Returns to Education in the Marriage Market: Access to Education, Age at Marriage and Assortative Matching in India

Akanksha Arora

March 2026

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 increases women's schooling by 0.12 years. Exposure to school construction also delays marriage by 0.04 years, reduces the spousal age gap by 0.03 years, and lowers the probability of child marriage by 0.3 percentage points. Marriage markets have become increasingly stratified at the bottom of the education distribution, while assortative mating has weakened at higher education levels. School construction contributes to these trends by strengthening educational sorting: each additional school raises the assortative mating index by 1.4 percent. Back-of-the-envelope calculations imply that school expansion accounts for approximately

28 percent of the observed rise in assortative mating. These findings suggest that education expansion amplified stratification in marriage markets instead of mitigating it.

1 Introduction

In recent decades, India has witnessed a dramatic rise in educational attainment, particularly among women (Figures 1). This narrowing gender gap in schooling has occurred alongside persistently low levels of female labor force participation, raising an important puzzle. Conventional economic theory suggests that individuals invest in human capital in response to labor market returns. Yet, in the Indian context, where women's employment opportunities remain constrained by restrictive social norms and structural barriers (Mukherjee (2017), Afridi et al. (2023), Kleven, Landais and Leite-Mariante (2025)), the observed increase in female education seems to exceed what labor market incentives alone would predict. What, then, accounts for this apparent "over-investment" in women's education?

This paper investigates whether part of the answer lies in the marriage market. Marriage remains a near-universal institution in India, shaping both individual welfare and household economic outcomes. If education influences not only employment opportunities but also marital prospects, it may provide substantial non-labor returns. In particular, education may affect the probability of marriage, the timing and patterns of marital matching (Behrman et al. (1999), Peters (2002), Desai and Andrist (2010), Lafortune (2013), Royer and Geruso (2014), Bhaskar et al. (2015), Ashraf et al. (2020), Anderberg et al. (2022), Agarwal, Bahure and Javadekar (2023)), as well as the distribution of bargaining power and welfare within households (Choo and Siow (2006), Mourifié and Siow (2017), Doorley, Dupuy and Weber (2019)). Understanding these dynamics is crucial for evaluating the broader returns to education in settings where women's labor force participation is limited.

The mechanisms through which education may generate marriage market returns are multifaceted. Education can facilitate assortative matching, where men and women with similar educational attainment are more likely to marry, potentially due to shared preferences (consumption complementarities) or joint gains in household production and child rearing (production complementarities). Rising average schooling levels can also mechanically increase the likelihood of matches between educated partners, even without changes in sorting behavior. Moreover, assortative mating carries significant distributional implications: while it may improve coordination within households, it can also concentrate advantages across generations, exacerbating inequality and limiting mobility.

This study leverages large-scale variation in school construction across Indian districts and cohorts to identify the causal impact of expanded educational access on women's marriage outcomes. Specifically, I examine how exogenous increases in schooling opportunities influenced educational attainment, assortative mating patterns, age at marriage, and intra-household dynamics. By combining administrative data on school expansion with data from the Demographic and Health Survey, I assess both individual- and household-level consequences of education in the marriage market.

2 Background on school construction in India

School access has expanded dramatically in India since Independence (Figure 2). The Minimum Needs Program in the mid-1970s aimed to provide a primary school within 1.5 km of every village. According to census data, the share of villages with primary schools increased from 53% in 1971 to 73% in 1991, middle schools from 8% to 19%, high schools from 4% to 8%. The District Primary Education Program (DPEP) in the mid-1990s increased funding for primary and upper-primary school building, hiring & training new teachers and decentralization of school administration. 160,000 new schools were built by 2000. This was followed by the Sarva Shiksha Abhiyan (SSA) in 2000s aimed at providing free and compulsory education for all children aged 6 to 14. By 2013, 98% of rural households had a primary school within 1 km and 92% had an upper primary school within 3 km. There has also been a huge increase in low-cost private schools across India in past few decades. Based on ASER data, 29% of children aged 6-14 in rural India attended fee-charging private schools in 2013.

3 Mechanisms

There are two key aspects to the characterization of returns to education in the marriage market: (i) marriage formation: trends in assortative mating (who marries whom with respect to education), and (ii) gains from marital surplus: payoffs to both genders within marriage conditional on educational attainment.

There are multiple mechanisms by which education could impact assortative matching in the marriage market. Schools and colleges act as matching platforms that reduce *search frictions* and enable matching in the marriage market (Kirkebøen, Leuven and Mogstad (2021)). This is less relevant in the Indian context where arranged marriages dominate. Individuals with similar educational attainment may share similar preferences and derive greater utility from partnership due to *consumption complementarities* (Becker (1981), Choo and Siow (2006), Lundberg and Pollak (2012), Chiappori, Dias and Meghir (2018)). Couples with similar (or strategically complementary) education levels can achieve greater household efficiency through specialization in labor and better coordination in raising and educating children (shown empirically in papers by Behrman and Rosenzweig (2002), Azam (2015), Sunder (2019), Beauchamp, Calvi and Fulford (2021)). This may be referred to as *production complementarities* (Becker (1981), Calvo, Lindenlaub and Reynoso (2024)). There is also a *mechanical effect* wherein as education levels rise for all, individuals are more likely to match with other more educated individuals mechanically without a change in assortativity. There are also *general equilibrium effects* due to assortative mating in the form of spillovers on individuals not directly treated. Hence, it is crucial to use a measure of assortative mating that accounts for the distribution of traits in the population (Eika, Mogstad and Zafar (2019), Anderberg et al. (2022)).

A distinction must be made between educational homogamy vs assortative mating. Changes in educational homogamy (fraction of couples with the same level of education) may be due to secular changes in educational attainment or due to changes in educational assortative mating.

To deal with this issue, I use a definition of assortative mating that accounts for shifts in the marginal distribution of education – men and women with same level of education marrying more or less frequently than at random in terms of education following Eika, Mogstad and Zafar (2019).

4 Contribution

There is an extensive literature on returns to education in the labor market. However, the literature on marriage market outcomes is fairly nascent. The latter is particularly consequential in the Indian context, where female labor force participation remains among the lowest in the world, raising the question of whether returns to education for women are channeled through the marriage market rather than direct participation in the labor market.

