Discussion Papers in Economics

Whose Education Matters? An Analysis Of Inter Caste Marriages In India.

Tridip Ray, Arka Roy Chaudhuri and Komal Sahai

September 2017

Discussion Paper 17-05

Indian Statistical Institute, Delhi
Economics and Planning Unit
7, S. J. S. Sansanwal Marg, New Delhi 110016, India

WHOSE EDUCATION MATTERS?
AN ANALYSIS OF INTER CASTE MARRIAGES IN INDIA*

Tridip Ray† Arka Roy Chaudhuri‡ Komal Sahai±

September 18, 2017

Abstract

Endogamy or intra-caste marriage is one of the most resilient of all the caste based practices in India. Even in 2011, the rate of inter caste marriages in India was as low as 5.82%. In this paper we explore whether education has any relationship with this age-old practice of marrying within one's own caste. Using a nationally representative data set, the Indian Human Development Survey, we find that, in sharp contrast with the findings in the existing literature on out-marriages in the Western countries, education levels of the spouses themselves do not have any association with the likelihood of their own marriage being an inter caste one. However, couples with a more educated mother of the husband have a significantly higher probability of being in an inter caste marriage. Increase in years of education of the husband's mother by 10 years would lead to an increase in the probability of inter caste marriage by 1.86 percentage points which is equivalent to approximately 36 percent of the sample mean. Our analysis highlights the importance of recognizing the institution of *arranged marriages* in any analysis of Indian marriage markets.

KEYWORDS: Inter caste marriages; education; arranged marriage institution; caste; India.

*We are grateful to Farzana Afridi, Diane Coffey, Sabyasachi Das, Ashwini Deshpande, Mukesh Eswaran, Abhiroop Mukhopadhyay, Bharat Ramaswami, Ranjan Ray, E. Somanathan, Dean Spears and seminar participants at ISI Delhi for their valuable comments and suggestions. We especially thank Sonalde Desai for her suggestions and for clearing several doubts regarding the IHDS data set. All remaining errors are ours.

†Indian Statistical Institute, New Delhi; tridip@isid.ac.in

‡Indian Statistical Institute, New Delhi; arka.rc@isid.ac.in

±Indian Statistical Institute, New Delhi; ksahai14r@isid.ac.in

1 Introduction

Ethnic endogamy as a practice to entrench clan, community or tribal boundaries has been around for centuries (Bidner and Eswaran 2015). The practice of endogamy or its obverse i.e exogamy has been studied in different contexts from racial intermarriage in the United States (Fryer 2007) to the practice of marriage within various ethnic groups in Africa (Conte 1979). Social scientists have analyzed various aspects of exogamy: the factors predicting intermarriage such as education (Furtado 2012; Gullickson 2006; Qian 1997), group size and geographic distribution (Blau et al. 1982; Hwang et al. 1994; Stevens and Swicegood 1987), cultural norms (Kalmijn and Van Tubergen 2010); and the consequences of intermarriage such as cultural and spatial assimilation (Ellis et al. 2006; Iceland and Nelson 2010; Luke and Luke 1998; Saenz et al. 1995), economic assimilation (Meng and Gregory 2005) and marital dissolution rates (Zhang and Van Hook 2009).

In the Indian context, endogamy has been one of the central features of the institution of caste.[1] Indian castes are largely endogamous groups with strict social norms regarding endogamy often enforced by imposing punishments such as social ostracism (Bidner and Eswaran 2015; Chowdhry 1997; Kaur 2010). Endogamy is to caste what caste is to the Indian society. This practice of within caste endogamy is not only central to the institution of caste, but is also one of the most resilient caste based practices till date. The rate of inter caste marriages, even as recent as in 2011, was as low as 5.82% and there has been no upward time trend over the past four decades. In this paper we study the relationship of caste endogamy with education. We explore whether education has any statistically significant association with this

[1] A huge body of literature has been developed to understand the origin, nature and contemporary aspects of the caste system in India (see, for example, Srinivas 1962; Beteille 1971; Dumont 1980.

age-old practice of marrying within one's own caste taking into account the nature of the Indian marriage market where marriages arranged by parents and close relatives is largely the norm.

