Dynamics of School Expansion and Inter-Caste Marriages in India

Working Paper

Nitin Kumar Bharti[†], Luca Maria Pesando[‡],
Koyel Sarkar[§], Neha Bailwal [¶]

September 15, 2023

Abstract

Marriage in India is largely controlled by hanging norms such as arranged marriage and sociocultural markers such as the caste system. Caste endogamy, i.e., marrying within the same caste, remains one of the strongest pillars of the caste system in the Indian society, with close to 86% of endogamous marriages in the 2020 marriage cohort. We perform a causal analysis of whether school expansion in rural India had any impact on the increase in inter-caste marriages (ICM), exploiting variation in school openings across different locations at different times. A 1 standard deviation (SD) change in school openings (per village) increases ICM by 5.67%. Exploring the underlying mechanisms, we do not find completed years of education to be the driver, suggesting contact theory, delayed age at marriage, and/or education assortativity are likely mechanisms. These results suggest that education, development, and broader modernization forces can be a driver of family change in India, albeit their relevance may be lower than in other low- and middle-income countries (LMICs).

Keywords: Inter-caste Marriage, Primary School Expansion, India
JEL code:.

*We would like to thank Dimple Chopde for her brilliant research assistance. We would also like to acknowledge financial support from the Global Dynamics Research Cluster at NYU Abu Dhabi, for funding the project. We also thank all the participants at the Indian Population Conference held at NYU-Abu Dhabi for providing helpful comments and suggestions.

[†]Email: nkb2039@nyu.edu. New York University Abu Dhabi.
[‡]Email: lucamaria.pesando@nyu.edu. New York University Abu Dhabi.
[§]Email: koyel.sarkar@nyu.edu. New York University Abu Dhabi.
[¶]Email: neha.bailwal95@gmail.com. Indian Statistical Institute Delhi.

1 Introduction

Divisions along caste lines remain a strong defining feature of Indian society. While several traditional systems of social stratification, such as slavery, racism, gender, and class are undergoing relatively rapid changes across the world, the caste system seems to retain its power and to keep regulating the lives of over 1.4 billion people irrespective of religion and ethnicity (Gundemeda, 2020; Vaid, 2014). One of the key characteristics of the caste system is the closed system of marriage or "caste endogamy" (Davis, 1941). Inter-caste marriages (ICM) are still relatively rare in rural and urban India. According to Caldwell et al. (1998, p. 146), even though "some erosion of arranged marriage has begun [...] and an increase has occurred in cross-caste marriage," these marriages still tend to be between castes of a similar hierarchical level (Caldwell et al., 1983). Despite increase in ICM and uneven prevalence across states, the desire to marry within caste remains strong (Banerjee et al., 2013). Open questions remain about the kind of societal changes and/or actual policies that may contribute to weakening the caste-endogamy system. Goli et al. (2013) suggests that increase in mixed marriage in India is associated with changing socioeconomic and cultural values, alongside the impact of Western education and economic diffusion. Despite knowing the importance of education in understanding and appreciating the diverse perspectives, promoting social integration and interaction, little is known about its impact on weakening caste boundaries. Therefore, in this paper, we look at the impact of education expansion on ICM in rural India.

The education sector in India has expanded massively since independence.[1] To attain universalization of elementary education, the main focus was on improving access to primary schools within each habitation, substantially expanding primary schools in rural India. This rapid proliferation of primary schools in rural areas has played a an important role in India achieving universal primary school enrollment in last few decades. Though the expansion of education is linked to developing a knowledge-based economy (Mander and Prasad, 2014), its impact on the dynamics of family formation is yet to be thoroughly assessed.

In this paper, we study whether exposure to primary school expansion increases the probability of ICM in India. Using longitudinal data on school openings from the District Information System for Education (DISE) merged with the Indian Census 2011 and two waves of the National Family Health Survey (NFHS), we address this question leveraging vari-

[1] Number of schools has grown more than ten times since independence, from about 141,000 to over 1.5 million (Ministry of Education 2021).

ation in primary school expansion across different locations at different times. We do so by focusing on rural areas – those that experienced "positive shocks" in school expansion at the right school-going age– as well as on men. [2]

We find that a one-unit increase in the number of primary schools within a village in a 5 km NFHS cluster during an individual's primary school-going age (0-9 years) leads to an increase in education and age of marriage for men by 0.6 years and 0.54 years, respectively. Additionally, the effect of primary school expansion on ICM is positive and statistically significant. Specifically, a one-standard deviation increase in school expansion raises ICM by 5.67 % of the sample mean. Notably, increased education years and delayed age of marriage do not appear to be potential channels, suggesting that contact theory may provide a plausible explanation.

Assessing the implications of schooling expansion on inter-caste marriage is fundamental for at least two reasons. First, inter-caste marriage is a family domain that has shown high inertia over the past century, leading to a reinforcement of the common perception that family forms and structures in India only weakly respond to development, urbanization, and modernization forces that are pervading low- and middle-income countries (LMICs). This aspect has key implications for policymaking, as it would suggest that social policies targeted towards families would do little to shape intrinsic features of the family system. Conversely, should schooling have the – at least minor – potential to affect socio-cultural markers of the Indian family system, then policymakers may push towards redirecting resources to promote, strengthen, and sustain human capital accumulation across the life course and across all strata of society – thus identifying a clear and effective policy lever.

Second, it is undeniable that the persistence of the caste system and, especially, the sustained high prevalence of caste endogamy in marriage has huge implications for societal inequalities in that it minimizes social mixing and integration, reinforcing societal homogeneity and contributing to the reproduction of intergenerational inequalities (Blau et al., 1984; Golebiowska, 2007; Goli et al., 2013; Singh et al., 2023; Munshi, 2017). ICM not only directly weaken the foundation of caste system but also lessen the social distance between minority and majority groups (Sharma, 2019; Song, 2009). Higher social integration and societal heterogeneity have been shown to correlate importantly with greater solidarity, trust, and higher social mobility (Roth and Peck, 1951; Schwartz et al., 2016). Therefore, our study has profound implications for the study of intra- and intergenerational social

[2]Majority of women migrate after marriage, and surveys do not capture childhood information, not allowing us to know their exposure to opening up of schools.

mobility in the Indian context.

Our study makes following important contributions to the existing literature. First, extant literature on education and ICM in India mainly focused on the role of parents education and its association with ICM (Ray et al., 2020; Sharma, 2019). Our study contributes to this discussion by highlighting how a spouse's exposure to school expansion during early childhood, particularly men, can influence ICM in rural India.

Second, while micro-level studies in India and other countries show a positive link between expanding education and changing marriage practices, including inter-caste marriages (Sarkar, 2022; Dommaraju, 2010; Medhe, 2019; Sharma, 2019; Singh et al., 2023), it remains uncertain whether this relationship is causal. Therefore, our study contribute to this strand of literature by providing casual impact of expansion of education on ICM using cohort analysis and two way fixed effects. Similar strategies leveraging school-supply "shocks" have been adopted to study polygamy in Cameroon (André and Dupraz, 2023), child health in Taiwan (Chou et al., 2010), wages and education in Indonesia (Duflo, 2001), and India (Khanna, 2023), and choice of school track in France (Garrouste and Zaiem, 2020).

The economics literature on the marriage market in India has primarily centered on dowry payments and marriage patterns (Anukriti et al., 2022; Chiplunkar and Weaver, 2023). Our paper extends this limited empirical literature on the marriage market by examining the dynamics of school expansion and inter-caste marriage (ICM) in rural India.

This paper also contributes to the growing literature on early childhood intervention and its long-term economic implications (Bharadwaj et al., 2013; Boucher et al., 2020; Bharti and Roy, 2023). We enhance this literature by investigating how exposure to school expansion during the formative years of men influences marriage market outcomes, specifically ICM. Our findings underscore the importance of targeted interventions in early childhood to weaken caste divisions in rural India.