With respect to estimating labor market returns in a developing country context, Duflo (2001)'s paper is significant. It evaluates the impact of a massive primary school construction program that took place between 1973 and 1978 in Indonesia using a difference-in-differences design that exploits variation across regions and birth cohorts. Each additional school built per 1,000 children led to an increase in average educational attainment of 0.12 to 0.19 years and wage increases of 1.5 to 2.7 percent. In my paper, I follow the methodology set out in Duflo (2001) using a long panel. My results on educational attainment are in line with Duflo (2001) but I also study the effect on marriage market outcomes.

In the Indian context, the closest papers to this one are by Khanna (2023) and Agarwal, Bahure and Javadekar (2023) which study the effect of primary school construction on men's labor market outcomes and women's marriage outcomes respectively. Both studies use a regression discontinuity design that exploits district-level cutoffs for increased funding for school construction during the 1990s. The regression discontinuity estimand provided by Khanna (2023) and Agarwal, Bahure and Javadekar (2023) is a local treatment effect for districts with female literacy around the cutoff whereas the difference-in-differences specification I use exploits variation across all districts and cohorts. Khanna (2023) and Agarwal, Bahure and Javadekar (2023) also focus specifically on the District Primary Education Program (DPEP) expansion in the 1990s whereas I use school construction data over a longer time frame. My dataset also includes construction of all schools and not only primary schools. In contrast to Khanna (2023) who focuses on labor market outcomes, I study the effect of school construction on the marriage market as do Agarwal, Bahure and Javadekar (2023).

My results on educational attainment are slightly larger than Khanna (2023) who finds an increase of 0.09 years using National Sample Survey data but smaller than estimates in Agarwal, Bahure and Javadekar (2023) who find an increase of 0.32 or 1 year (depending on sharp or fuzzy RDD respectively) using DHS data similar to this study. It should be noted that comparing my estimates with theirs is not strictly appropriate since these represent different estimands. The regression discontinuity estimates in Khanna (2023) and Agarwal, Bahure and Javadekar (2023) identify a local average treatment effect of DPEP eligibility for districts at the female literacy cutoff. In contrast, my difference-in-differences design exploits continuous variation in school construction intensity across districts and cohorts over a much longer horizon, identifying an average marginal effect of additional school availability. Since DPEP represented a large, discrete expansion targeted at low female literacy districts with high marginal returns to schooling, RDD estimates are expected to be larger than the average marginal effects estimated here. Also, increased school funding under DPEP may have been used for teacher hiring or infrastructure upgrades in existing schools in addition to new school construction whereas my estimates reflect the effect of constructing an additional school alone.

With respect to marriage outcomes, Agarwal, Bahure and Javadekar (2023) counter-intuitively find a fall in age at marriage by 0.75-1.55 years whereas I find an increase of 0.04 years. The observed increase in age of marriage in my study could reflect an enrollment or incapacitation effect of schooling — girls remain unmarried simply because they are still attending school which is less likely to matter at the primary school level (except for rare instances of child marriage). It could also be the case that a change in marriage or partner preferences occurs at a later stage of schooling. The effects on educational attainment found by Agarwal, Bahure and Javadekar (2023) are also much larger than the effects found in other studies including this paper.

I also build an assortative mating index for India documenting an increase in segregation of individuals with no education into their own closed marriage market, during a time of expanding school construction. While I follow the methodology of Eika, Mogstad and Zafar (2019), this measure has not previously been constructed for India. Earlier studies use a reduced form method correlating education levels of husbands and wives, which ignores changes in the marginal distribution. My results on assortative mating trends mirror those found in the US data by Eika, Mogstad and Zafar (2019) re-iterating a pattern of increased stratification in marriage markets at the bottom of the education distribution.

I also estimate the effect of school construction on assortative mating. While prior studies such as Anderberg et al. (2022) and Royer and Geruso (2014) have studied the effect of schooling on matching (both use the Raising of School-Leaving Age in the UK in 1972), I am able to improve upon these studies by exploiting both the spatial and temporal dimension of school construction programs instead of only variation across cohorts at the time of the policy change.

I find that school construction is not flattening the marriage market. School construction instead increases educational sorting among couples by 0.03 as measured by the assortative matching index. A back-of-the-envelope calculation shows that school expansion accounts for approximately 28 percent of the observed rise in assortative mating during this time period.

5 Data

Data on school construction and locations is available from Department of Education's UDISE database. UDISE is a comprehensive database of all 1.5 million public and private schools in India and includes detailed data on many school-level variables. However, I only use data on school location and year of construction to construct my treatment variable. The UDISE data was collected in 2018 through school surveys but my dataset includes all schools built since 1960. One concern is that this dataset misses any schools that had shut down before the time of survey.

I use education and marriage data from India's Demographic and Health Survey (DHS), a large-scale nationally representative survey of women aged 15-49 years. Since district identifiers are only available from 2015 onwards, I use data from the 2019 survey round.

Population data was drawn from the decadal variation in population tables published by the Census of India and then linearly interpolated across years within decades. These tables account for boundary changes while estimating population within a district.

A key challenge in conducting multiple-year analyses at the district level in India is reconciling the many changes in state and district boundaries that have occurred since the first comprehensive census of independent India in 1961. A key contribution of this paper is construction of this long-run district-level panel.

6 Empirical Strategy

School construction in India is plausibly exogenous since it was governed by centralized population-and-distance norms under successive Five-Year Plans and later Sarva Shiksha Abhiyans, with funding tied to predetermined district infrastructure targets rather than contemporaneous local education demand or marriage-market conditions. Once built, schools represent a long-lived, fixed infrastructure investment, making reverse causality from cohort-specific marriage outcomes implausible. Also, as exposure is determined by district of childhood residence and the timing of school construction during ages 6–15, treatment variation is fully predetermined with respect to adulthood marriage outcomes. Migration at school ages is rare, and outcome realizations occur many years after treatment. While women do migrate at the time of marriage, marriage outside the district is rare according to Census data.

Specifically, I estimate the impact of school construction on the following outcomes for women – educational attainment (years of schooling, educational qualifications), not married by age x (extensive margin), and age at marriage. I also estimate the effect for couples on the age gap (across-cohort sorting) and the educational assortative mating index (intensive margin).