There exists a substantial body of literature on out-marriages and a large part of this literature focuses on the relationship between out-marriages and the education of individuals. The evidence has been found to vary by the ethnic group under study. Qian (1997) looks at interracial marriages among whites, African Americans, Hispanics and Asian Americans in the USA in 1980 and 1990 and reports that the likelihood of an interracial marriage is positively related to the couple's educational attainment. Fryer (2007) finds that whites with some college education or more have shown a small increase in interracial marriage over the decades from 1940 to 2000 in the USA. While Qian and Lichter (2001) find the relationship to be positive for Latinos, Hwang et al. (1995) find that Asian women with lower levels of education are more likely to out-marry racially. Gullickson (2006), on the other hand, does not find any consistent relationship between the educational attainment of whites and the likelihood of interracial marriages.

To the best of our knowledge there has been no systematic attempt to understand whether education has any association with the age-old practice of inter caste marriages in India. We explore this in our research. But, at the outset, we recognize that we have to pay due attention to the institution of marriage markets in India. Marriage markets in India work very differently compared to the Western countries (Banerjee et al. 2013). A majority of marriages are arranged by the parents, and the spouses barely know each other before marriage. We find this pattern to be very prominent in our data set (second round of the Indian Human Development Survey, IHDS-II) too. 73% of marriages in our sample were reported to have been arranged

by parents and almost 70% women said that they met their husbands only on the day of their wedding/*gauna*[2]. Even among the women who reported that they chose their husbands themselves, 34% of them did not know their husbands before their wedding/*gauna* day. This pattern, quite surprisingly, holds for the inter caste marriages as well: close to 63% of those who said they were in an inter caste marriage reported their marriages to be arranged by parents. In fact recent studies using the IHDS have shown that apart from the fact that "arranged marriages" are the norm in India, the movement over time has not been towards "Western-style marriage, in which young people choose their own spouses" (Allendorf and Pandian 2016)[3]. The shift is rather towards increased say of women within the purview of arranged marriages (Allendorf and Pandian 2016; Banerji et al. 2013).

Recognition of this "arranged marriages" institution in the Indian marriage markets strongly suggests that any analysis of marriages in India must consider parental level attributes along with individual level ones. And the relationship between inter caste marriages and education should be no exception. To establish our point, we first explore whether education levels of the spouses themselves have any predictive power on the likelihood of inter caste marriages. We find that, contrary to the findings in the existing literature on out-marriages in the Western countries, especially in the USA, the education levels of the individuals themselves do not have any association with the probability of inter caste marriages. The result is very robust to the inclusion of a whole range of controls and fixed effects, and to variations in the sample.

[2] *Gauna* is a ceremony conducted after several years of a child marriage when the bride moves from her natal home to her husband's family.

[3] The term "arranged marriage" is used to refer to a marriage where parents or other relatives play the main role in selecting a spouse for their offspring, often keeping social attributes like caste and economic status of the family in view (Banerji et al. 2013).

As noted earlier, researchers have obtained mixed results on the relationship between education and intermarriage (Gullickson 2006; Hwang et al. 1995; Qian 1997; Qian and Lichter 2001). Furtado (2012) tries to address these seemingly contradictory results. She identifies two effects working in opposite directions. The first is the positive 'cultural adaptability effect' through which education broadens the mindset and makes migrant ethnicities more aware of and adaptable to the culture of the natives. The second effect, the 'assortative matching effect', however, may work in either direction. In a group with average education level below the average education level of the relevant population, a more educated individual will marry out and education will have a positive effect on exogamy for that group. The opposite will hold for a group with above average education level – a more educated individual in this group will marry within the group. Since the net effect can go in either direction, one may observe a positive, a negative or no relationship between education and exogamy depending on a particular group's characteristics. To ensure that the statistical insignificance of individuals' own education coefficients in our analysis is not being driven by the two opposing channels identified by Furtado (2012), we carry out another set of robustness checks adopting the methodology suggested by Furtado (2012) to the Indian context. Our results reaffirm our original findings: none of the opposing channels have any statistically significant association with the probability of inter caste marriages in India.

Thus the first part of our empirical analysis confirms our belief that in the presence of the institution of arranged marriages, individuals themselves have a very little role to play in their marriage decisions. To complete our investigations, we next explore whether any of the parental attributes (education, in our context) affect the likelihood of an inter caste marriage. So we add the education levels of the

parents of both the spouses (along with the education levels of the spouses themselves) to our set of explanatory variables. Here we find that the level of education of the husband's mother has a positive association with the likelihood of an inter caste marriage. Increase in years of education of the husband's mother by 10 years would lead to an increase in the probability of inter caste marriage by 1.86 percentage points which is equivalent to approximately 36 percent of the sample mean. The result is very robust to variations in the sample and to the addition of a number of controls as well as fixed effects. However, this part of the result is nuanced in the sense that among the parents on both sides, only the education of the husband's mother has any predictive power on the likelihood of inter caste marriage. We posit some potential channels and provide anecdotal evidence and theoretical arguments from the existing literature.