Lastly, previous studies at the national level that have explored the connection between education and ICM have relied on older datasets. Our study offers a more updated picture on the prevalence and trends of ICM, utilizing some of the latest datasets available compared to earlier studies.

This study is structured as follows. Section 2 provides context — recent trends in changes

in families — and data. Section 3 provides some descriptive statistics of the sample. Section 4 describes empirical strategy. Section 5 discusses the results in detail along with important heterogeneity. Section 6 concludes with discussion and policy implications.

2 Context and Data

2.1 Context

Marriage is treated as an alliance between two families rather than two individuals (Béteille, 1991), making parents the central figure in decision-making. Marriages are typically arranged by the family with careful consideration of caste, religion, language, and other ethnic and socioeconomic attributes of both spouses while discouraging contact between them prior to marriage.[3]

2.1.1 Family Change in India

There have been some notable changes over time and space, such as the nuclearization of families (Allendorf, 2013), increased involvement of couples in marriage-related decisions, and the importance given to love and compatibility even within arranged marriage settings (Caldwell et al., 1983; Uberoi, 2006; Fuller and Narasimhan, 2008). These have been attributed to the rise in education levels, employment, and urban living; family ties gradually loosened and changes in beliefs, norms, and practices (Ross, 1961; Kannan, 1963; Corwin, 1977; Thornton, 2001).

A major reported shift has occurred in the increasing age at marriage. The mean age at marriage was 18 years for women and 23 years for men in the 1991 census (Bhagat, 2016), which increased by two years in the 2011 census, elevating the median age at marriage to 20 years for women and 25 years for men. Marriages below the legal age for women (18 years) have dropped from 47% for age group 45-49 to 23% for age group 20-24, while marriages below the legal age for men (21 years) dropped from 27% for age group 45-49 to 18% for age group 20-24 years according to the NFHS 2019-21 report. The revival of intergenerational co-residence is another important development that could hinder shifts in couple formation from traditional norms to modern ways, including ICM, as studies have found household factors, such as household size, economic status, and parental education (Ray et al., 2020; Goli et al., 2013), as important factors driving these changes.

[3]Premarital courtships and love marriages were largely discouraged as these often led to irrational partnering outside the well-formulated norms.

2.1.2 Prevalence and trends in inter-caste marriage

Encouragement of ICM has been one of the tools of Indian governance systems to usher in multiculturalism, eradicate untouchability, and facilitate social integration. Inter-caste marriages were written into law in India by the Special Marriage Act of 1872 and recognized by the Hindu Marriage Act of 1955. In 2011, the Supreme Court of India declared that ICMs are in the national interest and provide a unifying factor for the nation, as there has never been a bar on inter-caste or inter-religious marriages in independent India.[4] These statements even translated into policy initiatives, such as the "Dr. Ambedkar Scheme for Social Integration through Inter-Caste Marriages 2013", with a monetary incentive to inter-caste couples. Regardless of such policies, efforts, statements, and public acknowledgments, most marriages continue to take place within the same caste, due to widespread opposition in the society.[5]

Figure 1 shows a close-to-monotonic increase in ICM from 10.4 in the earliest marriage cohort (1981-85) to 14.8 in the latest marriage cohort (2016-2020) using two waves of NFHS data from the whole of India. A 4.4 percentage-point increase over 40 years may be deemed negligible, yet it corresponds to a 42 percent increase.

IHDS datasets, using jati (definition for caste), the ICM stands at 5%, with state-level variation such as Punjab (12%), West Bengal (9%), and Gujarat (8%). ICMs are found to be higher in urban areas, among the economically, educationally, and culturally advanced groups (Goli et al., 2013) .[6] Focusing on the second round of the IHDS, Sharma (2019) finds the proportion of women marrying beyond caste boundaries to be higher for younger (less than 19) as well as older (above 30) brides.

2.1.3 Education and ICM

Education has the potential to mitigate rooted prejudices and stereotypes, and educational institutions can serve as platforms for social mixing and social integration, in turn shaping perceptions and attitudes toward the "other" (Ray et al., 2020). There are broadly 4

[4] In 2018, the Court reiterated that marriage between consenting adults is fully legal, notwithstanding their caste, and underlined the rights of adults to choose their partners as well as the need for society to learn to accept inter-caste and inter-faith marriages.

[5] 41% of adults in Delhi and 62% of adults in UP were in favor of laws banning intermarriage between high and low caste (Hathi, 2019).

[6] (Munshi and Rosenzweig, 2009), through their primary data, also finds similar estimates: among 25-40-year-olds, ICM was 7.6% in Mumbai in 2001, 6.2% in South Indian tea plantations in 2003, and 5.8% for the rural Indian population in 16 major states of India in 1999.

perspectives on why education can impact ICM:

First, according to the "enlightenment perspective," more education gives the ability to better analyze and understand different points of view, processing information through value systems that go beyond their own (Deary et al., 2008; Schoon et al., 2010), thus generating more tolerance towards others (Hathi, 2019; Hodson and Busseri, 2012).[7] Second, the "cultural adaptability hypothesis" postulates that individuals with higher levels of education and cognitive ability tend to have stronger motivation to adapt to foreign cultures and are more inclined to change the behavior patterns used in their original culture to adapt to the circumstances of the new environment, while lower-educated individuals appear more "passive" when adapting to foreign cultures (chun Lin et al., 2012; Ray et al., 2020).
Third, the "contact theory," popular in sociology and psychology, suggests that intergroup contact may effectively reduce prejudice and conflicts between majority and minority groups, leading to more acceptance and social mixing (Allport, 1979). Educational institutions provide a platform for social mixing by boosting intergroup contact. Fourth, "assortative matching behavior," more educated males will tend to marry more educated females. In a world of expanding education, the overall impact on ICM can go in any direction (Ray et al., 2020; Furtado, 2012). If the group's average education is higher than that of the overall population, a more educated individual from that group may "marry in."

The first three suggest a positive relationship between increasing education and ICM, whereas the fourth is unclear. The societal attitude captured through SARI, seems also mixed; there is widespread opposition by high-caste individuals to intermarriage with low-caste individuals, yet less opposition to intermarriage among higher-educated individuals (Hathi, 2019). Hence, a priori, the relationship is not straightforward.

3 Data

To assess the impact of school expansion on inter-caste marriages in rural India, we employ three separate datasets: (1) National Family Health Survey (NFHS) for the years 2015-16 and 2019-21, (2) District Information System for Education (DISE) data for 2011, and (3) Census data from the Office of the Registrar General and Census Commissioner,

[7] It follows from the argument that the roots of prejudice and intergroup antagonism are uninformed worldviews; thus, less education and lower cognitive ability are linked with starker prejudices, attitudes, and stereotypes (Wodtke, 2012).

Government of India, for the year 2011.

3.1 Data on inter-caste marriage

The NFHS is a nationally representative household survey offering comprehensive information on various aspects such as population, health, marriage, fertility, and nutrition. The household selection follows a stratified two-stage sampling approach. In the initial stage, Primary Sampling Units (PSUs), also referred to as clusters (equivalent to villages) are selected from the 2011 Census list using probability proportional to size. In the subsequent stage, an equal number of households (20) are randomly chosen within each PSU.

We use two rounds, namely NFHS-4 conducted in 2015-16 and NFHS-5 conducted in 2019-21, to get inter-caste couples for our analysis. The survey interviews all women aged 15-49 and men aged 15-54 within the selected households. Our analysis focuses exclusively on men, considering that women's village of residence tends to change following marriage.[8] We use the couple-level file in the NFHS, asking both husbands and wives the same questions about marriage. Our dependent variable is whether the spouses belong to different caste groups. We restrict the sample to rural areas, with men above 22 years and older.