To assess the validity of the parallel trends assumption, I estimate a pre-trend specification that interacts school construction exposure with cohort indicators, restricting the sample to early cohorts whose schooling years largely predate the expansion in school construction. Specifically, I estimate regressions of the form

$$Y_{idt} = \sum_{\tau} \beta_{\tau} \left(\text{cohort}_{\tau} \times \text{treatment}_{d}\right) + \delta_d + \varepsilon_{idt},$$

where Y_{idt} is the outcome of interest, treatment_d is standardized school construction exposure, and δ_d are district fixed effects. This specification tests whether districts with higher eventual exposure to school construction exhibit differential trends in educational outcomes prior to treatment. The estimated interaction coefficients are small in magnitude and statistically insignificant (Table 1). In particular, the coefficients on the interaction terms for the early cohort bins are close to zero, and a joint test fails to reject that they are equal to zero ($p = 0.789$). These results provide no evidence of differential pre-trends across districts with varying levels of school construction exposure, supporting the identifying assumption that, absent school construction, outcomes would have evolved similarly across districts.

At the same time, this test should be interpreted with caution. Because school construction expanded gradually over time and exposure is measured cumulatively, there is no sharp policy discontinuity, and the definition of "pre-exposure" cohorts is necessarily approximate. In addition, this exercise focuses on a subset of earlier cohorts and may not fully capture differential trends for later cohorts or along other dimensions of heterogeneity. The specification also absorbs district fixed effects but not district-specific cohort trends, so the test cannot rule out all forms of slow-moving unobservables. Thus, while the absence of detectable pre-trends is consistent with the parallel trends assumption, it should be viewed as supportive rather than definitive evidence for the identification strategy.

Following Duflo (2001), I use a two-way fixed effects estimation specified as:

$$\mathrm{Y}_{\mathrm{idt}} = \beta(I_{\mathrm{d}} {*} C_{\mathrm{i}}) + \gamma_{\mathrm{t}} + \delta_{\mathrm{d}} + \varepsilon_{\mathrm{idt}}$$

Y_{idt} is outcome of individual i in district d born in year t, δ_d is district fixed effect, γ_t is cohort fixed effect, I_d is treatment intensity (number of schools per 1000 school-aged children) in district d, C_i dummy equals 1 for exposed cohort, 0 for unexposed cohort. The exposure window for a cohort is considered to be between 6 to 15 years – this means that a woman's degree of treatment exposure or dosage is determined by the number of schools in the district she resided in when she was between 6 to 15 years old.

I use a continuous treatment measure since the policy generated large heterogeneity in construction intensity across district–cohorts. I estimate the effects of adding one additional school per 1000 school-age children to a district. This represents an 11.3 percent increase relative to the sample mean of 8.885 schools per thousand students.

Following Duflo (2001), I also present results with a binary treatment where 1 indicates more schools constructed than the (cohort-specific) mean and 0 fewer schools than the mean. The results are qualitatively similar.

Changes in educational homogamy (fraction of couples with same level of education) may be due to secular changes in educational attainment or due to changes in educational assortative mating. I show these results in Table 4 where I regress husband's education outcomes on wife's treatment exposure during childhood.

In addition, I improve upon this by using a definition of assortative mating that accounts for shifts in the marginal distribution of education – men and women with same level of education marrying more or less frequently than at random in terms of education, following Eika, Mogstad and Zafar (2019). This is a structural assortative mating index which is free of marginal education distribution bias. Marital sorting between a woman with education level e_f & a man with education level e_m is defined as the observed probability that a woman with education level e_f is married to a man with education level e_m relative to the probability under random matching with respect to education.

$$s(e_f, e_m) = \frac{Pr(E_f = e_f, E_m = e_m)}{Pr(E_f = e_f)Pr(E_m = e_m)} \tag{1}$$

The sorting index equals 1 under random matching and an index >1 indicates positive assortative matching. The higher the index, the stronger the extent of assortative matching.

In order to construct the index empirically, I follow the following steps:

- I use data on currently married couples from DHS Individual Recode files: years of education (discretized into categories), year of cohabitation (5-year bins), district of residence, sampling weights.

- For each district d and marriage cohort t, I compute the weighted joint distribution of spouses' education categories:

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

  where E_i^f and E_i^m denote the wife's and husband's education categories for couple i, and w_i is the DHS sampling weight.

- From the joint distribution, I obtain the marginal distributions of wives' and husbands' education within each district–cohort cell:

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

- I define the assortative mating index for each education pair (e_f, e_m) as the ratio of the observed joint probability to the probability implied by random matching:

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

- I focus on same-education matches by restricting attention to the diagonal elements $s_{dt}(e, e)$, which capture the intensity of educational homogamy within district d and cohort t.

- Finally, I construct an aggregate assortative mating index for each district and cohort as a weighted average of the diagonal elements:

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

  This aggregate index summarizes the overall strength of educational sorting in the marriage market within district d and cohort t, net of changes in the marginal education distributions.

7 Results

Table 2 shows the effect of school construction on woman's own educational outcomes and Table 3 on marriage outcomes. The treatment intensity measure in Panel A is the number of cumulative schools built per school-aged population in a woman's district by a given year and district between 1961 and 2011.

Note that the school census was collected in 2018, schools that were constructed earlier but shut down before the survey are not observed. This likely causes historical school exposure to be measured with error, especially for older cohorts whose exposure window lies further in the past. To the extent that school closures are unrelated to later marriage outcomes conditional on district and cohort fixed effects, this undercounting should attenuate the estimated effects toward zero. However, because older schools are more likely to be missing from the data, the resulting measurement error may be non-classical, so some bias in either direction remains possible. However, reports indicate that the bulk of school closures has occurred in the past decade whereas the time period under consideration was one of school expansion.

In all regressions, I include district and cohort fixed effects and cluster standard errors at the district level. The initial level of schools is absorbed by district fixed effects. Throughout the paper, I estimate the effects of adding one additional school per 1000 school-age children to a district. This represents an 11.3 percent increase relative to the sample mean of 8.885 schools per thousand students.

I find significant effects on women's own educational attainment. Table 2 Panel A shows that an additional school built per 1,000 school-age children in a district is associated with 0.12 years more schooling for women in cohorts exposed to school construction. This represents a 1.8 percent increase, relative to the sample mean of 6.7 years of education. One additional district school during a woman's schooling years also significantly raises the likelihood of her completing primary schooling by 1.9 percentage points, relative to a mean completion rate of 64.0 percent. I do not detect a significant change in secondary schooling rates.

