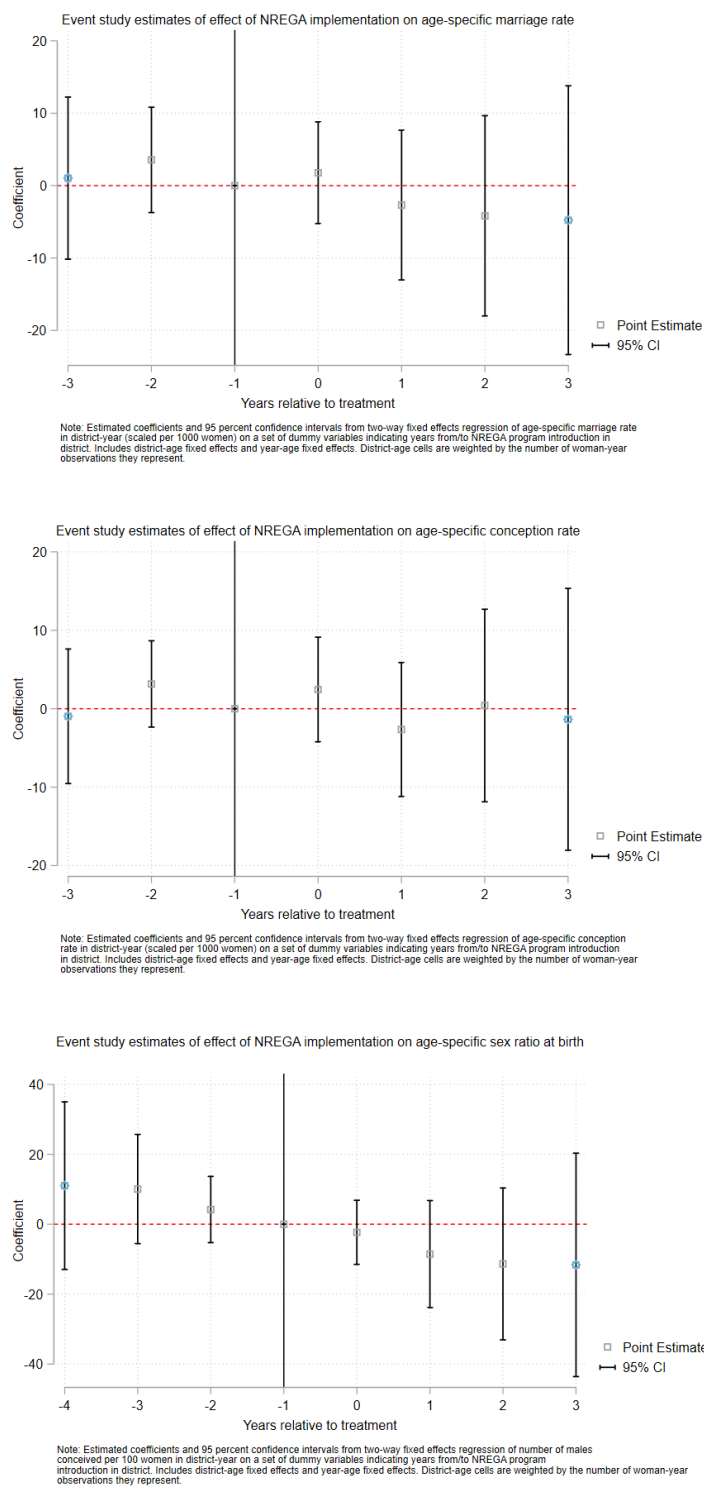


Figure 3.8: Event Study: Age-Adjusted Estimates for the Poorest Quartile

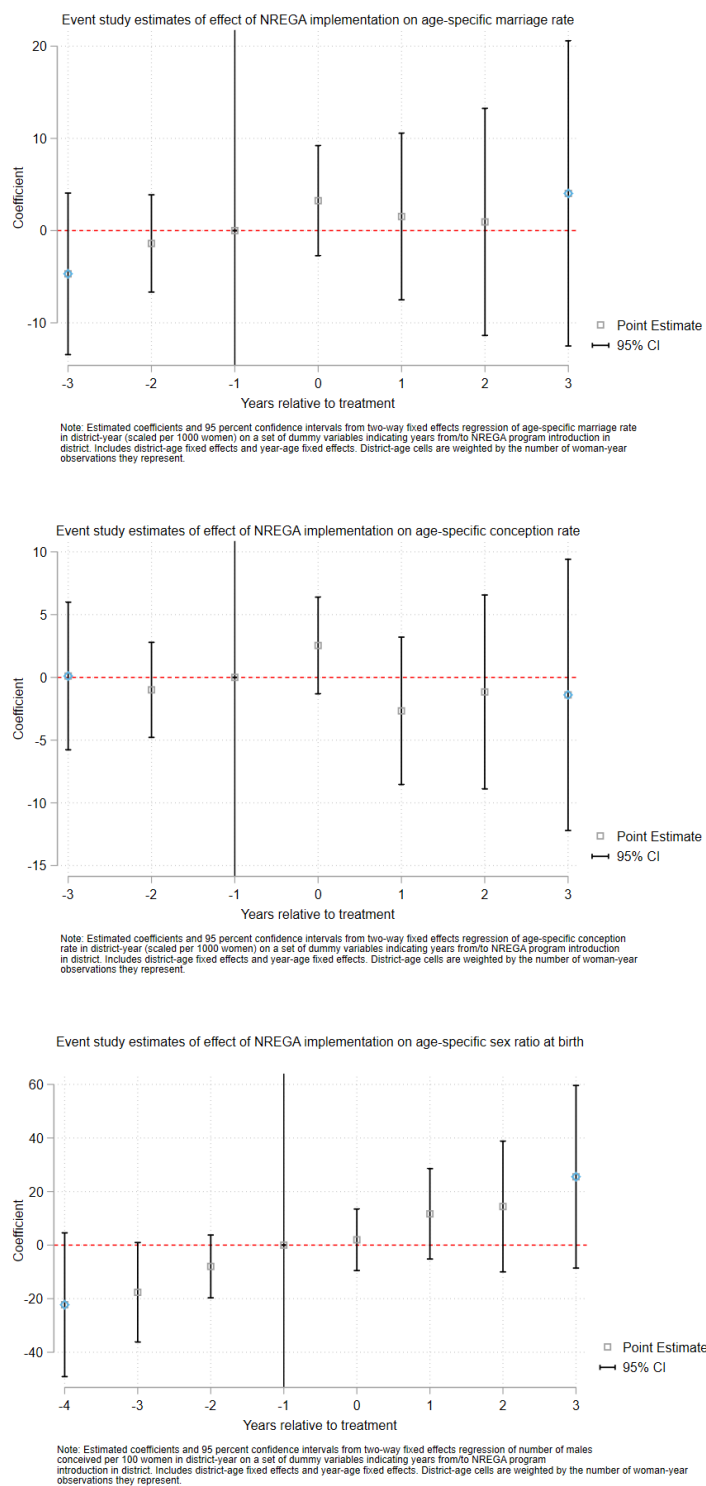


Figure 3.9: Event Study: Age-Adjusted Estimates for the Richest Quartile

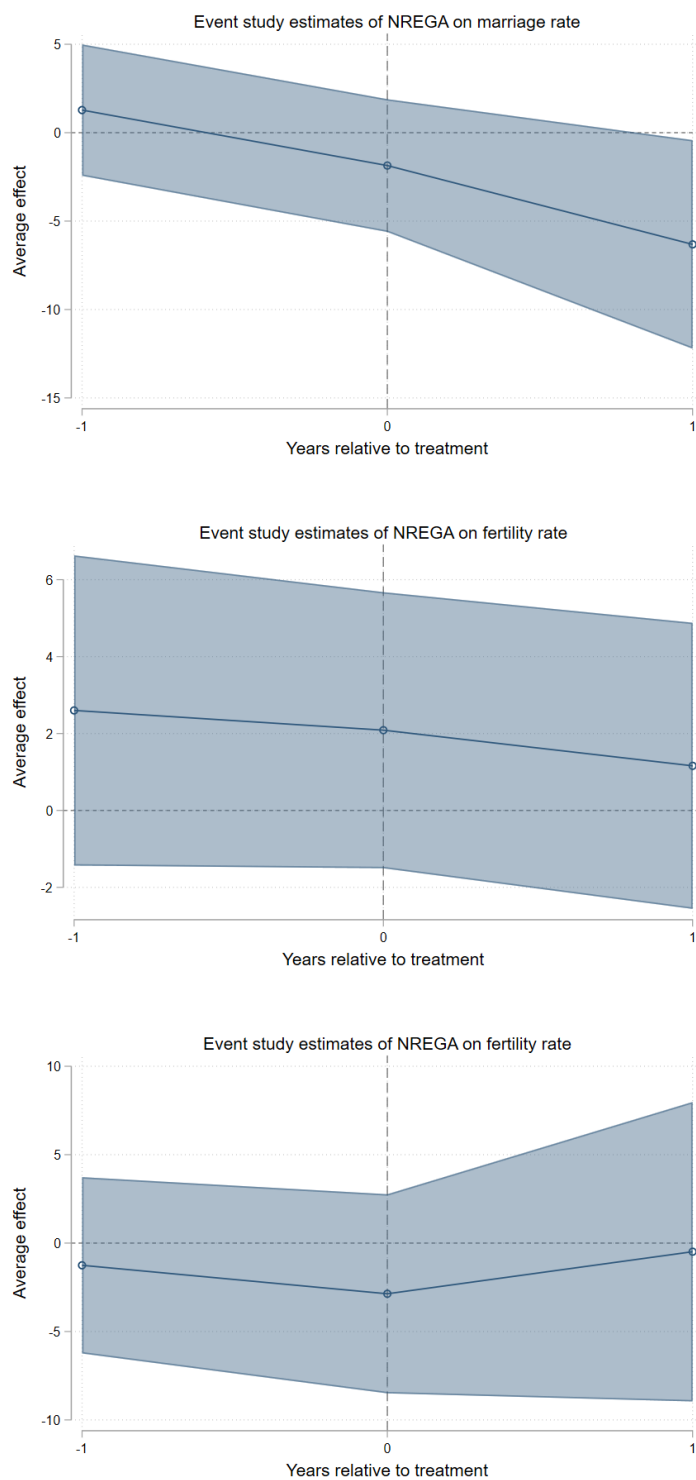


Figure 3.10: Event Study: de Chaisemartin-D'Haultfœuille Estimates

Table 3.1: Effect of NREGA Implementation on Marriage and Fertility

	(1) Marriage rate	(2) Conception rate	(3) Sex ratio
NREGA	0.215 (1.161)	1.636 (1.034)	1.400 (1.974)
Women-Years	1,241,020	2,770,312	2,770,146
Women	216,744	408,414	408,414
Marriages/Births/Sons	97,962	330,073	171,725

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district and year fixed effects. Dependent variables in columns (1) and (2) are district-year means based on indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.2: Effect of NREGA Implementation on Marriage and Fertility:
Age-Adjusted Estimates

	(1) Age-specific marriage rate	(2) Age-specific conception rate	(3) Age-specific sex ratio
NREGA	-1.279 (1.187)	1.355 (1.017)	-0.991 (2.243)
Women-Years	1,240,580	2,770,271	1,697,541
Women	216,744	408,414	408,414
Marriages/Births/Sons	97,962	330,073	171,725

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district-age and year-age fixed effects. Dependent variables in columns (1) and (2) are indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.3: Effect of NREGA Implementation on Marriage and Fertility with Controls