The NFHS also provides the geographical coordinates of each cluster. However, these coordinates are randomly displaced within a 5-kilometer radius to ensure respondent privacy.[9] Since we do not know the exact cluster, we create a 5km radius buffer zone around the given geographical coordinate of a cluster and create the main explanatory variable at this buffer zone level.[10]

3.2 Data on schools

We rely on the DISE-2011 data to look at the expansion of schools over time. DISE is an annual data set that covers a universe of schools (nearly 12 million) all over India, with school-level information, such as the year of opening of schools, type/levels of schools (primary, secondary, higher secondary), school management structures (Department of Education, Tribal/ Social Welfare, local body, private aided, etc.), and school's highest and

[8]Evidence from Indian Human Development Survey 2011-2012 (IHDS II) data suggests that migration among women is higher in rural India with more than 80 percent of women reporting that their childhood residence is different from their current village of residence.

[9]The geographic coordinates of 99% of the clusters in the rural stratum were displaced up to 5 km, with 1% being displaced up to 10km. The displacement was restricted to keep the clusters in the same districts.

[10]There is 99% chance that the actual surveyed village is within this buffer zone. Further, to avoid the measurement error, we only keep those villages, with at least 10% of their area falling into the 5km buffer zone.

lowest grade levels. For our analysis, we focus on primary schools only.

Furthermore, it incorporates geographical details about the school, including district, block, and village names, which link this data with the 2011 Population Census. However, without unique village-level codes in the DISE dataset, the linking process involves name-based fuzzy matching. The fuzzy matching of village names is done within a given district and subdistrict, to be precise. Further, a thorough manual examination is carried out to keep precise matches.

This merged DISE-Census data is further combined with NFHS 5km zone data using unique village-level census codes, and we keep only the relevant villages, i.e., falling within the 5km buffer zone. This provides us with the school-level information within the buffer zones. Using the year of opening of schools, we create an annual balanced panel at the buffer zone ("DHS-5km cluster") of the number of primary schools ranging from 1950 to 2012.

Lastly, we combine the couple dataset with the above-created panel dataset to compute our main regressor — change in the number of schools in the husband's first 9 years of age.[11]

3.3 Village level characteristics

We obtained village-level characteristics from the latest round of Census data, which provides comprehensive information on population composition and village amenities. Our village-level controls are constructed using primary census abstract and village directory data. To conduct our analysis, we merged the Census data with DISE data at the village level. It is important to note that there is no unique village identifier available that can be used to merge Census and DISE files. Therefore, we use fuzzy matching techniques to merge 2011 Census data with DISE data based on the state, district, block, and village name. Total number of Census villages are 6,40,950 in 2011. However, 43,330 villages are zero populated. We drop all such villages from our analysis. Since DISE covers only the villages that have at least one school, all the Census villages without the primary school also get dropped from our sample. Finally, based on fuzzy matching, we could match 404,926 number of villages. Subsequently, we merge this school and village-level data with NFHS by utilising the geographical coordinates.

[11] It is simply the difference between the number of schools at age 9 and the number of schools at age 0 (i.e., birth year)

4 Descriptive Statistics

In the sample, the proportion of ICM is notably lower compared to marriages within the same caste. Overall, the rate of ICM exhibits an upward trajectory between 1981 and 2021, driven by all the caste categories, except a slight declining trend in the General Caste for the last marriage cohort (Figure 1).

Appendix Figure A.I shows the cross-tabulations by husband and wife's caste, providing a glimpse into ICMs. As the presumed gap between the caste group increases, the rate of ICM declines. For example, 83% SC men marry SC women compared to a meager 8.8% marrying to OBC; 5.8% to ST, and only 2% to General caste. Conversely, when the husband is from the SC caste, the inter-caste marriage share shows a declining trend. Similarly, 88% ST men marry within their own caste, 7% to SC, 4% to OBC, and 1.2% to General.

In figure A.II, we look at the descriptive visualization of the relationship between inter-caste marriage rates and school exposure across different birth cohorts in rural India. When schools are fewer (below the median), it indicates lower school exposure. In this case, we observe a consistently higher level of ICM rates and school exposure. Notably, the increase in ICM rates appears more pronounced within specific birth cohorts, highlighting that the impact of school exposure on inter-caste marriages varies across different generations. To sum up, there is a positive association between school exposure and inter-caste marriage rates. Our empirical analyses that follow attempt to identify this pattern causally.

5 Empirical Strategy

To identify the effect of school expansion on inter-caste marriage, we need exogenous variation in school expansion such that individuals are assigned randomly to the change in number of primary schools in a cluster. We use a combination of two exogenous variations to estimate the causal effect. We first exploit variation in school expansion across different birth cohorts within the same cluster. Our second source of variation is coming from men in the same birth cohort exposed to different school expansions in different clusters.

5.1 Baseline Specification

Our baseline specification is as follows:

$$y_{idb} = \alpha + \beta post_{db} + \beta_1 post_{d,b-1} + \delta_d + \rho_b + X_i + \epsilon_{idb} \quad (1)$$

where, y_{idb} is 1 for inter-caste marriage for individual i born (and living) in NFHS-5km cluster d, belonging to birth cohort b. Our main coefficient of interest is β, which captures the causal impact of primary school expansion on inter-caste marriage. The covariate $post_{db}$ is a continuous variable capturing the intensity of the primary school expansion for individuals i aged 0-9 years. It is measured as the change in the number of primary schools per village[12] for individual i in the first nine years of the birth (i.e., during age 0 to 9 years). The assumption is that opening new primary schools will impact the cohort who was 0-9 years old (the age of attending primary school); the later cohort will miss the opportunity to attend newly opened primary schools. δ_d and ρ_b are the cluster and birth cohort fixed effects. The covariate $post_{d,b-1}$ is the number of primary schools per village one year before the birth- capturing the impact of the prevalence of existing primary schools on individual i. X_i includes individual-level controls like dummy for religion (Hindu and Muslim) and household current wealth.

Our main identifying assumption is that change in the number of primary schools in a cluster is exogenous conditional on cluster and birth year fixed effect. We control for unobserved differences in birth year by including birth year fixed effect (ρ_t). Clusters with large school expansion may be different from clusters that experienced less change in the number of schools. We take into account such unobserved differences in clusters by controlling for cluster fixed effects captured by δ_d.

While we believe that change in the number of primary schools in a cluster is exogenous, there could be a concern that individuals may migrate to clusters with large school expansion. Given the low level of migration for men in rural India, we believe that migration is less of a concern in our setting (Munshi and Rosenzweig, 2016).[13]

6 Main Impact of the Expansion of Primary Schools

This section of the paper presents our main results. Table 1 displays the impact of school expansion on variables of interest — years of education, the age of marriage for husbands,

[12] We normalize the change in the number of schools with the number of villages falling in the NFHS cluster as per census 2011. It is to take into account different sizes of NFHS clusters. It is synonymous with normalizing with respect to an area of the cluster. The other possible contender, being population, has the problem of reverse causality (**cite**)

[13] Data from IHDS reveals that over 90 percent of the sampled households have been residing in their current village of residence for more than 50 years."

the age of marriage of wives, and inter-caste marriage — using our baseline specification.

Column (1) of Table 1 shows that a one-unit increase in primary schools in a village in a 5 km NFHS cluster during an individual's primary school-going age (0-9 years) results in an increase of 0.6 years in education on men (8% increase over the mean) at 1% significance level. It is important to highlight that it is not the total impact of the existence of schools, but the opening up of new schools conditional on the existing stock of schools. The β_1 coefficient captures the impact of the existing stock of schools per village, which is 0.439 at 5% significance level (6% over the sample mean).