In Table 2 Panel B, I also show results using a binary treatment where 1 indicates schools higher than cohort-specific mean and 0 indicates lower than cohort-specific mean. The results are qualitatively similar for educational outcomes. Women in districts exposed to above-average school construction have 0.45 more years of education, which corresponds to a 6.8% increase relative to the mean. The likelihood of completing primary schooling is 4.3% higher, relative to a mean completion rate of 64.0 percent. Again, I do not detect a significant change in secondary schooling rates.

Table 3 shows the effect of school construction on a woman's age at marriage, the age gap between husbands and wives, and the probability for a woman not being married by age 18. I find small but statistically significant effects on all outcomes at the 5% level. School construction is associated with a rise in the age at marriage by 0.04 years whereas the mean age at marriage for the estimation sample was 19 years. The age gap between husbands and wives falls by 0.03 years (mean age gap is 4.8 years). [1] The probability of child marriage falls by 0.3 percentage points (64% of sample is unmarried by age 18). Panel B shows the results for the binary treatment – again, the results are qualitatively similar. Moving to a district with higher number of schools than the cohort-specific average from one with lower than average schools results in age at marriage being 0.28 years higher, age gap being 0.1 year lower and the probability of being unmarried by age 18 becomes 2.3 percentage points higher.

Table 4 shows results from a regression of the husband's educational outcomes where the treatment is a woman's exposure to school construction during childhood. While there is a positive effect on the husband's educational attainment, the effects are not significant except for secondary schooling completion in the specification with binary treatment. This pattern is consistent with marriage-market selection. As women's

[1] In order to calculate the age gap between husband and wife, I supplement the data on husband's age from the women's survey (which has many missing values) with data from the household roster, however the latter is only available for co-residing spouses. I also cross-validate the spousal age distribution in the household roster data with the subset of observations with non-missing values in the women's survey and find that the correlation is close to 1.

educational attainment rises, particularly at key thresholds, marriage markets re-sort such that more educated women match with more educated men. The fact that the effect appears at the secondary margin — where male educational attainment is more heterogeneous — supports an assortative matching interpretation. However, it should be noted that husbands' and wives' exposure to education to school construction may be highly correlated since they live in the same district (data on women's childhood residence is not available) and age gaps between couples have been narrowing over time resulting in treatment exposure occurring at the same time for men and women in more recent years. As a result, the estimates in Table 3 should not be interpreted as causal effects of women's education on men's schooling. Instead, they reflect equilibrium outcomes in local marriage markets where both men and women may have been exposed to educational expansion. Consistent with this interpretation, I focus on market-level measures of assortative mating in Table 6, which explicitly account for joint exposure and capture changes in sorting behavior net of changes in marginal educational attainment.

Table 5 shows the assortative mating index and average school construction exposure for 5-year binned marriage cohorts. In figure 4, I plot these trends in assortative mating for each educational category across marriage cohorts. Low-education individuals in India are increasingly segmented into internally homogeneous marriages, far beyond what education expansion alone would generate mechanically. For individuals with no education, the index starts around 1.2–1.4 in early 1980s, rises almost monotonically and reaches 5–5.7 by late 2010s. This may imply that non-educated individuals increasingly form a closed marriage market as education expands for others. This is structurally identical to what Eika, Mogstad and Zafar (2019) finds for the U.S. low-educated after 1980, but the magnitude here is much larger. Individuals with primary education also sort positively but are less sharply stratified than the no education group. The upper primary group has positive but weaker and fairly stable assortative mating. For high school group, there has been a decline in assortative mating in recent years. As education expands, high school may be becoming the new "mass" category, not an elite one, so homogamy weakens. Again, this mirrors the decline in high-skill assortative mating that Eika, Mogstad and Zafar (2019) document for U.S. college graduates after 1980. Overall, India's marriage market is becoming dramatically more segregated at the bottom of the education distribution, while assortative mating weakens at the high-school level. This may imply rising intergenerational persistence of low human capital, even as overall schooling expands.

I also estimate the causal effect of school construction on the educational assortative mating index. Table 6 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 increase of one unit in school construction exposure raises the assortative mating index by approximately 0.03 which represents a 1.4 percent increase per additional school. Note that the baseline level of the index indicates that India's marriage market is highly positively assortative even absent new schools – the mean of 2.1 indicates that couples with the same education level are about twice as common as they would be under random matching.

To assess 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 cohorts (1980–1989) and the late cohorts (2010–2019), the assortative mating index increases from 1.53 to 2.14, corresponding to a rise of 0.61. Over the same period, average school construction exposure increases from 5.08 to 10.77 schools per 1,000 children, a change of 5.70. Using the estimated coefficient from Table 6, this increase in exposure implies a predicted rise in the assortative mating index of 0.17. This accounts for approximately 28 percent of the observed increase in assortative mating over time. These results suggest that school construction contributed meaningfully to the strengthening of educational sorting in marriage markets, although a substantial share of the overall rise remains driven by other factors.

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. 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, strengthening educational sorting at the bottom of the distribution while weakening assortative matching at higher education levels where schooling becomes less scarce. In future work, I plan to look at the effect on intergenerational persistence of education to verify if these these trends in educational stratification in the marriage market persist through children's education.

References

Afridi, Farzana, Abhishek Arora, Diva Dhar, and Kanika Mahajan. 2023. "Women's Work, Social Norms and the Marriage Market."

Agarwal, Madhuri, Vikram Bahure, and Sayli Javadekar. 2023. "Marrying Young: The Surprising Effect of Education."

Anderberg, Dan, Jesper Bagger, V Bhaskar, and Tanya Wilson. 2022. "Marriage market equilibrium with matching on latent ability: Identification using a compulsory schooling expansion."

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. 2015. "Like Father, Like Son? Intergenerational Educational Mobility in India." *Demography*, 52(6): 1929–1959.

Beauchamp, Jonathan, Rossella Calvi, and Scott Fulford. 2021. "Women's Education and Child Health in Developing Countries." *World Development*, 140: 105371.

Becker, Gary S. 1981. *A Treatise on the Family.* Cambridge, MA:Harvard University Press. NBER Books, National Bureau of Economic Research, Inc.

Behrman, Jere, Andrew Foster, Mark Rosenzweig, and Prem Vashishtha. 1999. "Women's Schooling, Home Teaching, and Economic Growth." *Journal of Political Economy*, 107(4): 682–714.

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.

Bhaskar, V., et al. 2015. "Caste and Mate Selection in Modern India." *Economic Journal*, 125(586): 1051–1081.