	(1) Age-specific marriage rate	(2) Age-specific conception rate	(3) Age-specific sex ratio
NREGA	-7.557** (3.520)	1.641 (3.285)	7.846 (8.001)
Women-Years	853,113	1,987,692	1,212,239
Women	151,858	294,056	294,056
Marriages/Births/Sons	70,836	240,382	125,304

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district-age fixed effects, year-age fixed effects, and pre-program district controls interacted with time fixed effects. Dependent variables in columns (1) and (2) are indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.4: Effect of NREGA Implementation on Marriage by Age Group with Controls

Age Group:	(1) 15-19	(2) 20-24	(3) 25-29	(4) 30-34	(5) 35-39	(6) 40-44
NREGA	-1.237 (6.635)	-7.417 (9.377)	-7.649 (15.125)	-25.629*** (8.304)	10.080 (6.752)	3.184 (3.948)
Women-Years	445,494	243,554	80,052	37,649	31,952	13,802
Women	128,718	76,614	28,726	12,825	9,003	4,887
Marriages	41,797	23,955	4,217	686	160	20

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation by binned age group, with district-age and year-age fixed effects and district controls interacted with time fixed effects. The dependent variable is an indicator for marriage occurring in a given district-year, multiplied by 1,000. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.5: Effect of NREGA Implementation on Fertility by Age Group with Controls

Age Group:	(1) 15-19	(2) 20-24	(3) 25-29	(4) 30-34	(5) 35-39	(6) 40-44
NREGA	5.827 (6.060)	-19.142* (11.309)	15.509* (9.068)	-3.902 (6.624)	-1.034 (4.999)	-5.340 (6.862)
Women-Years	488,797	420,534	364,482	327,416	272,306	111,764
Women	146,713	129,670	112,746	101,794	77,051	38,719
Births	48,476	96,679	57,717	25,536	10,401	1,559

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation by binned age group, with district-age and year-age fixed effects and district controls interacted with time fixed effects. The dependent variable is an indicator for conception occurring in a given district-year, multiplied by 1,000. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.6: Effect of NREGA Implementation on Marriage and Fertility for the Poorest Quartile

	(1) Age-specific marriage rate	(2) Age-specific conception rate	(3) Age-specific sex ratio
NREGA	0.940 (2.804)	0.238 (2.292)	-0.883 (3.239)
Women-Years	286,340	703,016	377,731
Women	50,541	102,730	102,730
Marriages/Births/Sons	25,731	116,860	59,991

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district-age and year-age fixed effects for the bottom 25 percent of households in the wealth distribution. Dependent variables in columns (1) and (2) are indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.7: Effect of NREGA Implementation on Marriage and Fertility for the Richest Quartile

	(1) Age-specific marriage rate	(2) Age-specific conception rate	(3) Age-specific sex ratio
NREGA	1.693 (2.446)	1.877 (1.595)	-2.783 (4.570)
Women-Years	341,754	700,950	184,391
Women	58,582	103,280	103,280
Marriages/Births/Sons	24,147	57,037	30,426

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district-age and year-age fixed effects for the top 25 percent of households in the wealth distribution. Dependent variables in columns (1) and (2) are indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3.8: Effect of NREGA Implementation on Marriage and Fertility: de Chaisemartin-D'Haultfoeuille Estimates

	(1) Marriage rate	(2) Conception rate	(3) Sex ratio
NREGA	-3.620* (2.070)	1.729 (1.627)	-1.947 (3.032)
Women-years	1,241,020	2,770,312	2,770,312
Women	216,744	408,414	408,414

Notes: The table presents estimates from regressions of the dependent variable on an indicator for NREGA implementation, with district and year fixed effects, using the de Chaisemartin-D'Haultfoeuille estimator. Dependent variables in columns (1) and (2) are district-year means based on indicators for marriage or conception occurring in a given district-year, multiplied by 1,000. The dependent variable in column (3) is the sex ratio at birth, defined as the number of males conceived per 100 females. Standard errors are clustered at the district level. Significance levels: * 0.10, ** 0.05, *** 0.01.

Appendix A

Appendix for Chapter 1

A.1 Data and Variable Construction

A.1.1 District Panel Construction

Few studies in India undertake a long time series analysis at the district level due to changes in district boundaries over time. Some studies use only districts with unchanged boundaries, which limits the sample, whereas other studies assume that district splits are subsets of the original parent district. However, in reality, new district boundaries can cut across multiple parent districts. This complicates the construction of a longer district-level panel. One contribution of this study is the construction of a consistent district panel over the period 1961–2011. I use the 640 districts in the 2011 Census, which also gives me a larger district-level sample than studies that only use unchanged districts.

If district boundary changes are not carefully accounted for, population data may contain large measurement error. To address this issue, I use harmonized population counts across districts drawn from the “Decadal Variation in Population” tables provided by the Census of India, which account for district boundary changes over time. I then interpolate these population estimates across intercensal years.

For outcome data, I use the 2019 Demographic and Health Survey (DHS). I map the shapefile for the 2019 DHS cluster geocodes to 2011 Census district administrative boundaries (available via SHRUG) using QGIS. My treatment variable comes from the UDISE school database. I scraped geocodes for all 1.5 million schools and mapped them to DHS clusters using QGIS. I then map DHS households to DHS clusters, which are now linked to Census district boundaries.

A.1.2 Construction of the Assortative Mating Index

This appendix describes the construction of the educational assortative mating index used in the main analysis. The measure follows Eika, Mogstad and Zafar (2019) and captures the extent to which individuals with a given education level marry each other more frequently than would occur under random matching, conditional on the marginal distributions of education.

The analysis uses data on currently married couples from the DHS Individual Recode files. For each couple, I observe the wife’s and husband’s educational attainment, district of residence, year of cohabitation, and DHS sampling weights. Educational attainment is grouped into five categories: no schooling, primary (1–5 years), upper primary (6–8 years), secondary (9–12 years), and college education (more than 12 years). College-educated individuals are excluded from some disaggregated trend figures because small cell sizes generate volatile series in earlier marriage cohorts. Marriage cohorts are grouped into five-year bins.