The rest of the paper is organized as follows. In section 2, we describe the related literature and locate our paper within the literature. Section 3 describes the data. The descriptive analysis in Section 4 prepares the contextual background and provides the descriptive statistics. Section 5 details the regression analysis specifying the empirical strategy and discussing the main results and robustness checks. Section 6 gives a brief discussion of the possible channels behind the results and section 7 concludes.

2 Related Literature

At a broad level, we contribute to the huge literature on exogamy (interracial and inter ethnic marriages), primarily in the United States, where the basic premise is that inter group marriage is the key indicator of assimilation between various ethnic and racial groups. Fryer (2007) reports that even in the year 2000 interracial

marriages account for only around 1 percent of white marriages, 5 percent of black marriages, and 14 percent of Asian marriages. Wong (2003) reports that only 5.5% of black males married white females in 1990, attributing it mainly to the existing "mating taboo" across racial lines. Batson et al. (2006) look at the 2000 US Census data and find that native born African Americans are more likely to marry whites, as compared to Black ethnics from other geographic origins, in spite of their lower socioeconomic status. Using the same data set, Qian et al. (2012) find that ethnic endogamy is strong and marital or cohabiting unions with whites differs significantly by group characteristics like native origins and educational attainment. Another strand of literature looks at the characteristics of an inter ethnic and/or an interracial union both in the USA and elsewhere. Fu (2001) finds that in the USA marriages of blacks and Mexican Americans with whites suggests a racial hierarchy where whites are considered superior, putting the minority counterparts of the marriage market at a disadvantage. Gullickson and Torche (2014) find similar evidence of status exchange across racial lines even in Brazil where the fluidity of racial boundaries is greater. Our paper aims to throw more light on the existing trends on out-marriages for a developing country like India. We look at the relationship between education and inter caste marriages paying due attention to the institutional context in the Indian marriage market.

Our paper is closely related to the literature on exogamy in South Asia. However, such studies have been relatively scarce and primarily based on localized samples. Allendorf and Thornton (2015) conduct a longitudinal survey in Chitwan valley, Nepal, and provide the first individual-level test of the "Developmental Idealism" theory in sociology[4]. They report that individuals who showed greater belief in

[4]Developmental Idealism is a collection of beliefs and values which relate family behaviours to socioeconomic development. Certain beliefs and values, like choosing one's own spouse and living

Developmental Idealism in 2008 were also more likely to implement its values in 2012, like choosing their own spouse, even from a different caste. Banerjee et al. (2013) study a section of middle-class population in West Bengal, India, and find them to have a strong preference for within caste marriages and willing to pay a high price in terms of other attributes like education for a same caste match. Dugar et al. (2012) study the responses of higher caste females to lower caste males in matrimonial advertisements of potential grooms which they place in newspapers and systematically vary their caste and income. They find that discrimination exists, but it decreases with an increase in the income of lower status males. Ahuja and Ostermann (2016) find a marriage premium to attributes like income and class, along with caste in their survey of 1070 females from upper castes and Scheduled castes from the three Indian states of Uttar Pradesh, Maharashtra and Tamil Nadu. Interest in intermarrriage increases with income among SCs while it decreases with income among the upper castes. Fuller and Narasimhan (2008) study a Tamil *Brahmin* subcaste called the *Eighteen-Village Vattimas* and report that sub caste endogamy remains the norm among this caste[5]. Marriages are still arranged by the family while there is an increased say of the individuals themselves. However, as mentioned earlier, these studies are not based on nationally representative samples and hence cannot be used to infer a national picture. Our paper is the first attempt to study the aspects of inter caste marriages using a large nationally representative data set, the Indian Human Development Survey (IHDS). This novel and unique data set enables us to study the patterns in caste endogamy at the national level and examine its

in nuclear families with small number of children, are identified to be more desirable than others in due course of development. These values are considered to be inherently good and to make people more healthy, wealthy and happy (Allendorf and Thornton 2015).

[5]The *Eighteen-Village Vattimas* were erstwhile landlords in the state of Tamil Nadu. They now belong to the urban middle class.

relationship with individual and household characteristics, especially education. We also explicitly recognize and account for the marriage market structure in place in India.