Calvo, Paula A., Ilse Lindenlaub, and Ana Reynoso. 2024. "Marriage Market and Labour Market Sorting." *The Review of Economic Studies*, 91(6): 3316–3361.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir. 2018. "The Marriage Market, Labor Supply, and Education Choice." *Journal of Political Economy*, 126.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Desai, Sonalde, and Lester Andrist. 2010. "Gender Scripts and Age at Marriage in India." *Demography*, 47(3): 667–687.

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.

Khanna, Gaurav. 2023. "Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India." *Journal of Political Economy.* forthcoming.

Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad. 2021. "College as a Marriage Market." Discusses search frictions and local marriage markets due to college attendance.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2025. "The Child Penalty Atlas." *Review of Economic Studies*, 92: 3174–3207.

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.*

Mukherjee, Priya. 2017. "The Effects of Social Identity on Aspirations and Learning Outcomes: A Field Experiment in India."

Peters, H. Elizabeth. 2002. "Competing Premarital Investments and the Return to Education in the Marriage Market." *Journal of Human Resources*, 37(3): 679–697.

Royer, Heather, and Michael Geruso. 2014. "The Impact of Education on Fertility and Childbearing." *Journal of Labor Economics*, 32(4): 799–826.

Sunder, Naveen. 2019. "Parents' Schooling and Intergenerational Human Capital: Evidence from India." Job Market Paper, Cornell University.

9 Figures and tables

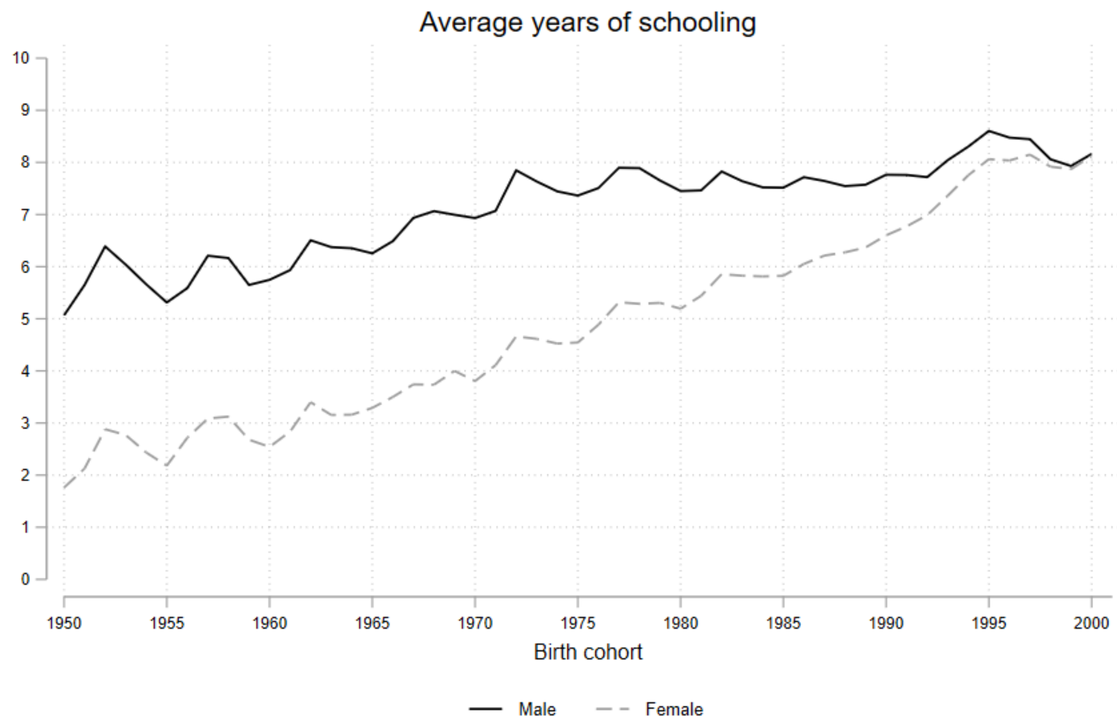

Figure 1: Average years of schooling by birth cohort

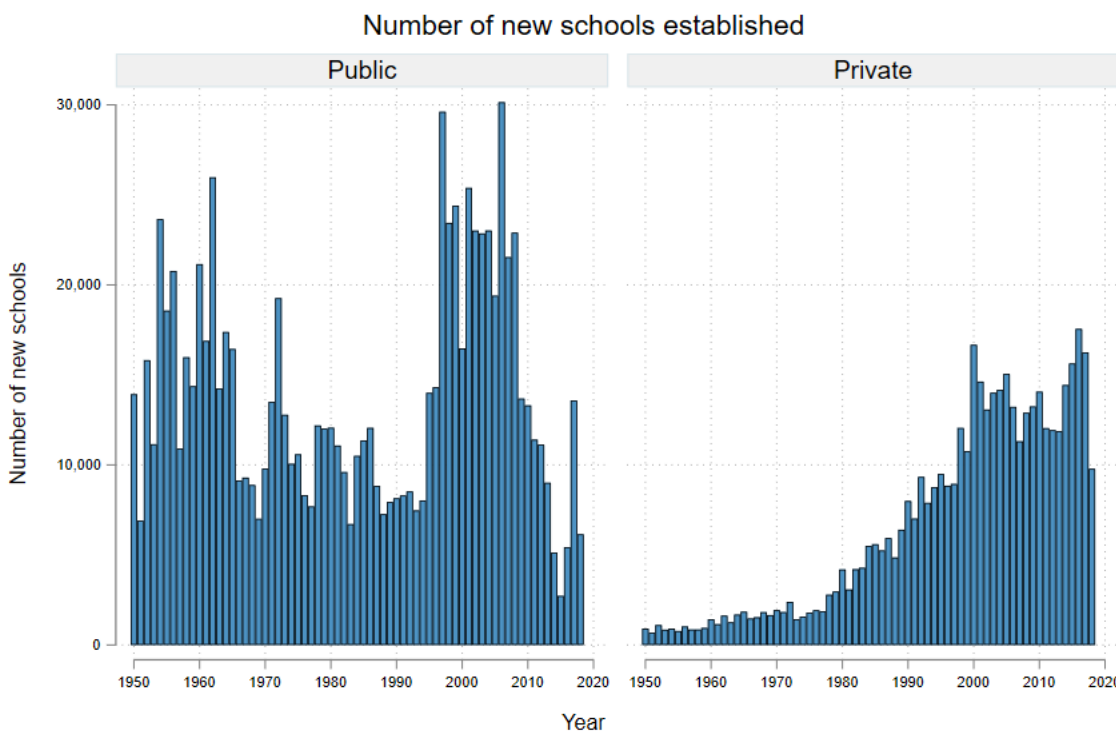

Figure 2: School construction over time (UDISE 2018)