For each district d and marriage cohort t , I first compute the weighted joint distribution of spouses’ educational attainment:

$$P_{dt}(e_f, e_m) = \frac{\sum_{i \in (d,t)} w_i \cdot 1\{E_i^f = e_f, E_i^m = e_m\}}{\sum_{i \in (d,t)} w_i},$$

where E_i^f and E_i^m denote the education categories of the wife and husband in couple i , and w_i denotes the DHS sampling weight. From the joint distribution, I recover the marginal distributions of women's and men's education within each district-cohort cell:

$$P_{dt}^f(e_f) = \sum_{e_m} P_{dt}(e_f, e_m), \quad P_{dt}^m(e_m) = \sum_{e_f} P_{dt}(e_f, e_m).$$

The assortative mating index for education pair (e_f, e_m) is then defined as:

$$s_{dt}(e_f, e_m) = \frac{P_{dt}(e_f, e_m)}{P_{dt}^f(e_f) \cdot P_{dt}^m(e_m)}.$$

Under random matching with respect to education, the joint probability equals the product of the marginals and the index equals one. Values greater than one indicate positive assortative matching, meaning that individuals with those education levels marry each other more frequently than would occur under random matching.

The main analysis focuses on same-education matches, corresponding to the diagonal elements $s_{dt}(e, e)$. I additionally construct an aggregate assortative mating index for each district and marriage cohort as a weighted average of the diagonal elements, where weights are proportional to the observed share of homogamous matches in each education category:

$$S_{dt}^{\text{agg}} = \sum_e \omega_{dt,e} s_{dt}(e, e), \quad \omega_{dt,e} = \frac{P_{dt}(e, e)}{\sum_k P_{dt}(k, k)}.$$

This aggregate measure summarizes the overall intensity of educational assortative matching within district d and cohort t , net of changes in the marginal distributions of education across cohorts.

A.1.3 Construction of Domestic Violence Outcomes

Outcome	Construction
Decision-making	Indicator equal to the average of four binary measures capturing whether the respondent has any say in major household decisions. For each domain, an indicator is coded as 1 if the decision is made by the respondent alone or jointly with her partner, and 0 otherwise: (1) respondent's healthcare; (2) large household purchases; (3) visits to relatives or family; (4) decisions over husband's earnings.
Controlling partner	Binary indicator equal to 1 if the respondent reports at least one controlling behavior by her husband, and 0 otherwise: (1) husband/partner is jealous if respondent talks with other men; (2) accuses respondent of unfaithfulness; (3) does not permit respondent to meet female friends; (4) tries to limit respondent's contact with family; (5) insists on knowing where respondent is at all times; (6) does not trust respondent with money.
Emotional abuse	Binary indicator equal to 1 if the respondent has ever experienced emotional violence by her husband, and 0 otherwise: (1) ever been humiliated; (2) threatened with harm; (3) insulted or made to feel bad.
Physical violence	Binary indicator equal to 1 if the respondent has ever experienced physical violence by her husband, and 0 otherwise: (1) moderate physical violence (pushed, shook, had something thrown, slapped, punched, had arm twisted, or hair pulled); (2) severe physical violence (kicked or dragged, strangled or burnt, threatened with a weapon).
Sexual violence	Binary indicator equal to 1 if the respondent has ever experienced sexual violence by her husband, and 0 otherwise: (1) ever physically forced into unwanted sex; (2) forced into other unwanted sexual acts; (3) physically forced to perform sexual acts she did not want to perform.

Notes: All domestic violence variables are measured for women selected into the DHS domestic violence module. The decision-making outcome is coded so that higher values indicate greater participation in household decision-making; the domestic violence outcomes are coded so that higher values indicate worse relationship quality.

A.2 For Online Publication

A.2.1 Additional Figures

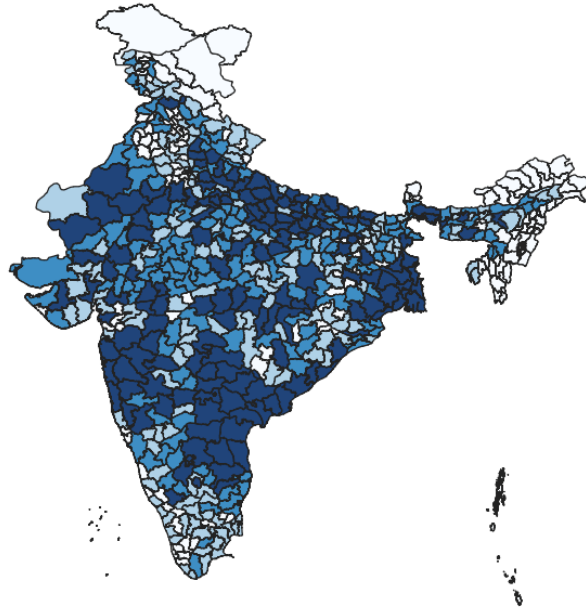


Figure A.1: Spatial Distribution of Schools across Districts

Notes: The figure maps the number of schools observed in the UDISE 2018 school census by district. Districts are grouped into quartiles of the distribution, with darker shading indicating a larger number of schools. Because UDISE records schools operating at the time of the 2018 survey, the map reflects the spatial distribution of observed schools rather than historical cohort-specific school exposure.