3 Data

We use data from the latest round of the Indian Human Development Survey (IHDS II), jointly conducted by researchers from the University of Maryland and the National Council of Applied Economic Research (NCAER), New Delhi. The IHDS is a nationally representative household panel survey conducted in 384 districts, composed of 1420 villages and 1042 urban neighborhoods across all states and union territories of India. The second round of the survey, IHDS-II, was conducted in 2011-12[6]. The survey has detailed socio-economic and human development related questions for a household as a whole, for young children in the household and for one ever married woman in the age group of 15-49 years in each household called the 'eligible woman'.

We combine data from two schedules of the survey for our analysis: the household schedule and the eligible woman's schedule. The household questionnaire contains detailed questions about various socio-economic characteristics of the household. In the eligible woman's schedule one eligible woman was interviewed regarding health, education, fertility, family planning, marriage and gender relations in the household and the community. In the households where the eligible woman from the first round of the survey died between the survey waves or was no more in the eligible age group, a new eligible woman was interviewed, along with the old one, if present. Thus there

[6]IHDS II re-interviewed 83% of the original as well as split households residing within the village which were interviewed in IHDS-I, and an additional sample of 2134 households.

can be a maximum of 2 eligible women in each household. In households with more than one potential eligible woman, one was selected using a standard random number procedure in IHDS-I (Desai et al. 2009).

Even though caste and various caste based practices are common in India, there has been little systematic attempt so far to collect data on these aspects in a nationally representative survey. IHDS, for the first time, asks questions that help us explore along this direction. Our outcome variable, whether a marriage is an inter caste marriage, is defined using the following question in the eligible woman's questionnaire: "Is your husband's family the same caste as your natal family?" The dependent variable "ICmarriage" takes value 1 if the answer to this question is "No". A caveat is that we do not know the specific caste of the woman before her marriage, so we cannot differentiate one inter caste marriage from another on the basis of the distance between the marrying castes. However, the marriage is recognized by the responding woman as inter caste and therefore, is "closer to the lived reality of an inter caste marriage"[7].

Our main variables of interest are the years of education of the spouses and their respective parents. The years of education of the husband is obtained from the household roster. The years of education of the wife and both sets of parents are obtained from the eligible woman's questionnaire. Our set of control variables include the caste and the urban or rural location of the husband's household at the time of the survey. A household is considered living in an urban (rural) area in

[7]According to The Hindu (New Delhi, 13 November 2014, "Just 5% of Indian marriages are inter caste: survey"), the IHDS said that "..what female respondents interpreted as a "different caste" is likely to have been subjective, but ultimately closer to the lived reality of an inter-caste marriage". In her interview to The Hindu, Sonalde Desai (Senior Fellow at NCAER and Professor of Sociology at the University of Maryland) who led the IHDS, said: "So the IHDS took a simple approach and asked women whether their natal family belongs to the same caste as their husband's family, allowing us to bypass the complex issue of defining what caste means and get subjective perceptions from our respondents".

accordance with the Census 2011. We include assets and annual per capita income of the household at the time of the survey to proxy for the assets and income level of the household at the time of the marriage. Income per capita is the annual per capita income of the household measured in rupees. The assets scale is created by the IHDS itself, by summing 33 dichotomous items measuring household possessions and housing quality. These variables are obtained from the household questionnaire. We also control for the age at marriage of the bride and the comparative economic status of the two families at the time of their marriage. These variables are obtained from the eligible woman's questionnaire.

We also use the Employment and Unemployment Survey (68th round) of the National Sample Survey of India (NSS) conducted in 2011-12. We use this data set to construct average and caste-wise average years of education of females in the marriageable age group (12 to 35 years) for each district. We calculate the marriageable age group by looking at the distribution of age at marriage of the eligible women in our IHDS sample. 96.8% of women report their age at marriage to be from 12 to 35 years. We also calculate the proportion of population belonging to the same caste as that of a husband in our sample in his district of residence using this NSS data set. The average and caste-wise average education variables along with population proportion variables are used to separate any positive or negative channels through which education of the spouses may be associated with likelihood of an inter caste marriage, as explained in the introduction, via cultural adaptability and assortative matching effects.

4 Descriptive Analysis

Our specific aim in this paper is to look at the relationship between inter caste marriages and education. First we look at a broad range of descriptive statistics to get a better idea about the existing trends and dynamics of the marriage market in India in general and inter caste marriages in particular. The IHDS asks a host of questions, the responses to which aid us in understanding marriage dynamics in India and putting our particular research question in context.