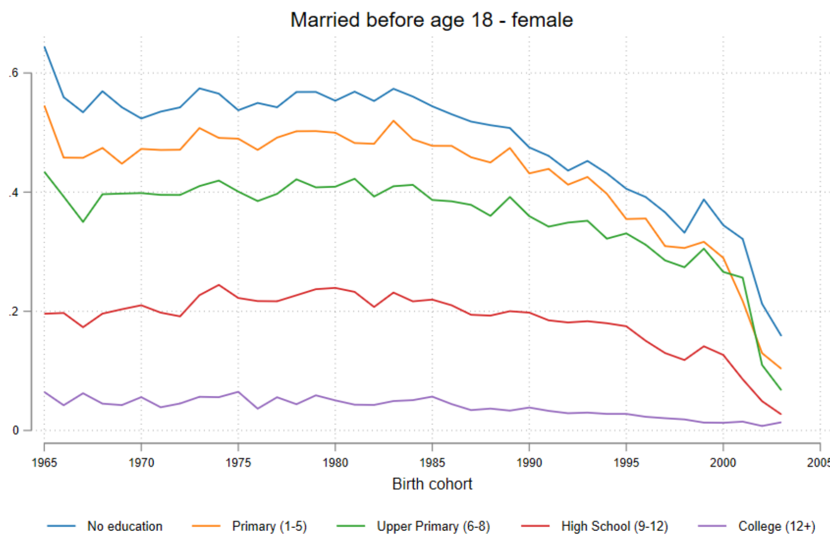

Figure 3: Child marriage

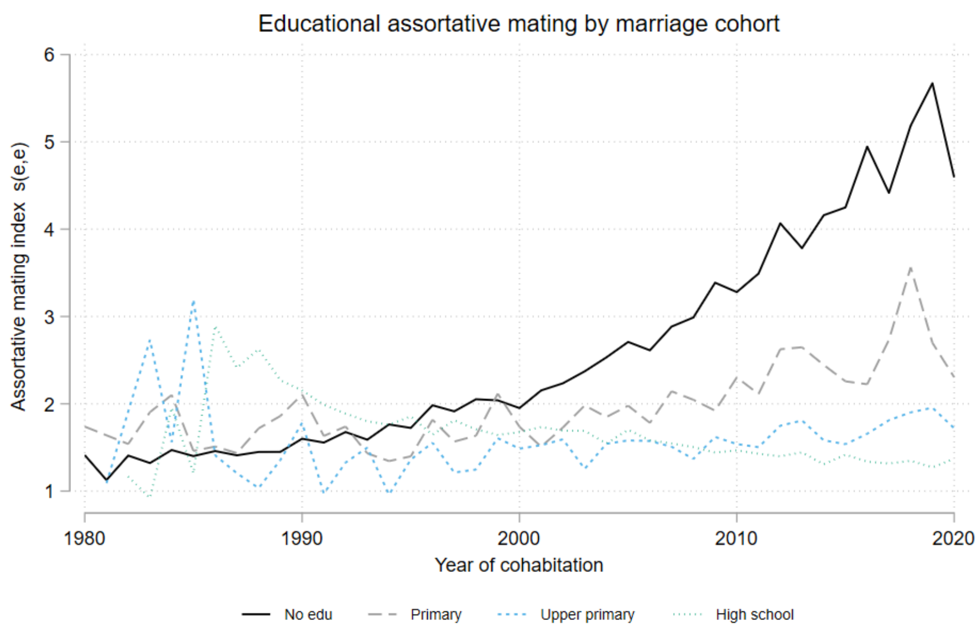

Figure 4: Assortative mating index (DHS 2019)

Table 1: Pre-Trend Test for Women's Years of Education

	(1)
	Years of education
Differential effect:	
School construction exposure × cohort bin	0.000
	(.)
School construction exposure × cohort (1970)	0.031
	(0.442)
School construction exposure × cohort (1975)	-0.059
	(0.451)
N	79928
Mean of dependent variable	4.216

Notes: The table reports estimates from a pre-trend specification that interacts standardized school construction exposure with early cohort-bin indicators, restricting the sample to older cohorts whose schooling years largely predate the main expansion in school construction. District fixed effects are included and standard errors are clustered at the district level. A joint test fails to reject that the interaction coefficients are equal to zero ($F = 0.24$, $p = 0.789$).

Table 2: Effect of School Construction on Women's Education Outcomes

	(1) Years of education	(2) Primary completed	(3) Secondary completed
Panel A: Continuous treatment			
Schools built	0.122***	0.019***	-0.001
	(0.021)	(0.002)	(0.002)
Mean	6.686	0.640	0.337
N	454,655	454,655	454,655
Panel B: Binary treatment			
Schools built > mean	0.433***	0.043***	0.007
	(0.107)	(0.008)	(0.011)
Mean	6.686	0.640	0.337
N	454,655	454,655	454,655

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. Standard errors (in parentheses) are clustered at the district level. Mean of cumulative school construction exposure in the estimation sample: 8.885.

Table 3: Effect of School Construction on Women's Marriage Outcomes

	(1) Age at marriage	(2) Age gap	(3) Not married by age 18
Panel A: Continuous treatment			
Schools built	0.036**	-0.025***	0.003**
	(0.015)	(0.009)	(0.002)
Mean	19.027	4.839	0.640
N	424,649	361,968	454,655
Panel B: Binary treatment			
Schools built > mean	0.282***	-0.102**	0.023**
	(0.080)	(0.042)	(0.010)
Mean	19.027	4.839	0.640
N	424,649	361,968	454,655

Notes: Each column reports estimates from regressions of the dependent variable on school construction exposure, with district and cohort fixed effects. In columns (1)–(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. Standard errors (in parentheses) are clustered at the district level. Mean of cumulative school construction exposure in the estimation sample: 8.604.