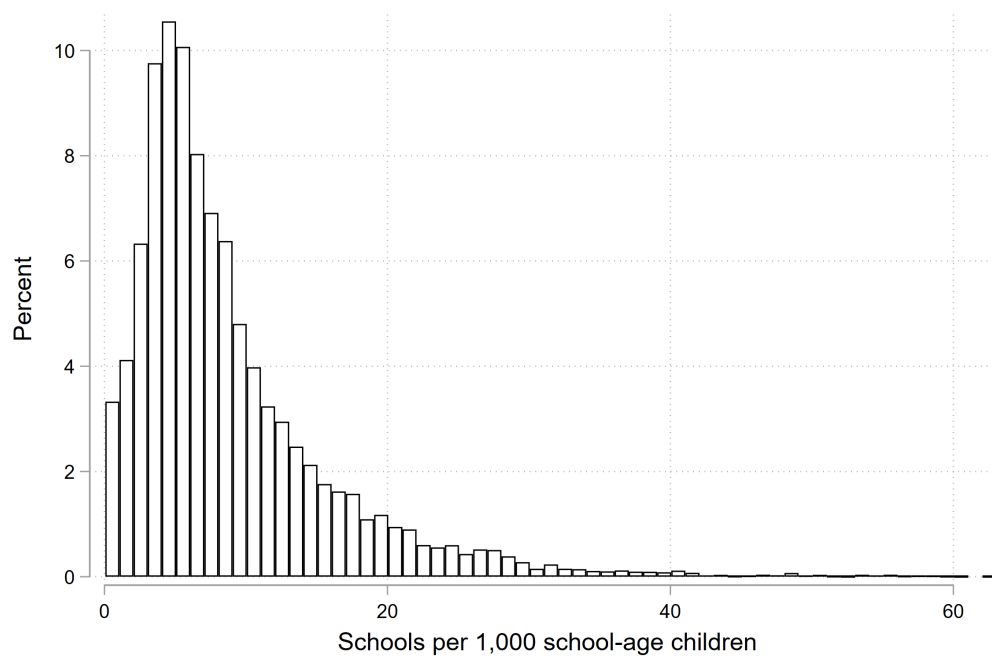


Figure A.2: Distribution of School Construction Exposure

Notes: The figure plots the distribution of cumulative school construction exposure during ages 6–15, measured as schools per 1,000 school-age children in the respondent's district.

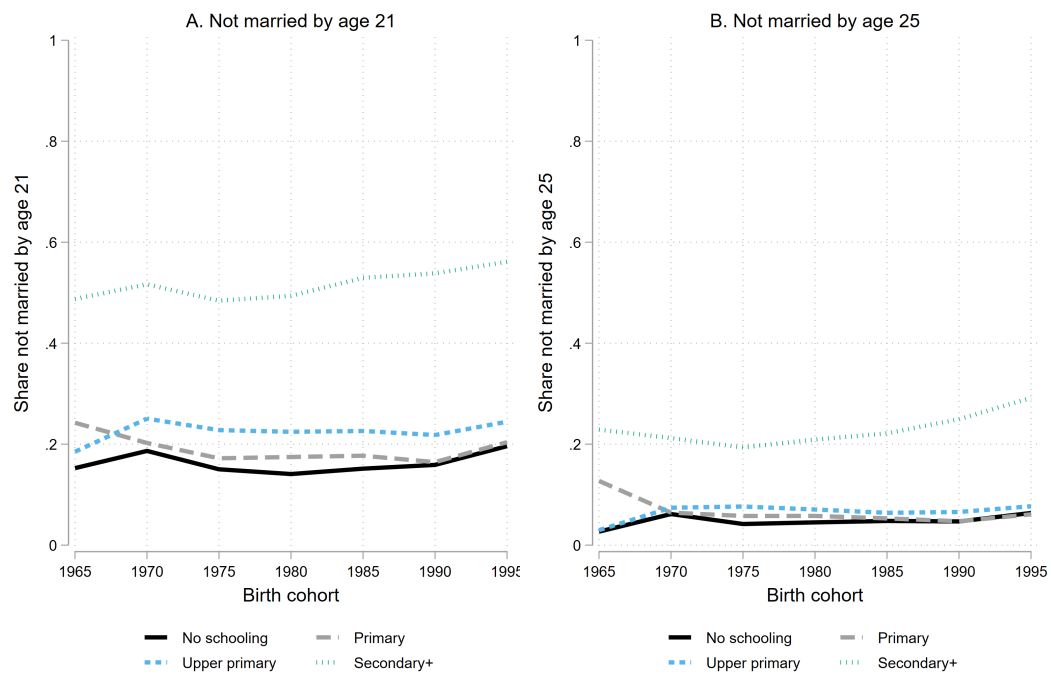


Figure A.3: Marriage Timing by Education Level

Notes: The figure plots the share of women not yet married by ages 21 and 25 across five-year birth cohorts separately by educational attainment category. Educational categories are defined as no schooling, primary (1–5 years), upper primary (6–8 years), and secondary or higher (9+ years). Women younger than the relevant age threshold at the time of interview are excluded from the corresponding calculation.