Column (2) of Table 1 shows that a one-unit increase in primary schools in a village in a 5 km NFHS cluster during an individual's primary school-going age (0-9 years) results in an increase of 0.54 years in the age of marriage on men (2.4% increase over the mean) at 1% significance level. However, we do not observe any statistically significant impact on wives' marriage age (Col (4), same table). The subdued effect (compared to education) on the variable related to marriage is not surprising due to families being major decision-makers and strong cultural norms in the rural setup.

We test wives' years of education in Column (3) of Table 1. The coefficient is positive at 0.352 (7% increase over the mean) at 1% significance level, suggesting increasing education-level assortativity.

Column (5) of Table 1 depicts the impact of school expansion on inter-caste marriage. Our main coefficient of interest β shows the effect of an increase in primary schools on the marriage outcome of men exposed to school expansion during the age of 0-9 years, compared to men who did not experience any change in the number of primary schools. The effect of school expansion on inter-caste marriage is positive and statistically significant at 5% significance level (.028). One sd increase in the school expansion increases ICM by 5.67% of the sample mean.[14]

6.1 Robustness

As a first robustness check, we exclude specific social groups from the sample and examine whether our main coefficient on ICM remains robust. We find that our results remain qualitatively similar when we exclude the ST group from our sample or consider only

[14]The standard deviation of our main regressor is 0.26; the impact in the ICM for 1sd change is going to be 0.028*0.26=.00737; which is 5.67% of the mean (.00737/.13)

Hindus who are non-ST, as shown in Column (2) and Column (5) of Appendix Table B.II. However, our coefficient of interest loses significance if we exclude Muslims or take Hindus with ST as our sample (Columns (3) and (4) of Appendix Table B.II). This suggests interesting heterogeneity across castes, which we explore in the later parts.

The next robustness check shows that results are not driven by any possible selection of sample coming from inefficiency in the matching rate. In Appendix Table B.III, as we move from Column (1) to Column (5), we concentrate on a sample with an improved matching rate. For example, in Column (1), we consider a sample where we could match at least 50 percent of villages in the NFHS cluster with DISE Census data. Similarly, in Column (5), we narrow the sample to cases where we could match 90 percent or more of the villages. We find the coefficient to be quite stable (0.019-0.027) across all columns at 10% significance level, except when we restrict too much in Columns (4) and (5), leading to a drastic reduction in sample size.

6.2 Heterogeneity

6.2.1 Heterogeneity by intensity of school expansion

Table 3 examines heterogeneity by intensity of school expansion. We categorize the change in the number of primary schools in a cluster into four categories: less than 25 % change in number of primary schools, between 25-50 % change in number of primary schools, between 50-75 % change in number of primary schools, and more than 75 % change in number of primary schools. We observe that the impact of school expansion on the years of education and age of marriage for husbands and wives is most pronounced and statistically significant in the fourth category. However, we do not find any differential impact on ICM based on the intensity of school expansion.

6.2.2 Heterogeneity by Social groups

We examine the impact of school expansion on ICM for different social groups. In Table 4 and 5, we observe that the impact of school expansion is positive and statistically significant for men belonging to the General and OBC groups. Interestingly, while the husband's years of education no longer show a significant association with ICM, the wife's years of education are negatively associated with ICM among the General castes.[15] While the effect

[15]Once again, the stability of our coefficient of interest, even after including the years of education and age of marriage for both husbands and wives, suggests that these factors are not the potential channels of explanation.

of school expansion for SC is positive (but insignificant), for ST it is negative (insignificant), as shown in Table 6 and Table 7. This could be because ST groups are geographically isolated, and their integration with other communities is much lower compared to other caste groups.

Next, we interact men's caste group with our main regressor to check for any differential impact by caste groups. The coefficients are plotted in Figure 2 at 5% significance level. All the coefficients are significant at 5% level. The magnitude for ST is negative at 0.072, showing that primary school expansion had an overall negative impact on the ICM within ST. The coefficients are positive for all other three caste groups; SC - 0.093, OBC and General - 0.012.

7 Possible Mechanisms

In this section we explore some possible channels that might be driving our main results on ICM.

7.1 Is the effect driven by increasing years of education?

First, we test the most obvious channel of gains in years of education (due to the expansion of primary schools). First, we add the husbands' years of education as a control in our baseline specification and find that the main coefficient remains almost the same as before (Column (2) of Table 2; Column (1) shows the baseline results) at 5% significance level,[16] suggesting a lesser role of direct increase of education to be the main channel.

Next, we split the sample by the four completed levels of education of men — no education, primary, secondary, and higher — as captured during NFHS surveys. The coefficient of interest is positive and of similar magnitude (to the baseline specification) — for three (out of the four) groups. The coefficient for those who have studied till the primary level is negative. However, all the coefficients are statistically insignificant even at 10% confidence level (Refer Appendix Table B.IV). Once again, the lack of meaningful differential results suggests that gains in education are likely not the main channel.

Finally, we interact men's complete levels of education with our main regressor, to check for any differential impact. Estimated coefficients are plotted in Appendix Figure A.III,

[16] Notably, the coefficient on the husbands' years of education is negative at 1% significance level, suggesting that higher education among husbands is associated with lower rates of inter-caste marriage.

with "no education" as the base category. The coefficients on "primary", "secondary", and "higher" are close to zero, showing no differential impact of completed years of education.

7.2 Is the effect driven by the increasing age at marriage?

Increasing the age at marriage allows more time to choose one's spouse by oneself and possibly could offer less resistance from families in unconventional matchings, such as ICM (Ghimire and Axinn 2006). Hence, we test this mechanism explicitly.

First, we add the age at marriage at the husband (men) directly in our main specification and observe if there is any change in our main coefficient. Col(4) of Table 2 shows that adding this variable as a control doesn't change our main coefficient's magnitude or significance level. Adding the age at marriage of the wife (Col(5) of Table 2) also doesn't change the result. This exercise suggests that age at marriage may also not be the main channel in our sample.

7.3 Is the effect driven by "contact theory"?

Intergroup contact is recognised as one of the most effective means to reduce prejudice and discrimination (Allport, 1979). When members of majority groups engage with individuals from minority groups, it leads to better interaction and understanding, ultimately leading to a reduction in prejudice and discrimination against the minority groups.

The literature suggests that childhood, providing the formative years of development, plays a crucial role in shaping individuals' social preferences and choices. For example, increased exposure of children to other ethnicities during their early years is associated with a higher likelihood of forming interethnic friendships (Boucher et al., 2020).

We believe that with the expansion of primary schools, children have had the opportunity to come into contact with peers and teachers from different social groups in their early childhood, potentially leading to long-lasting impacts on their social preferences, behavior, and choices. We plan to explore the contact theory further as a potential explanation in this project.