Table 4: Effect of School Construction in Wife's District on Husband's Education Outcomes

	(1) Years of education	(2) Primary completed	(3) Secondary completed
Panel A: Continuous treatment			
Schools built	0.050	0.003	0.002
	(0.035)	(0.002)	(0.003)
Mean	8.065	0.746	0.416
N	63,815	63,815	63,815
Panel B: Binary treatment			
Schools built > mean	0.393	0.002	0.042***
	(0.250)	(0.014)	(0.013)
Mean	8.065	0.746	0.416
N	63,815	63,815	63,815

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. Standard errors (in parentheses) are clustered at the district level. Mean of cumulative school construction exposure in the estimation sample: 8.634.

Table 5: Trends in Assortative Mating and School Construction over Marriage Cohorts

Binned marriage cohorts	Assortative mating index	Mean school construction exposure
1980	1.307	4.841
1985	1.762	5.315
1990	2.073	6.231
1995	2.212	7.098
2000	2.355	8.409
2005	2.327	9.526
2010	2.326	10.491
2015	1.962	11.058

Notes: Table shows assortative mating index S^{agg} and mean school construction exposure for each of the 5-year binned marriage cohorts.

Table 6: Effect of School Construction on Educational Assortative Mating Index

	(1)
	Assortative mating index
Schools built	0.030***
	(0.012)
Mean	2.104
N	4,646

Notes: District-by-marriage-cohort regression with district and cohort fixed effects. Cells weighted by number of couples. Mean of cumulative school construction exposure in estimation sample: 7.896.

A Appendix: 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 process of construction of a longer district level panel. One contribution of this study is the construction of a consistent district panel over a long time period 1961-2011. I use the 640 districts in the 2011 Census which also gives me a larger district level sample than other studies that only use unchanged districts.

If district boundary changes are not carefully accounted for, population data may have large measurement errors. To deal with this issue, I use the harmonized population counts across districts drawn from "Decadal Variation in Population" tables which account for district boundary changes over time which are provided by the Census of India. I then interpolate these population estimates across inter-censal years.

For data on outcomes, 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 is from the UDISE school database. I scraped geocodes for all 1.5 million schools and mapped it to DHS clusters using QGIS. I then map DHS households to DHS clusters which are now linked with Census district boundaries.

B Appendix: For Online Publication

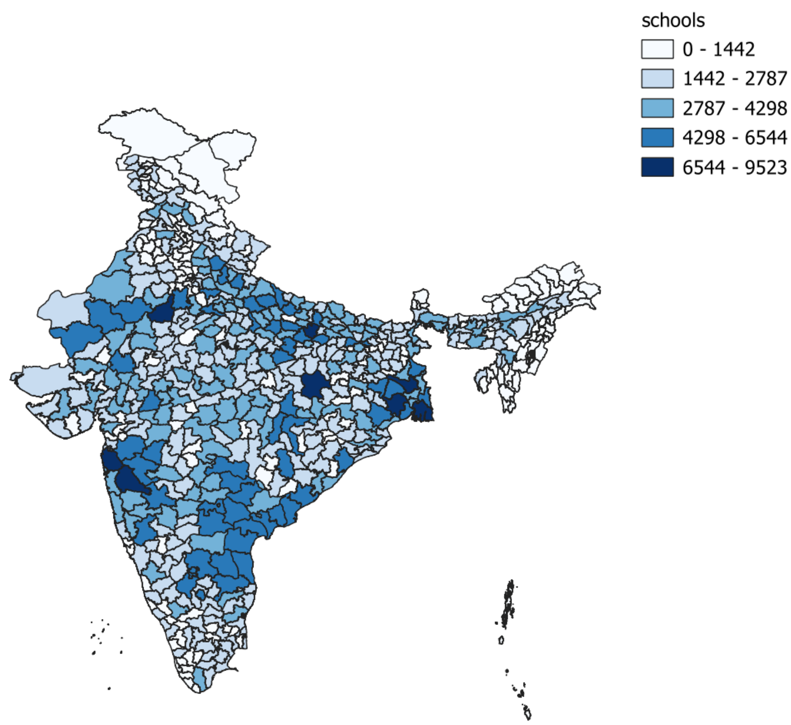

Figure 5: School construction across districts (UDISE 2018)