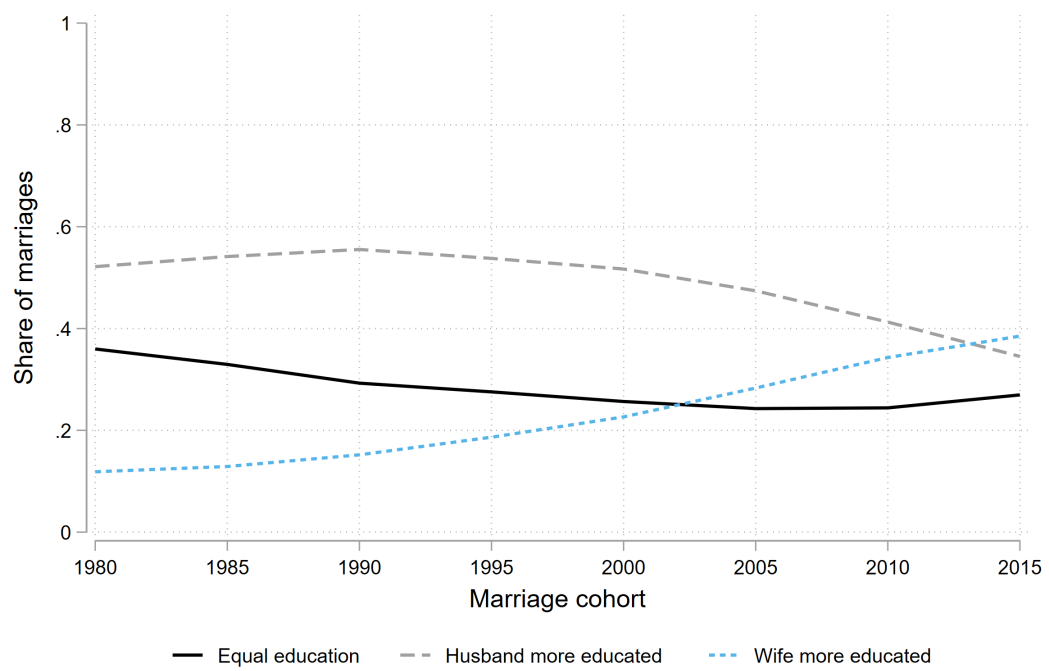


Figure A.4: Direction of Educational Matching by Marriage Cohort

Notes: The figure plots the share of marriages in which spouses have equal education, the husband has more education, or the wife has more education. Marriage cohorts are grouped into five-year bins.

A.2.2 Additional Tables

Table A.1: Education and Work in Matrimonial Advertisements

	Men (seeking brides)	Women (seeking grooms)
Number of profiles	8,355	6,430
<i>Attributes signaled (%)</i>		
Education signaled	76.10	75.85
Work signaled	46.99	21.12
<i>Attributes demanded (%)</i>		
Education demanded	53.56	47.57
Work demanded	18.77	47.57
Education or work demanded	59.34	56.25

Notes: This table summarizes characteristics of matrimonial advertisements from an Indian newspaper. “Signaled” refers to attributes mentioned in the profile description, while “demanded” refers to attributes specified for the prospective bride or groom.

Table A.2: Effect of School Construction on Marriage Timing

	(1)	(2)	(3)
	Not married by 18	Not married by 21	Not married by 25
<i>Panel A: Continuous treatment</i>			
Schools built	0.003** (0.002)	0.000 (0.002)	-0.002 (0.001)
Mean of treatment	8.904	8.904	8.722
Elasticity of outcome	0.046	0.004	-0.099
<i>Panel B: Binary treatment</i>			
Schools built > mean	0.024** (0.010)	0.011 (0.008)	-0.001 (0.005)
Mean of treatment	0.342	0.342	0.340
Elasticity of outcome	0.013	0.011	-0.002
Mean of outcome	0.642	0.359	0.147
N	462,963	462,963	427,195

Notes: Each column reports estimates from regressions of the indicated marriage-timing outcome on school construction exposure in the woman's district during her school-age years, with district and birth-cohort fixed effects. Panel A uses cumulative schools constructed per 1,000 population aged 6–15 years. Panel B uses an indicator for whether exposure is above the cohort-specific mean. Elasticities are computed as $\beta \times \bar{X}/\bar{Y}$. Outcomes are indicators for whether a woman remained unmarried by the specified age. Standard errors are clustered at the district level.

Table A.3: Effect of School Construction on Direction of Educational Matching in Marriage

	(1) Hypergamy	(2) Hypogamy
<i>Panel A: Continuous treatment</i>		
Schools built	0.003 (0.002)	-0.000 (0.001)
Mean of treatment	8.759	8.759
Elasticity of outcome	0.046	-0.016
<i>Panel B: Binary treatment</i>		
Schools built > mean	0.003 (0.009)	-0.009 (0.006)
Mean of treatment	0.345	0.345
Elasticity of outcome	0.002	-0.013
Mean of outcome	0.504	0.231
N	360,548	360,548

Notes: Each column reports estimates from separate regressions where the dependent variable is an indicator for the type of educational match within marriage. In column (1), the outcome equals one if the husband has more years of schooling than the wife, and zero otherwise. In column (2), the outcome equals one if the wife has more years of schooling than the husband, and zero otherwise. Marriages in which spouses have equal education are included in the omitted category in both specifications. The key explanatory variable is district-level cumulative school construction exposure, measured as the number of schools per 1,000 school-age children in a woman's district during her school-age years. Panel A uses the continuous measure of exposure, while Panel B uses an indicator equal to one if exposure is above the cohort-specific mean. All regressions include district and birth-cohort fixed effects and are estimated using DHS sample weights. Standard errors are clustered at the district level. The sample is restricted to ever-married women with non-missing information on both spouses' education. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment variable to the mean of the outcome.

Appendix B

Appendix for Chapter 2

B.1 Data and Variable Construction

B.1.1 District Panel Construction

Few studies in India undertake a long time series analysis at the district level due to changes in district boundaries over time. Some studies use only districts with unchanged boundaries, which limits the sample, whereas other studies assume that district splits are subsets of the original parent district. However, in reality, new district boundaries can cut across multiple parent districts. This complicates the construction of a longer district-level panel. One contribution of this study is the construction of a consistent district panel over the period 1961–2011. I use the 640 districts in the 2011 Census, which also gives me a larger district-level sample than studies that only use unchanged districts.

If district boundary changes are not carefully accounted for, population data may contain large measurement error. To address this issue, I use harmonized population counts across districts drawn from the “Decadal Variation in Population” tables provided by the Census of India, which account for district boundary changes over time. I then interpolate these population estimates across intercensal years.

For outcome data, I use the 2019 Demographic and Health Survey (DHS). I map the shapefile for the 2019 DHS cluster geocodes to 2011 Census district administrative boundaries (available via SHRUG) using QGIS. My treatment variable comes from the UDISE school database. I scraped geocodes for all 1.5 million schools and mapped them to DHS clusters using QGIS. I then map DHS households to DHS clusters, which are now linked to Census district boundaries.