We begin by looking at the trend in the rate of inter caste marriages over time. Figure 1 plots the rate of inter caste marriages by the year of marriage. The Modernization theory in Sociology, first developed by the American social scientists in the 1950s, explains the process of transition of a nation from a traditional political structure to a democratic one. The causal chains for this transition consist of "a progressive accumulation of social changes" like industrialization, urbanization, education and so on (Przeworski and Limongi 1997). One of the predictions of the Modernization theory is that with the advent of industrialization and urbanization, various non-Western family behaviours will converge towards the Western nuclear family model. As a result, there will be a decline in arranged marriages. Consequently there will also be a decline in the importance of attributes like group identities and statuses of the two families (Allendorf and Pandian 2016). However, such a clear upward trend is not visible from the figure: the rate of inter caste marriages has hovered around 5% since 1970 to 2012. The average for 2000-2012 is marginally higher than 1971-80 and 1981-90, but is not statistically different from the decade 1990-2000[8]. Thus, there has hardly been any significant movement in out-of-caste

[8]The years 2011 and 2012 have been clubbed with the decade of 2000-2010 as 2012 is the last year of survey.

marriages over the last 4 decades.

Next, in Table 1, we look at the distribution of inter caste marriages by various characteristics of the households[9]. The first panel shows that *Brahmins* have the highest rate of out-of-caste marriages, followed by Other Forward castes (OFC), while Other Backwards Classes (OBC) and Scheduled castes (SC) have the lowest rate[10]. However, it should be noted that the rate of exogamy for *Brahmins* is not statistically different from any other caste categories. The only significant differences are between the rates of OFCs and OBCs, and OFCs and SCs. This could most likely be due to the fact that OFC is a very broad administrative category comprising of many castes which may be more open to marrying among themselves compared to other caste categories[11].

The second panel of Table 1 shows the rate of inter caste marriages by the location of a household. The Modernization theory predicts that urbanization will be the harbinger of change towards a Western family model with reduced importance of caste and other group characteristics. However, contrary to this prediction, we see that the urban households do not have a higher probability of having an inter caste marriage. The difference between the urban and rural households in the rate of inter caste marriages is not statistically significant. A finer division tells us that within the urban sector, it is the metropolitan urban areas that have the lowest rate, while other urban areas have a higher rate (3.84% and 5.41% respectively)[12]. Within the rural sector, developed villages have a higher rate, while less developed villages have

[9]Here the household corresponds to that of the husband.

[10]Refer to the Appendix for a description of the social and administrative categorizations of the caste system in India.

[11]It is important to note that the data on whether a marriage is inter caste or not is self reported. A reported inter caste marriage may not necessarily involve two broad administrative caste categories. Our definition of an inter caste marriage is not based on marriage across these administrative categories.

[12]Authors' calculations from IHDS-II, not reported in the table.

a lower rate of inter caste marriages (5.72% and 4.86% respectively). This more detailed pattern again confirms that more urbanized areas do not necessarily have a higher rate of out-marriages in India.

The next two panels of Table 1 show the rate of inter caste marriages by asset and annual per capita income quartiles of the households respectively. We observe that for the asset quartiles, the rate goes down as we move up the asset distribution. The rate of inter caste marriages is significantly higher in the first asset quartile than that in the fourth quartile. A similar pattern can be observed for the income quartiles. The last panel of Table 1 shows the rate of inter caste marriages by the comparative economic statuses of the wife's and the husband's family at the time of their marriage. No difference is observed here irrespective of whether the husband's family had the same, better or worse status than the wife's family at the time of their marriage.

The observations so far make it clear that caste endogamy is much more pervasive than expected in the face of economic development and expansion of market forces. There has been no secular increase in the rate of inter caste marriages over time, nor is there an urban advantage. In fact, metropolitan areas have the lowest rate of inter caste marriages. The rate does not differ by the comparative economic status of the two families at the time of marriage either. Rather, the rate seems to go down as one moves up the asset or income quartiles.

We now move on to look at the decision making process at the time of marriage. Marriages in India, as mentioned above, generally involve the entire family in the decision making process (Banerjee et al. 2013). The IHDS II asks the eligible woman various questions related to the choice of her husband. In particular, it asks questions about who chose her husband, did she have any say in choosing her husband and

did she have any interaction with her husband before marriage. Table 2 shows the pattern. The second column of the table reports the percentages among all marriages while the third column reports that among inter caste marriages only. Among all marriages, a striking 73% of women say that parents (or other relative) chose their husbands, and in fact almost 70% of them met their husbands only on the day of their wedding/*gauna*. Even more striking is the fact that even among the women who said they chose the husbands themselves, only 67.53% said they had met their husbands before marriage and 34.13% did not know their husbands before the wedding/*gauna* day[13]. This is clearly indicative of the fact that not only marriages in India are arranged by parents in a large majority of the cases, the figure of 73% may well be an underestimate and many women who said that they chose their husbands themselves may essentially have negligible roles to play in the decision making process. Spouses barely know each other before marriage. Only a quarter of the women had met their husbands or had seen their photos before marriage; even fewer had talked to them before getting married to them (third panel of Table 2).