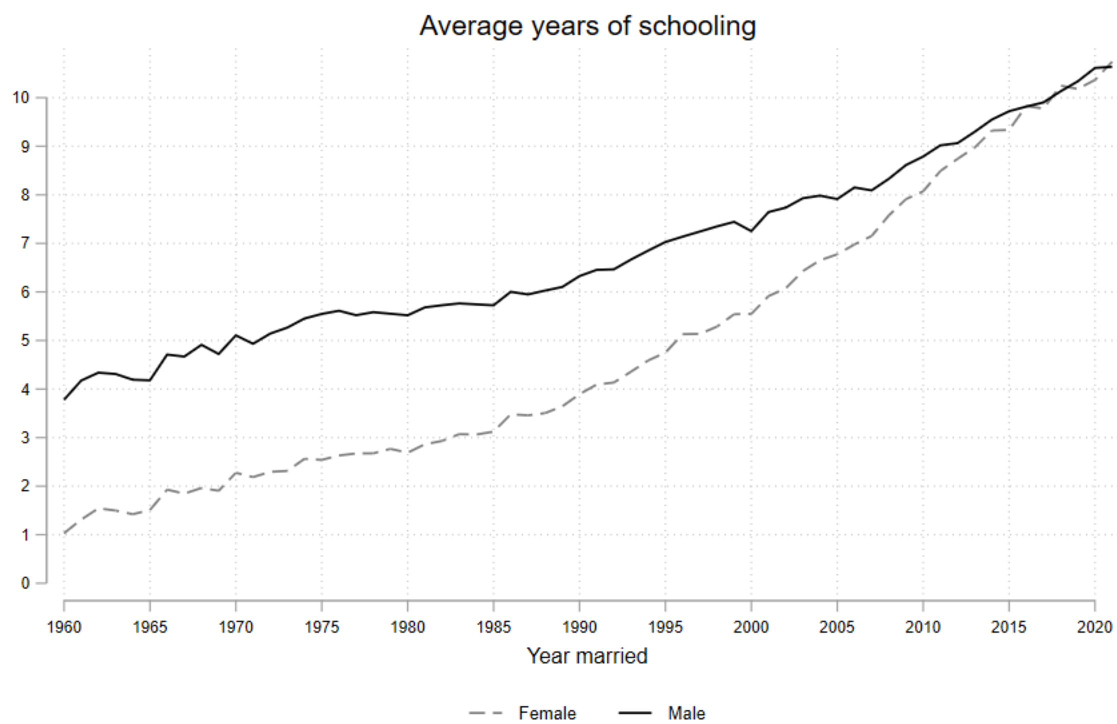

Figure 6: Average years of schooling among married couples

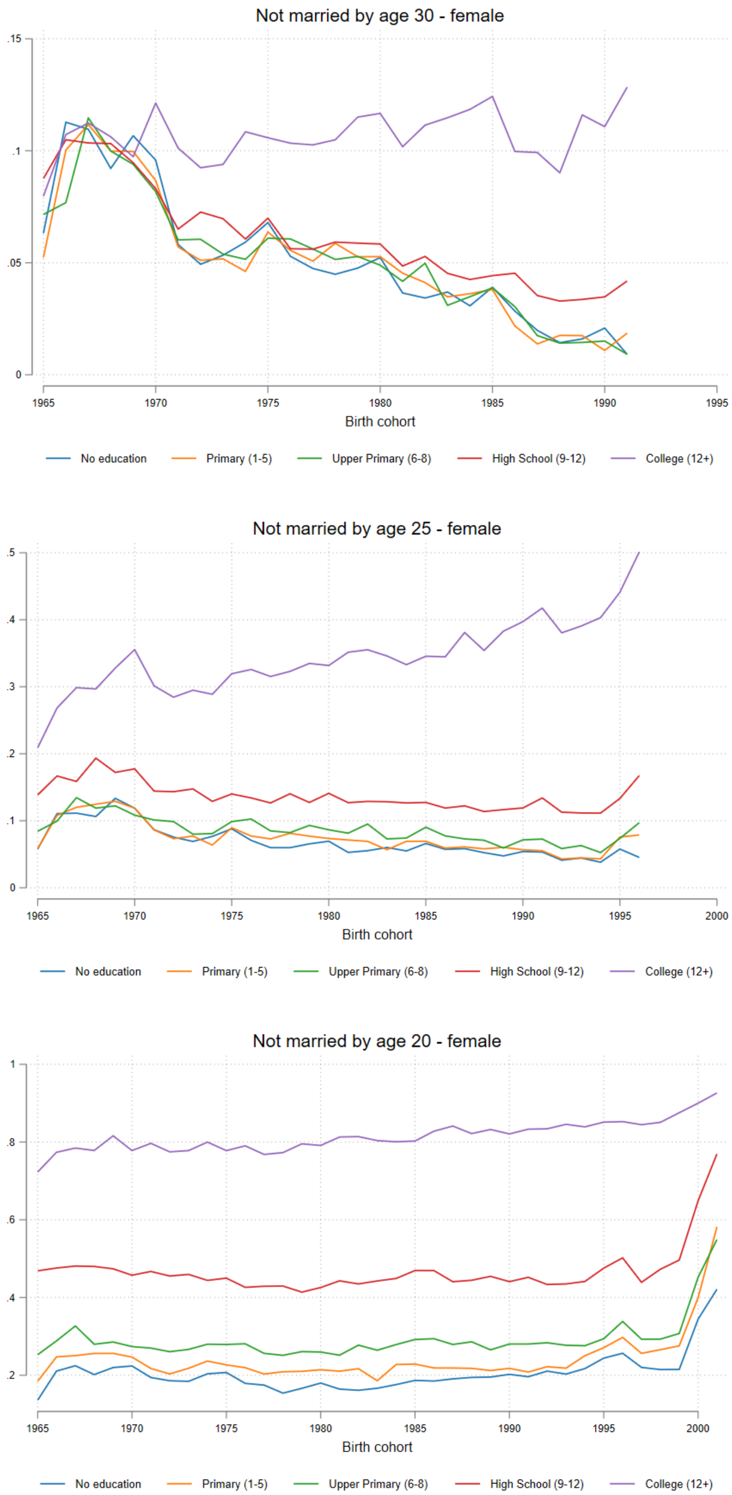

Figure 7: Not married by given age – women

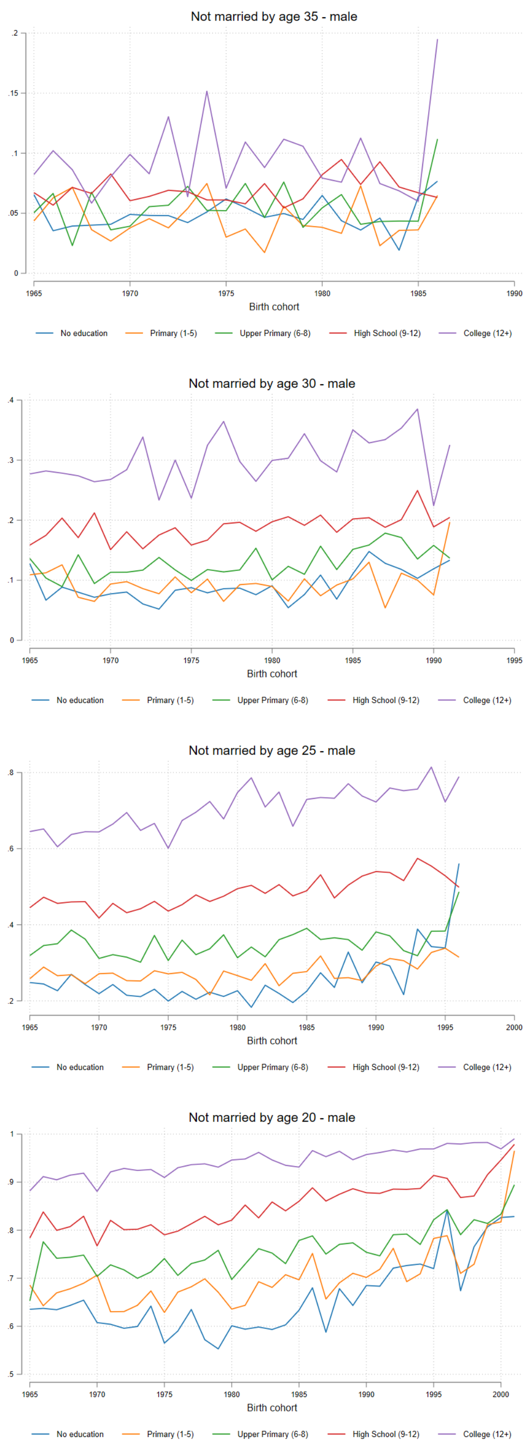

Figure 8: Not married by given age – men