B.2 Additional Tables

Table B.1: Pre-trend Tests for Fertility and Intergenerational Outcomes

	(1) Age at first birth	(2) Completed fertility	(3) Child years of schooling
1970 cohort $\times$ eventual exposure	-0.125 (0.276)	-0.019 (0.112)	0.154 (0.265)
1975 cohort $\times$ eventual exposure	-0.145 (0.285)	-0.050 (0.114)	0.143 (0.269)
Mean of outcome	21.261	3.156	4.716
N	77,598	80,409	25,131
Joint F-statistic	0.142	0.861	0.211
Joint p-value	0.867	0.423	0.810

Notes: The table reports cohort-specific pre-trend tests using early maternal cohorts only. Eventual school exposure is defined as district-level average cumulative school construction exposure among later cohorts and standardized to have mean zero and standard deviation one. All regressions include district and cohort-bin fixed effects. The child years of schooling specification additionally includes child age fixed effects and a child gender indicator. Standard errors are clustered at the district level.

Table B.2: Effect of School Construction on Completed Fertility: Robustness to Age Restrictions

	(1) Age 35+	(2) Age 40+	(3) Age 45+
Schools built	-0.047*** (0.009)	-0.037** (0.017)	-0.037 (0.047)
Mean of treatment	5.087	4.627	4.333
Elasticity of outcome	-0.083	-0.058	-0.052
Mean of outcome	2.861	2.993	3.099
N	233,302	146,388	74,421

Notes: The table reports estimates of the effect of school construction exposure on completed fertility under alternative age restrictions. The treatment variable is the cumulative number of schools built per 1,000 school-age children in a woman's district of residence during ages 6–15. Each column restricts the sample to women above the indicated age threshold. All specifications include district and birth cohort fixed effects. Standard errors are clustered at the district level. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment to the mean of the outcome.

Table B.3: Intergenerational Effect of School Construction on Child Education:
Robustness to Age Window (6–18 Years)

	(1) Years of schooling	(2) Schooling gap	(3) Behind grade level
Schools built	0.046*** (0.008)	0.044*** (0.007)	-0.010*** (0.002)
Mean of treatment	8.638	8.638	8.638
Elasticity of outcome	0.084	-0.733	-0.227
Mean of outcome	4.744	-0.514	0.392
N	480,400	480,400	480,400
Child age FE	Yes	No	No
Child gender FE	Yes	Yes	Yes

Notes: The table reports robustness estimates of the intergenerational effect of school construction exposure on children’s education outcomes using a broader sample of children aged 6–18. The treatment variable is the cumulative number of schools built per 1,000 school-age children in the mother’s district of residence during ages 6–15. Outcome variables include years of schooling, schooling gap, and an indicator for being behind the age-appropriate grade level. Child age fixed effects are included only for the years of schooling regression, while a gender indicator is included in all specifications. For schooling gap and behind-grade indicators, age is accounted for in the construction of the dependent variable and is therefore not included as a control. All specifications include district and maternal birth cohort fixed effects. Standard errors are clustered at the district level. Elasticities are calculated as the coefficient multiplied by the ratio of the mean of the treatment to the mean of the outcome.

Bibliography

- Abanokova, Kristina, Francisco Campos, and Pushkar Maitra.** 2020. “The Long-Run Effects of Education Expansion on Fertility and Family Formation.” *World Development*, 127: 104772.
- Afridi, Farzana, Abhiroop Mukhopadhyay, and Soham Sahoo.** 2016. “Female labor force participation and child education in India: evidence from the National Rural Employment Guarantee Scheme.” *IZA Journal of Labor and Development*, 5(1): 1–53.
- Agarwal, Madhuri, Vikram Bahure, and Sayli Javadekar.** 2023. “Marrying Young: The Surprising Effect of Education.”
- Almar, Frederik, Benjamin Friedrich, Ana Reynoso, Bastian Schulz, and Rune M. Vejlin.** 2025. “Educational Ambition, Marital Sorting, and Inequality.” *NBER Working Paper*, , (33683).
- Anderberg, Dan, Jesper Bagger, V. Bhaskar, and Tanya Wilson.** 2022. “Marriage Market Equilibrium with Matching on Latent Ability: Identification Using a Compulsory Schooling Expansion.” Working Paper.
- Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena.** 2020. “Bride Price and Female Education.” *Journal of Political Economy*, 128(2): 591–641.
- Azam, Mehtabul.** 2012. “The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment.”
- Azam, Mehtabul.** 2015. “Like Father, Like Son? Intergenerational Educational Mobility in India.” *Demography*, 52(6): 1929–1959.
- Azam, Mehtabul, and Chan Hang Saing.** 2016. “Assessing the Impact of the District Primary Education Program in India.” *World Development*, 81: 1–17.
- Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune.** 2013. “Marry for What? Caste and Mate Selection in Modern India.” *American Economic Journal: Microeconomics*, 5(2): 33–72.
- Becker, Gary S.** 1960. “An Economic Analysis of Fertility.” *Demographic and Economic Change in Developed Countries*, 11: 209–240.

- Becker, Gary S.** 1981. *A Treatise on the Family*. Cambridge, MA:Harvard University Press. NBER Books, National Bureau of Economic Research, Inc.
- Becker, Gary S., and H. Gregg Lewis.** 1973. "On the Interaction between the Quantity and Quality of Children." In *Economics of the Family: Marriage, Children, and Human Capital*, ed. Theodore W. Schultz, 81–90. Chicago:University of Chicago Press.
- Becker, Gary S., and Nigel Tomes.** 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, 84(4, Part 2): S143–S162.
- Becker, Gary S., and Nigel Tomes.** 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87(6): 1153–1189.
- Behrman, Jere R., and Mark R. Rosenzweig.** 2002. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review*, 92(1): 323–334.
- Behrman, Jere R., Andrew D. Foster, Mark R. Rosenzweig, and Prem Vashishtha.** 1999. "Women's Schooling, Home Teaching, and Economic Growth." *Journal of Political Economy*, 107(4): 682–714.
- Black, Sandra E., and Paul J. Devereux.** 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, 4: 1487–1541.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes.** 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review*, 95(1): 437–449.
- Breierova, Lucia, and Esther Duflo.** 2004. "The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers?" *NBER Working Paper*, , (10513).
- Carranza, Eliana.** 2014. "Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India." *American Economic Journal: Applied Economics*, 6(4): 197–225.
- Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir.** 2018. "The Marriage Market, Labor Supply, and Education Choice." *Journal of Political Economy*, 126(S1): S26–S72.
- Choo, Eugene, and Aloysius Siow.** 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.
- Cook, C. Justin, and Manisha Shah.** 2022. "Aggregate Effects from Public Works: Evidence from India." *Review of Economics and Statistics*, 104(4): 797–806.