While the second column of Table 2 gives a good sense of the arranged marriage set up of the Indian marriage market, the third column strengthens this idea. As mentioned earlier, it looks at the decision making process only in the subset of inter caste marriages. The Modernization Theory, as mentioned earlier, predicts that there will be a decline in arranged marriages leading to a decline in caste endogamy. This implies that inter caste marriages would usually be "love" marriages and thus the involvement of parents would be minimal, at least at the decision making stage[14]. This prediction too does not hold in the data. Table 2 shows that even among the subset

[13]Authors' calculations from IHDS-II, not reported in the table.

[14]Banerji et al. (2013) use the term "love" marriages to refer to self-arranged marriages wherein couples choose their own partners with little or no inputs from their parents.

of only inter caste marriages, almost 63% of them are arranged by parents/other relatives only. And even here, among those who say that they chose their husbands on their own, 13.78% met their husbands only on the day of the wedding/*gauna*. Another notable point is that among all inter caste marriages, an overwhelming 98.07% of couples lived with their parents immediately after marriage. Thus, when a marriage takes place, inter caste or not, the parents have the primary say in a majority of the cases. This observation lends reasonable amount of support to the idea that the effect of parental attributes should be relevant in any analysis of marriages in India, along with that of individual attributes.

Finally we turn to our main attribute of interest, namely education. Figures 2 and 3 plot the rate of inter caste marriages for different educational categories of the wife and the husband, and wife's mother, wife's father, husband's mother and husband's father respectively. Figure 2 shows that this rate is not statistically different among the different educational categories of the spouses themselves. This substantiates the observations from the previous tables that individuals themselves have little role to play in the decision of their own marriage. From panels A and B of Figure 3, it can be observed that the rate of inter caste marriages does not vary by the educational categories of the fathers of the spouses[15]. The plots for the mothers, however, paint a very different picture. The rate of inter caste marriages appears to be significantly higher at higher educational categories of the mothers of the spouses[16]. This corroborates well with the earlier observation that parental attributes should be important in the analysis of marriages in India where the institution of arranged marriages plays a dominant role. In what follows, we further explore along

[15]The mean differences between any pair of educational categories of the fathers are statistically insignificant in general.

[16]The mean differences are statistically significant and positive for a number of pairs of educational categories.

these directions in a rigorous regression analysis of the relationship between inter caste marriages and education.

5 Regression Analysis

5.1 Regression specification

Our aim in this paper is to explore the relationship between education and the probability of a marriage to be an inter caste one. But whose education should matter? While a large body of the literature on interracial and inter ethnic marriages in the Western countries has focused on the education of individual spouses, marriage markets in India work very differently compared to the Western countries. Our observations from the descriptive analysis make it clear that marriages in India are arranged by the parents in a majority of the cases. Thus we must pay due attention to parental education along with the education of individual spouses.

To establish our point we proceed in two steps. First, we explore whether education levels of the spouses themselves can predict the occurrence of inter caste marriages. Considering a married couple as our unit of observation, we run the following regression:

$$ICmarriage_{id} = \alpha + \beta_1.husband's\ \ edu_{id} + \beta_2.wife's\ \ edu_{id} + \theta.X_{id} + \delta_d + \tau_t + \varepsilon_{id}. \tag{1}$$

$ICmarriage_{id}$ is a binary variable which takes value 1 if a couple i in district d is in an inter caste marriage and 0 if in an intra caste one. Our primary independent variables of interest are the education variables: *husband's* edu_{id} denotes the years of education attained by the husband in the married couple i in district d and

wife's edu_{id} is that attained by the wife.