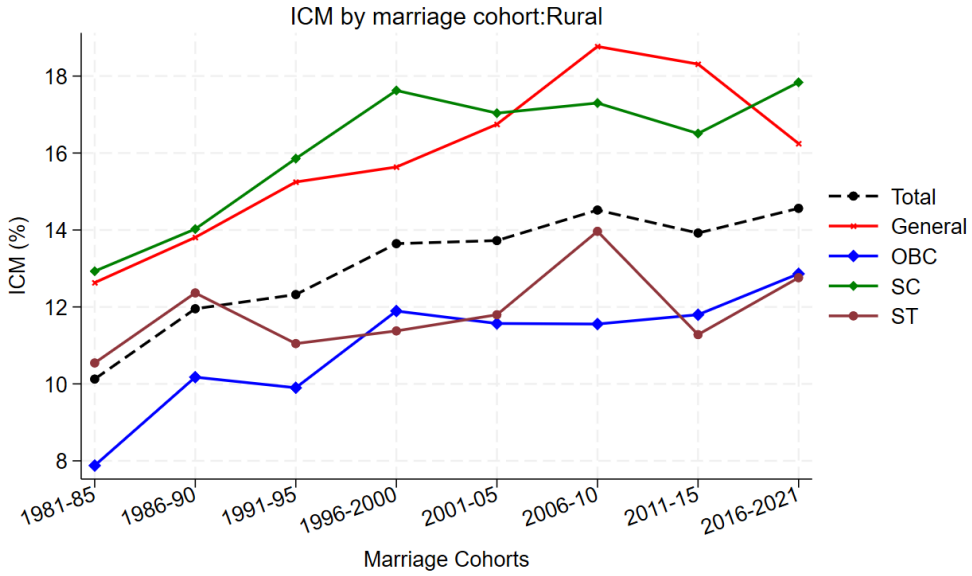
8 Conclusion

We identify the impact of school expansion on inter-caste marriages in rural India. We show that increasing years of education and increasing age at marriages are not the most

likely channels. We hypothesize that since contact with other castes (in schools) occurred at an early age, it may have developed a positive attitude towards different caste groups. However, we still have to explore this channel in more detail, which we are planning to present in the next draft.

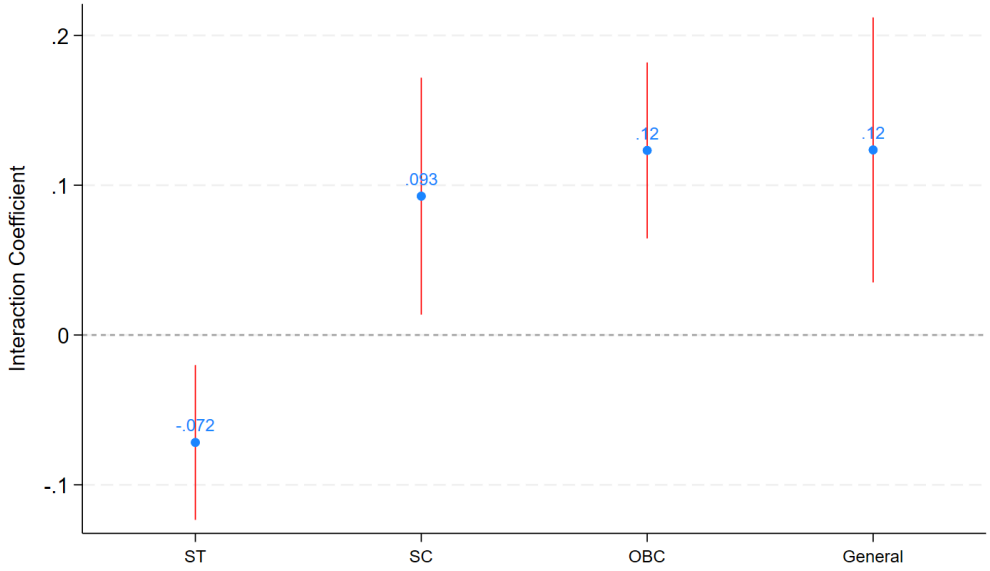
9 Figure

Figure 1: Evolution of ICM



Notes: The figure plots the evolution of inter-caste marriages by birth cohort in rural India from the 1980s till now. It also shows the ICM within each caste category (using the husband's caste). Calculations using NFHS combined sample without using survey design weights.

Figure 2: Differential Impact of Men's Caste Group



Notes: The figure plot coefficient on the dummy of caste group after the ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever - following Equation 1. The dependent variable is a dummy of inter-caste marriage.

10 Table

Table 1: Impact of Schools on Years of Education, Age of Marriage and Inter-Caste Marriage

VARIABLES	(1) h_educ_yr	(2) age_marr_H	(3) w_educ_yr	(4) age_marr_W	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.601***	0.541***	0.352***	0.143	0.028**
	(0.159)	(0.179)	(0.132)	(0.128)	(0.014)
# of Schools/Vill pre-birth	0.439**	-0.390	-0.108	-0.534***	0.003
	(0.198)	(0.238)	(0.173)	(0.182)	(0.012)
Observations	54,789	54,233	54,789	54,233	51,150
R-squared	0.476	0.399	0.563	0.362	0.316
MeanDepVar	7.22	22.93	5.20	18.54	0.13
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
CLUSTER	statedist	statedist	statedist	statedist	statedist
NUM_clusters	528	528	528	528	528

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is the years of education of husbands (or men) in Column (1); age at marriage of husbands in Column (2); years of education of wives in Column (3); age at marriage of wives in Column (4); and, dummy of inter-caste marriage in Column (5). The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 2: Impact of Schools on Inter-Caste Marriage

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.028**	0.029**	0.029**	0.027*	0.027*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
# of Schools/Vill pre-birth	0.003	0.004	0.004	0.004	0.004
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
h_educ_yr		-0.002***	-0.002***	-0.002***	-0.002***
		(0.001)	(0.001)	(0.001)	(0.001)
w_educ_yr			-0.000	-0.000	-0.000
			(0.001)	(0.001)	(0.001)
age_marr_H				0.001	0.001
				(0.001)	(0.001)
age_marr_W					-0.001
					(0.001)
Observations	51,150	51,150	51,150	50,613	50,613
R-squared	0.316	0.317	0.317	0.318	0.318
MeanDepVar	0.13	0.13	0.13	0.13	0.13
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	528	528	528	528	528

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) is the baseline regression. Column (2) adds husbands' years of education as controls. Column (3) adds wives' years of education; Column (4) and (5) further add age at marriage for husband and wife. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 3: Impact of schools expansion intensity on years of education, age at marriage and ICM

VARIABLES	(1) h_educ_yr	(2) age_marr_H	(3) w_educ_yr	(4) age_marr_W	(5) icm
Δ in Schools/Vill b.n. 0-9 age (2nd quartile)	0.025	-0.044	-0.016	0.086	-0.003
	(0.090)	(0.095)	(0.070)	(0.070)	(0.007)
Δ in Schools/Vill b.n. 0-9 age (3rd quartile)	0.152	0.210**	0.137	0.171**	0.002
	(0.102)	(0.101)	(0.085)	(0.083)	(0.008)
Δ in Schools/Vill b.n. 0-9 age (4th quartile)	0.399***	0.413***	0.353***	0.277***	0.011
	(0.127)	(0.131)	(0.101)	(0.094)	(0.010)
# of Schools/Vill pre-birth	0.474**	-0.316	-0.023	-0.446**	0.004
	(0.198)	(0.263)	(0.163)	(0.193)	(0.011)
Observations	54,789	54,233	54,789	54,233	51,150
R-squared	0.476	0.399	0.563	0.362	0.316
MeanDepVar	7.22	22.93	5.20	18.54	0.13
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	528	528	528	528	528

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is the years of education of husbands (or men) in Column (1); age at marriage of husbands in Column (2); years of education of wives in Column (3); age at marriage of wives in Column (4); and, dummy of inter-caste marriage in Column (5). The main explanatory variable here is the change in schools (per village) in the first 9 years of age of husbands - split into 4 quartiles, with the first quartile as the base category. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 4: Impact of Schools on Inter-Caste Marriage for General caste