- Currie, Janet, and Enrico Moretti.** 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Quarterly Journal of Economics*, 118(4): 1495–1532.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille.** 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964–2996.
- Desai, Sonalde, and Lester Andrist.** 2010. "Gender Scripts and Age at Marriage in India." *Demography*, 47(3): 667–687.
- Doepke, Matthias.** 2015. "Gary Becker on the Quantity and Quality of Children." *Journal of Demographic Economics*, 81(1): 59–66.
- Doepke, Matthias, Moshe Hazan, and Yishay D. Maoz.** 2013. "The baby boom and world war II: A macroeconomic analysis." *Review of Economic Studies*, 82(3): 1031–1073.
- Doorley, Karina, Arnaud Dupuy, and Simon Weber.** 2019. "The Empirical Content of Marital Surplus in Matching Models." *Economics Letters*, 176: 51–54.
- Duflo, Esther.** 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.
- Eika, Lasse, Magne Mogstad, and Basit Zafar.** 2019. "Educational Assortative Mating and Household Income Inequality." *Journal of Political Economy*, 127(6): 2795–2835.
- Fernández, Raquel, Nezih Guner, and John Knowles.** 2005. "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Quarterly Journal of Economics*, 120(1): 273–344.
- Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Moore.** 2021. "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms." *American Economic Review*, 111(7): 2342–75.
- Geruso, Michael, and Heather Royer.** 2014. "The Impact of Education on Fertility and Childbearing." *Journal of Labor Economics*, 32(4): 799–826.
- Goldin, Claudia.** 1994. "The U-Shaped Female Labor Force Function in Economic Development and Economic History."
- Greenwood, Jeremy, Ananth Seshadri, and Guillaume Vandenbroucke.** 2005. "The baby boom and baby bust." *American Economic Review*, 95(1): 183–207.

- Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos.** 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review Papers and Proceedings*, 104(5): 348–353.
- Heath, Rachel, and A. Mushfiq Mobarak.** 2015. "Manufacturing growth and the lives of Bangladeshi women." *Journal of Development Economics*, 115: 1–15.
- Imbert, Clément, and John Papp.** 2015. "Labor market effects of social programs: Evidence from India's employment guarantee." *American Economic Journal: Applied Economics*, 7(2): 233–263.
- Imbert, Clément, and John Papp.** 2020. "Short-term Migration, Rural Public Works, and Urban Labor Markets: Evidence from India." *Journal of the European Economic Association*, 18(2): 927–963.
- Jensen, Robert.** 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *The Quarterly Journal of Economics*, 127(2): 753–792.
- Khanna, Gaurav.** 2023. "Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India." *Journal of Political Economy*. Forthcoming.
- Lafortune, Jeanne.** 2013. "Making Yourself Attractive: Premarital Investments and the Returns to Education in the Marriage Market." *American Economic Journal: Applied Economics*, 5(2): 151–178.
- Lundberg, Shelly, and Robert A. Pollak.** 2012. "Personality and Marital Surplus." *IZA Journal of Labor Economics*, 1(3). Discusses how shared preferences and consumption complementarities contribute to positive assortative mating, beyond pure production complementarities.
- Mourifié, Ismael, and Aloysius Siow.** 2017. "The Cobb–Douglas Marriage Matching Function: Marriage Matching with Peer and Scale Effects." *Unpublished Manuscript*.
- Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar.** 2018. "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India." *NBER Working Paper*, 1–76.
- Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens.** 2006. "The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics*, 24(4): 729–760.
- Osili, Una Okonkwo, and Bridget Terry Long.** 2008. "Does Female Schooling Reduce Fertility? Evidence from Nigeria." *Journal of Development Economics*, 87(1): 57–75.

- Peters, H. Elizabeth.** 2002. “Competing Premarital Investments and the Return to Education in the Marriage Market.” *Journal of Human Resources*, 37(3): 679–697.
- Qian, Nancy.** 2008. “Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance.” *Quarterly Journal of Economics*, 123(3): 1251–1285.
- Schwartz, Christine R.** 2013. “Trends and Variation in Assortative Mating: Causes and Consequences.” *Annual Review of Sociology*, 39: 451–470.
- Shah, Manisha, and Bryce Millett Steinberg.** 2019. “Workfare and Human Capital Investment: Evidence from India.” *Journal of Human Resources*, 1117–9201R2.
- Sivasankaran, Anitha.** 2014. “Work and Women’s Marriage, Fertility and Empowerment: Evidence from Textile Mill Employment.”
- Sunder, Naveen.** 2019. “Parents’ Schooling and Intergenerational Human Capital: Evidence from India.” Job Market Paper, Cornell University.
- Vogl, Tom S.** 2013. “Marriage Institutions and Sibling Competition: Evidence from South Asia.” *Quarterly Journal of Economics*, 128(3): 1017–1072.
- Voigtländer, Nico, and Hans-Joachim Voth.** 2013. “How the West “Invented” Fertility Restriction.” *American Economic Review*, 103(6): 2227–2264.
- Zimmermann, Laura.** 2020. “Why Guarantee Employment? Evidence from a Large Indian Public-Works Program.”