To complete our analysis, in the next step we explore whether any of the parental education levels, along with the education levels of individual spouses, affects the likelihood of an inter caste marriage. So we run the following regression where we add the years of education of the parents of both the spouses to the set of explanatory variables considered in equation (1):

$$\begin{aligned} ICmarriage_{id} = \alpha &+ \beta_1.husband's\ \ edu_{id} + \beta_2.wife's\ \ edu_{id} \\ &+ \gamma_1.husband's\ \ mother's\ \ edu_{id} + \gamma_2.husband's\ \ father's\ \ edu_{id} \\ &+ \gamma_3.wife's\ \ mother\ \ edu_{id} + \gamma_4.wife's\ \ father's\ \ edu_{id} \\ &+ \theta.X_{id} + \delta_d + \tau_t + \varepsilon_{id}. \end{aligned} \tag{2}$$

Similar to equation (1), *husband's mother's* edu_{id}, *husband's father's* edu_{id}, *wife's mother's* edu_{id} *and wife's father's* edu_{id} are the completed years of education of the husband's parents and wife's parents respectively for couple i in district d.

In both equations (1) and (2), X_{id} is a vector of couple and household level control variables, such as caste category of the husband's household (Brahmins, OFC, OBC or SC), age at marriage of the wife and dummies for the comparative economic status of the two families at the time of the marriage. It also includes the per capita income and the assets index of the household and its location (rural or urban).

In addition to the above, we include district fixed effects, δ_d, and year of marriage fixed effects, τ_t. Marriages in India occur overwhelmingly within the district (Desai and Andrist 2010). Thus we control for any time invariant unobserved factors at the level of a district. The year of marriage fixed effects, on the other hand, control for all unobservables across districts in the year a couple got married. The year of

marriage variable, though available in our data set, has 30.66% missing values. We instead construct our own variable for the year of marriage using the year of birth of the eligible woman respondent and her age at marriage. The results are qualitatively very similar even if we use the year of marriage variable provided in our data set.

In our data set, households belonging to all religions have reported their castes. However, the caste system was originally a Hinduism phenomenon. To incorporate both these observations, the sample for our main analysis consists of only those households who have stated their religion as Hinduism, Buddhism, Jainism or Sikhism. Our choice is driven by the fact that all these religions come under the Hindu Marriage Act under the Constitution of India. We exclude scheduled tribes (STs) from our main sample mainly because even though a significant number of tribals report their religion as Hinduism, "there is sufficient heterogeneity and distinctiveness within tribal communities that they cannot be considered a part of the *varna* system" (Deshpande 2011)[17] , [18]. For our analysis we consider the 20 major states of India[19]. Thus, from a total of 33,369 couples, this selection leaves us with a sample of 25,070 couples. Standard errors are clustered at the Primary Sampling Unit (PSU) level. Table 3 provides the summary statistics for all the variables used in the regressions.

[17]It is possible that reporting or having the very knowledge of your religion or caste be dependent on the level of education of the respondent. For example, more educated individuals may not report their caste because they may be relatively liberal. While this is a possibility, we find that only 0.02% observations (8 in number) in the entire sample report their religion to be "no religion", and only 1.52% observations (505 in number) report their castes as Others or are missing.

[18]Refer to the Appendix for a description of the social and administrative categorizations of the caste system in India.

[19]This list includes the following states: Himachal Pradesh, Punjab, Uttarakhand, Haryana, Delhi, Rajasthan, Uttar Pradesh, Bihar, Assam, West Bengal, Jharkhand, Orissa, Chhattisgarh, Madhya Pradesh, Gujarat, Maharashtra, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu. We exclude the small states of North-East and Goa. We also exclude Jammu and Kashmir due to its long history of conflict.

5.2 Results

5.2.1 Inter caste marriages and own education

Table 4 reports our first set of results on the relationship between inter caste marriages and spouses' own education. Columns 1 and 2 report results from the estimation of equation (1). In column 1, we report regression coefficients from the parsimonious specification with only caste controls and the education levels of the spouses. We find that the education of neither the husband nor the wife is associated with the likelihood of an inter caste marriage. In column 2, we add a host of control variables. These controls include age at marriage of the wife, economic status of the wife's natal family as compared to husband's family at the time of their marriage, the annual per capita income of the husband's household, its asset index and its rural or urban location. The addition of these controls has no effect on the coefficients of the spouses' own education – they remain statistically insignificant. Both these regressions include year of marriage fixed effects and district fixed effects.