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.065*	0.070*	0.072*	0.073*	0.073*
	(0.039)	(0.040)	(0.039)	(0.040)	(0.040)
# of Schools/Vill pre-birth	-0.007	-0.003	-0.005	-0.004	-0.007
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
h_educ_yr		-0.004*	-0.002	-0.002	-0.002
		(0.002)	(0.002)	(0.002)	(0.002)
w_educ_yr			-0.005***	-0.004**	-0.004**
			(0.002)	(0.002)	(0.002)
age_marr_H				-0.001	0.001
				(0.002)	(0.002)
age_marr_W					-0.004
					(0.002)
Observations	7,703	7,703	7,703	7,655	7,655
R-squared	0.588	0.589	0.590	0.588	0.588
MeanDepVar	0.15	0.15	0.15	0.15	0.15
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	442	442	442	441	441
Caste	General	General	General	General	General

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Further, the sample is restricted to the General caste population. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) is the baseline regression. Column (2) adds husbands' years of education as controls. Column (3) adds wives' years of education; Column (4) and (5) further add age at marriage for husband and wife. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 5: Impact of Schools on Inter-Caste Marriage for OBC

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.036**	0.036**	0.036**	0.036**	0.036**
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)
# of Schools/Vill pre-birth	-0.020	-0.019	-0.019	-0.021	-0.020
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
h_educ_yr		-0.001	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
w_educ_yr			0.000	0.000	0.000
			(0.001)	(0.001)	(0.001)
age_marr_H				-0.000	-0.001
				(0.001)	(0.001)
age_marr_W					0.001
					(0.002)
Observations	20,603	20,603	20,603	20,379	20,379
R-squared	0.411	0.412	0.412	0.412	0.412
MeanDepVar	0.11	0.11	0.11	0.11	0.11
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	513	513	513	513	513
Caste	OBC	OBC	OBC	OBC	OBC

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Further, the sample is restricted to the Other Backward Class (OBC) group population. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) is the baseline regression. Column (2) adds husbands' years of education as controls. Column (3) adds wives' years of education; Column (4) and (5) further add age at marriage for husband and wife. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 6: Impact of Schools on Inter-Caste Marriage for SC

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.042	0.044	0.044	0.039	0.039
	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)
# of Schools/Vill pre-birth	0.014	0.015	0.014	0.014	0.014
	(0.040)	(0.040)	(0.040)	(0.041)	(0.041)
h_educ_yr		-0.002	-0.001	-0.002	-0.002
		(0.002)	(0.002)	(0.002)	(0.002)
w_educ_yr			-0.001	-0.001	-0.001
			(0.002)	(0.002)	(0.002)
age_marr_H				-0.000	-0.001
				(0.001)	(0.002)
age_marr_W					0.001
					(0.002)
Observations	8,792	8,792	8,792	8,687	8,687
R-squared	0.525	0.525	0.525	0.524	0.524
MeanDepVar	0.16	0.16	0.16	0.16	0.16
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	476	476	476	476	476
Caste	SC	SC	SC	SC	SC

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Further, the sample is restricted to the Scheduled Caste population. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) is the baseline regression. Column (2) adds husbands' years of education as controls. Column (3) adds wives' years of education; Column (4) and (5) further add age at marriage for husband and wife. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table 7: Impact of Schools on Inter-Caste Marriage for ST

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	-0.022	-0.023	-0.023	-0.025	-0.025
	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)
# of Schools/Vill pre-birth	0.013	0.012	0.013	0.013	0.013
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
h_educ_yr		-0.003**	-0.004***	-0.003**	-0.003**
		(0.001)	(0.001)	(0.001)	(0.001)
w_educ_yr			0.001	0.000	0.000
			(0.001)	(0.001)	(0.001)
age_marr_H				0.001	0.001
				(0.001)	(0.002)
age_marr_W					-0.000
					(0.002)
Observations	8,267	8,267	8,267	8,111	8,111
R-squared	0.583	0.584	0.584	0.584	0.584
MeanDepVar	0.09	0.09	0.09	0.09	0.09
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	290	290	290	289	289
Caste	ST	ST	ST	ST	ST

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Further, the sample is restricted to the Scheduled Tribe population. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) is the baseline regression. Column (2) adds husbands' years of education as controls. Column (3) adds wives' years of education; Column (4) and (5) further add age at marriage for husband and wife. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

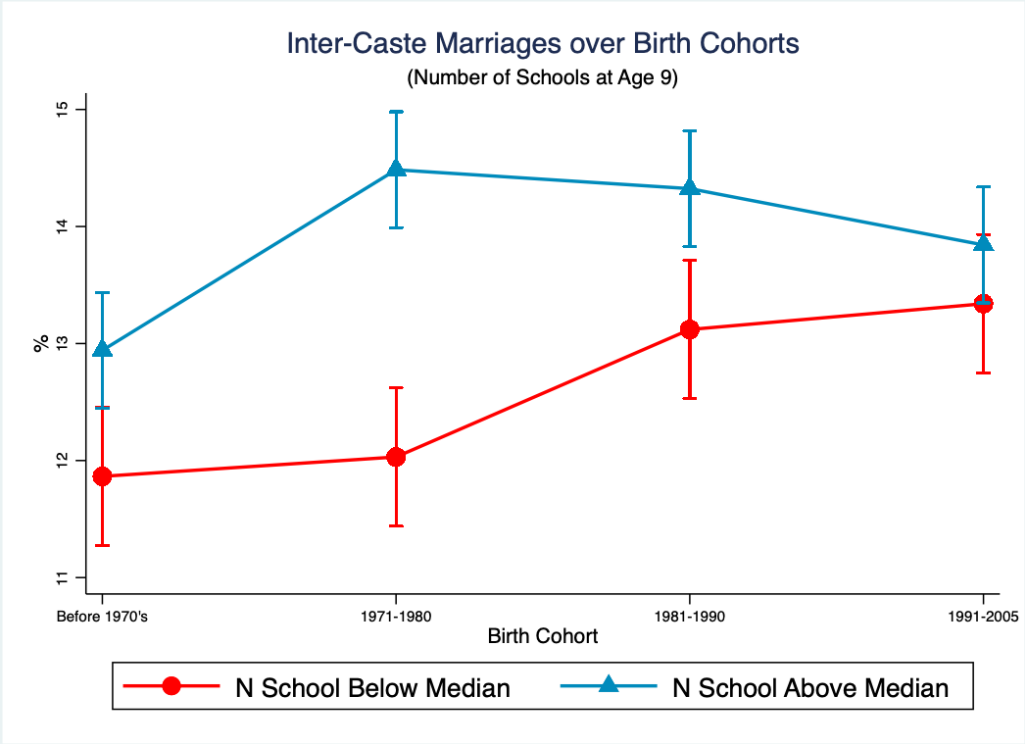
A Appendix Figures

Figure A.I: ICM by husband and wife caste group

Wife's Caste	Husband's Caste			
	Scheduled Tribe (ST)	Scheduled Caste (SC)	Other Background Class (OBC)	General Caste (Gen)
Scheduled Tribe (ST)	**87.9 %** (10,584)	**5.8 %** (855)	**1.9 %** (551)	**1.5 %** (178)
Scheduled Caste (SC)	**7.0 %** (836)	**83.3 %** (12,243)	**4.4 %** (1,314)	**3.0 %** (357)
Other Background Class (OBC)	**4.0 %** (476)	**8.8 %** (1,294)	**88.8 %** (26,471)	**12.5 %** (1,503)
General Caste (Gen)	**1.2 %** (141)	**2.1 %** (302)	**5.0 %** (1,481)	**83.1 %** (9,997)

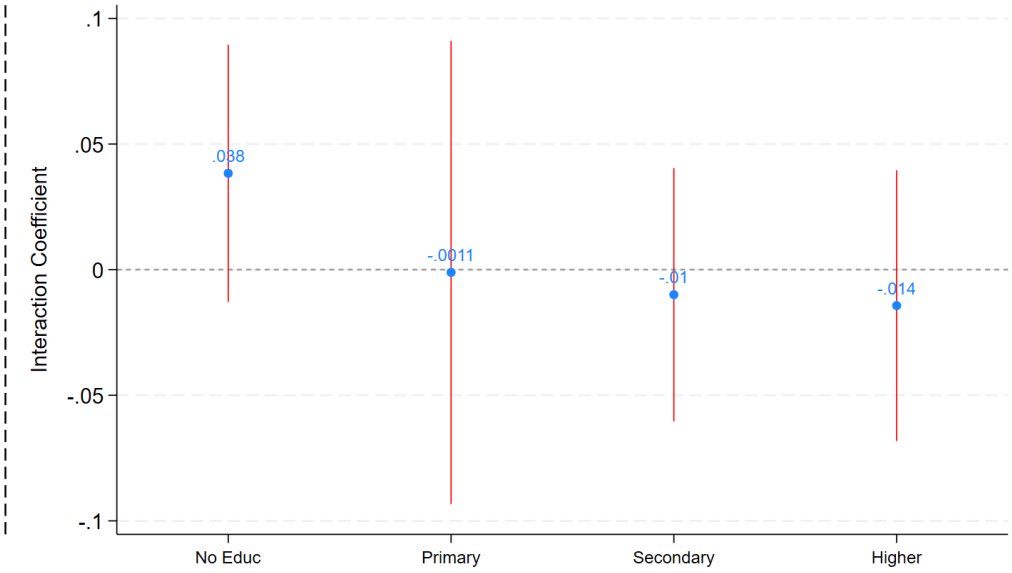
Notes: The table shows the rate of ICM by husband and wife caste categories. For e.g., if we look at the first row, then 87% of the ST marriages occur within the same caste, and out of 13% ICM, 7% is with SC, 4.5% with OBC, and 1.5% with the General caste group.

Figure A.II: ICM by husband and wife caste group



Notes: The figure shows the rate of ICM by two categories- below and above the median, where the median is defined by the total number of schools at the age of 9. It shows that the rate of ICM is higher for all the birth cohort of men who where exposed to more number of primary schools in their school-going age.

Figure A.III: Differential Impact of Men's Completed Level of Education



Notes: The figure plot coefficient on the dummy of men's completed education level after the ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever - following Equation 1. The dependent variable is a dummy of inter-caste marriage.

B Appendix Tables

Table B.I: Data Preparation: Matching Rate

States	NFHS (5km buffer zone)	DISE-CENSUS	NFHS & DISE-CENSUS	Unmatched Villages In NFHS but not in DISE-CENSUS	Matching Rate
Andhra Pradesh	6,881	24,334	6,267	614	91%
Assam	19,807	15,532	12966	6,841	65%
Bihar	40,026	29,062	26,289	13,737	66%
Chhattisgarh	12,298	16407	10,423	1,875	85%
Goa	268	205	171	97	64%
Gujarat	11,558	16,697	10,718	840	93%
Haryana	6,140	5,904	5,308	832	86%
Himachal Pradesh	10,552	6,949	4,007	6,545	38%
Jammu & Kashmir	5,419	4,866	4072	1,347	75%
Jharkhand	27,531	22,023	18,690	8,841	68%
Karnataka	17,370	22,577	13,539	3,831	78%
Kerala	923	730	663	260	72%
Madhya Pradesh	36,452	43,052	29,030	7,422	80%
Maharashtra	19,957	36,508	16,964	2,993	85%
Manipur	418	1572	309	109	74%
Odisha	35,667	34297	24,607	11,060	69%
Puducherry	86	82	78	8	91%
Punjab	10,048	9,580	7,738	2,310	77%
Rajasthan	25,124	29,721	17,178	7,946	68%
Sikkim	450	155	155	295	34%
Tamil Nadu	10,642	11,651	7,908	2,734	74%
Tripura	843	775	748	95	89%
Uttar Pradesh	89,252	63,374	53,441	35,811	60%
Uttarakhand	13,877	8,873	7,334	6,543	53%
Total	***401,589***	***404,926***	***278,603***	*122,986*	***69%***

Notes: The table presents the matching rate by state. Column (1) is the name of the state. Column (2) is the number of villages falling within all the 5km buffer zones created. Column (3) is the number of villages in each state coming from the DISE-Census fuzzy matching exercise. Col (4) is the number of villages after matching Col (2) and Col (3). Col (5) is the number of villages unable to be matched. Col (6) is the matching rate. Overall we are able to map 70% of the villages from Col (2).

Table B.II: Robustness check: Impact of schools on ICM with different subpopulation

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.013	0.040***	0.016	0.017	0.035*
	(0.020)	(0.016)	(0.016)	(0.017)	(0.020)
# of Schools/Vill pre-birth	-0.007	0.001	0.013	0.012	0.015
	(0.017)	(0.014)	(0.014)	(0.015)	(0.016)
Observations	46,857	41,761	46,889	43,491	35,583
R-squared	0.329	0.332	0.327	0.330	0.349
MeanDepVar	0.13	0.14	0.13	0.13	0.14
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	528	525	520	515	513
Condition	W_Abv_22	Excl_ST	Excl_Muslim	Hindu_Only	Hindu_non_ST

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) keeps the wife above 22 years; Column (2) excludes Scheduled Tribe from the sample; Column (3) excludes Muslims; Column (4) keeps hindu only subsample (and no religion dummy controls); and Column (5) is with the sub-sample who are Hindu and not Scheduled Tribe. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table B.III: Robustness checks: Impact of schools on ICM with different level of matching

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm	(5) icm
Δ in Schools/Vill b.n. 0-9 age	0.028*	0.025*	0.027*	0.026	0.019
	(0.015)	(0.015)	(0.016)	(0.017)	(0.021)
# of Schools/Vill pre-birth	0.005	0.002	0.000	-0.001	0.003
	(0.013)	(0.013)	(0.014)	(0.015)	(0.015)
Observations	43,992	40,019	32,753	24,056	12,320
R-squared	0.316	0.318	0.320	0.318	0.306
MeanDepVar	0.14	0.14	0.14	0.14	0.14
DHS FE	Yes	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist	statedist
NUM_clusters	524	514	501	476	373
Condition	$> 50\%$	$> 60\%$	$> 70\%$	$> 80\%$	$> 90\%$

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) keeps where matching rate was at least 50% (i.e. we were able to find the school-level DISE information for more than 50% of the NFHS 5km cluster); Column (2) with matching rate above 60%; Column (3) with matching rate above 70%; Column (4) with matching rate above 80%; and Column (5) with matching rate above 90%. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

Table B.IV: Differential impact on ICM by education level

VARIABLES	(1) icm	(2) icm	(3) icm	(4) icm
Δ in Schools/Vill b.n. 0-9 age	0.028	-0.049	0.023	0.030
	(0.033)	(0.038)	(0.018)	(0.053)
# of Schools/Vill pre-birth	0.084*	-0.014	0.012	0.123
	(0.048)	(0.053)	(0.018)	(0.083)
Observations	7,116	5,665	26,031	2,760
R-squared	0.489	0.514	0.402	0.525
MeanDepVar	0.14	0.13	0.13	0.12
DHS FE	Yes	Yes	Yes	Yes
Birth Year FE	yes	yes	yes	yes
Lived Always	yes	yes	yes	yes
Cluster	statedist	statedist	statedist	statedist
NUM_clusters	443	478	528	407
Condition	no education	primary	secondary	higher

Notes: The table reports ordinary least square estimations based on the NFHS-4 and 5, married couple-level sample, living in rural areas, husbands above 22 years old, and who have lived in the same place forever. Robust standard errors in parentheses are clustered at the district level. The dependent variable is a dummy of inter-caste marriage. The main explanatory variable is the change in schools (per village in the first 9 years of age of husbands. We also control for the number of schools (per village) one year before birth, household current wealth, and dummy for Hindu and Muslim religions. Column (1) retains men who have no education; Column (2) is with men who have completed at least primary level education; Column (3) is with men who have completed at least secondary level education; and Column (4) is with men who have completed at least higher level education. All estimations include NFHS 5km buffer, birth year of men, and NFHS round fixed effects.

References

Allendorf, K. (2013). Going Nuclear? Family Structure and Young Women's Health in India, 1992-2006. *Demography*, 50(3):853–880.

Allport, G. W. (1979). *The Nature of Prejudice*. Basic Books, Inc.

André, P. and Dupraz, Y. (2023). Education and Polygamy: Evidence From Cameroon. *Journal of Development Economics*, 162(February):103068.

Anukriti, S., Kwon, S., and Prakash, N. (2022). Saving for dowry: Evidence from rural india. *Journal of Development Economics*, 154:102750.

Banerjee, A., Duflo, E., Ghatak, M., and Lafortune, J. (2013). Marry for What? Caste and Mate Selection in Modern India. *American Economic Journal: Microeconomics*, 5(2):33–72.

Béteille, A. (1991). The Reproduction of Inequality: Occupation, Caste and Family. *Contributions to Indian Sociology*, 25(1):3–28.

Bhagat, R. (2016). The Practice of Early Marriages among Females in India: Persistence and Change.

Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, 103(5):1862–1891.

Bharti, N. K. and Roy, S. (2023). The early origins of judicial stringency in bail decisions: Evidence from early childhood exposure to hindu-muslim riots in india. *Journal of Public Economics*, 221:104846.

Blau, P. M., Beeker, C., and Fitzpatrick, K. M. (1984). Intersecting Social Affiliations and Intermarriage. *Social Forces*, 62(3):585–606.

Boucher, V., Tumen, S., Vlassopoulos, M., Wahba, J., and Zenou, Y. (2020). Ethnic mixing in early childhood: Evidence from a randomized field experiment and a structural model.

Caldwell, J. C., Reddy, P. H., and Caldwell, P. (1983). The Causes of Marriage Change in South India. *Population Studies*, 37(3):343–361.

Chiplunkar, G. and Weaver, J. (2023). Marriage markets and the rise of dowry in india. *Journal of Development Economics*, 164:103115.

Chou, S. Y., Liu, J. T., Grossman, M., and Joyce, T. (2010). Parental Education and Child Health: Evidence From a Natural Experiment in Taiwan. *American Economic Journal: Applied Economics*, 2(1):33–61.

chun Lin, Y., yih Chen, A. S., and chen Song, Y. (2012). Does Your Intelligence Help to Survive in a Foreign Jungle? The Effects of Cultural Intelligence and Emotional Intelligence on Cross-Cultural Adjustment. *International Journal of Intercultural Relations*, 36(4):541–552.

Corwin, L. A. (1977). Caste, Class and the Love-Marriage: Social Change in India. *Journal of Marriage and the Family*, 39(4):823.

Davis, K. (1941). Intermarriage in Caste Societies. *American Anthropologist*, 43(3):376–395.

Deary, I. J., Batty, G. D., and Gale, C. R. (2008). Bright Children Become Enlightened Adults. *Psychological Science*, 19(1):1–6.

Dommaraju, P. (2010). Instruction Des Femmes et Évolution du Mariage en Inde. *Population*, Vol. 64(4):757–775.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91(4):795–813.

Fuller, C. J. and Narasimhan, H. (2008). Companionate Marriage in India: The Changing Marriage System in a Middle-Class Brahman Subcaste. *Journal of the Royal Anthropological Institute*, 14(4):736–754.

Furtado, D. (2012). Human Capital and Interethnic Marriage Decisions. *Economic Inquiry*, 50(1):82–93.

Garrouste, M. and Zaiem, M. (2020). School supply Constraints in Track Choices: A French Study Using High School Openings. *Economics of Education Review*, 78(August):102041.

Golebiowska, E. A. (2007). The Contours and Etiology of Whites' Attitudes Toward Black-White Interracial Marriage. *Journal of Black Studies*, 38(2):268–287.

Goli, S., Singh, D., and Sekher, T. V. (2013). Exploring the Myth of Mixed Marriages in India: Evidence from a Nation-Wide Survey. *Journal of Comparative Family Studies*, 44(2):193–206.

Gundemeda, N. (2020). Caste in Twenty First Century India: Sociological Reflections on University Students' Perceptions in South India. *Asian and African Studies*, 29(1):89–111.

Hathi, P. (2019). Can More Education Reduce Opposition to Intermarriage? Comparing India and the US.

Hodson, G. and Busseri, M. A. (2012). Bright Minds and Dark Attitudes: Lower Cognitive Ability Predicts Greater Prejudice Through Right-Wing Ideology and Low Intergroup Contact. *Psychological Science*, 23(2):187–195.

Kannan, C. T. (1963). *Intercaste and Inter-Community Marriages in India*. Allied Publishers.

Khanna, G. (2023). Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India. *Journal of Political Economy*, 131(2):549–591.

Mander, H. and Prasad, M. (2014). India Exclusion Report. Technical report, Books for Change, Bangalore, India.

Medhe, R. (2019). Role of Education in Inter-Caste Marriages of Women from Scheduled Castes in India. *Journal of Geographical Studies*, 2(2):79–86.

Munshi, K. (2017). Caste and the Indian Economy.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(01):46–98.

Munshi, K. and Rosenzweig, M. R. (2009). Why is Mobility in India so Low? Social Insurance, Inequality, and Growth.

Ray, T., Roy Chaudhuri, A., and Sahai, K. (2020). Whose Education Matters? An Analysis of Inter Caste Marriages in India. *Journal of Economic Behavior and Organization*, 176:619–633.

Ross, A. D. (1961). *The Hindu Family in Its Urban Setting*. University of Toronto Press., Toronto, CA.

Roth, J. and Peck, R. F. (1951). Social Class and Social Mobility Factors Related to Marital Adjustment. *American Sociological Review*, 16(4):478–487.

Sarkar, K. (2022). Can Status Exchanges Explain Educational Hypogamy in India? *Demographic Research*, 46(May):809–847.

Schoon, I., Cheng, H., Gale, C. R., Batty, G. D., and Deary, I. J. (2010). Social Status, Cognitive Ability, and Educational Attainment as Predictors of Liberal Social Attitudes and Political Trust. *Intelligence*, 38(1):144–150.

Schwartz, C. R., Zeng, Z., and Xie, Y. (2016). Marrying Up by Marrying Down: Status Exchange Between Social Origin and Education in the United States. *Sociological Science*, 3:1003–1027.

Sharma, A. (2019). Influence of Parental Education on Nuptial Behavior: An Analysis of Hindu Community in India. *Journal of Population and Social Studies*, 27(2):168–180.

Singh, M., Shekhar, C., and Shri, N. (2023). Patterns in Age at First Marriage and Its Determinants in India: A Historical Perspective of Last 30 Years (1992–2021). *SSM - Population Health*, 22(February):101363.

Song, M. (2009). Is Intermarriage a Good Indicator of Integration? *Journal of Ethnic and Migration Studies*, 35(2):331–348.

Thornton, A. (2001). The Developmental Paradigm, Reading History Sideways, and Family Change. *Demography*, 38(4):449–465.

Uberoi, P. (2006). *Freedom and Destiny: Gender, Family, and Popular Culture in India*. Oxford University Press, Oxford, UK.

Vaid, D. (2014). Caste in Contemporary India: Flexibility and Persistence. *Annual Review of Sociology*, 40:391–410.

Wodtke, G. T. (2012). The Impact of Education on Intergroup Attitudes: A Multiracial Analysis. *Social Psychology Quarterly*, 75(1):80–106.