This result stands in sharp contrast to the findings in the existing literature on out-marriages in the Western countries, especially in the USA, where individual's own education shows up as a predictor of one's marriage being within or outside one's race or ethnicity. Recall the two potentially opposing effects of education – the 'cultural adaptability effect' and the 'assortative matching effect' – identified by Furtado (2012). This absence of relationship might be a result of these opposing effects canceling each other. Therefore, to probe further along this direction, we

adopt the model suggested by Furtado (2012) to the Indian context:

$$\begin{aligned} ICmarriage_{icd} = \kappa + \lambda.husband's\ \ edu_{icd} + \pi_1.(avg\ \ edu_{cd} - avg\ \ edu_d) \\ + \pi_2.husband's\ \ edu_{icd}.(avg\ \ edu_{cd} - avg\ \ edu_d) \\ + \mu_1.pop\ \ pr_{cd} + \mu_2.pop\ \ pr_{cd}^2 + \sigma.X_{id} + \Psi_s + \tau_t + \xi_{icd}. \end{aligned} \tag{3}$$

The dependent variable is a dummy which takes value 1 if husband i of caste c in district d is in an inter caste marriage[20]. $avg\ \ edu_{cd}$ is the average education level of females in the marriageable age group (12 to 35 years) in husband i's caste c in district d, while $avg\ \ edu_d$ is the average education level of all females in the marriageable age group in husband i's district d. The coefficient π_2 captures the assortative matching effect of education. A man with a higher level of education is more likely to find a higher educated woman from his own caste if the average education level of the women of his caste is higher than the district average. Thus the expected sign of π_2 is negative if the assortative matching effect of education is at work. $pop\ \ pr_{cd}$ is the proportion of female population in the marriageable age group of husband i's caste c in district d. We include state fixed effects, Ψ_s, instead of district fixed effects because our regressors are district level variables.

Given the potentially opposite directions of the effects of education, the net effect may be positive, negative, or even null, depending upon a caste's characteristics. We, therefore, want to ensure against the possibility that the coefficients in columns 1 and 2 in Table 4 are not statistically significant due to this phenomenon. The results are reported in columns 3 to 5 in Table 4.

In column 3, we again report results from regressing the inter caste marriage dummy on the years of education of the individuals, but use the sample of only

[20]Since we do not know the caste of the wife in a couple, our sample consists of the husbands only.

husbands here to make it comparable to the regressions in the next two columns. The coefficient of this variable shows the overall association of the husband's education with the probability of his marriage being an inter caste one. This coefficient mirrors the results in the first two columns[21].

In column 4, we add the population proportion of same caste females as that of the husband's in his district of residence. This variable captures the enclave effect: the population proportion gives the likelihood that the individual will encounter a potential spouse of the same caste in his relevant region of search, which we assume to be the district. It can be seen that the addition of this control and its square term has no effect on the coefficient of the husband's education variable. The coefficients on the variables themselves are also statistically insignificant.

Finally in column 5, we add the assortative matching term to the previous specification. This term, as explained above, is the interaction between the years of education of the husband and the education gap between the average education of the females (in the marriageable age group) in his caste and that of all females in his district. Again we find that the estimated coefficient of this variable is statistically insignificant and it does not affect the coefficient of husband's own education (it remains statistically insignificant, although its sign has reversed). Thus, even after we explicitly take into account the potential channels through which education might have an effect, we find that neither of these channels predict the likelihood of an inter caste marriage. This set of regressions reinforces our result that the education of the individuals does not predict the likelihood of a marriage being an inter caste one.

[21]The difference comes from the fact that here we use state fixed effects instead of district fixed effects because our regressors are district level variables.

5.2.2 Inter caste marriages and parental education

Now we move on to the estimation of equation (2) where, to take into account the practice of arranged marriages in India, we add the education level of the parents of both the spouses to our set of explanatory variables. For the sake of comparison, column 1 in Table 5 reproduces the column 1 of Table 4. In column 2, we report results from the estimation of equation (2) where we add the education levels of the parents of the spouses. We find that the education of the husband's mother has a positive and statistically significant association with the probability of an inter caste marriage. A one-year increase in education of the husband's mother increases the probability of an inter caste marriage by 0.18 percentage points. The results in both the columns 1 and 2 are consistent with our descriptive analysis where we observed that parents have the major say in any marriage in India and individuals themselves have a very little role to play.

In columns 3 and 4 we successively add controls to the base specification. In column 3 we add the age at marriage of the wife and dummies indicating whether the economic status of the wife's natal family was better, same or worse than that of the husband's family at the time of the marriage. The addition of these variables has little effect on the coefficient of the husband's mother's education.

Finally in column 4, our preferred specification, we add current income and assets of the household, and whether the household was located in an urban or rural area. These variables also have no effect on our main variable of interest. Thus, we can conclude this section with the key finding that husband's mother's education predicts the likelihood of an inter caste marriage and that it is robust to the inclusion of a number of controls and fixed effects. The magnitude of this relationship is also quite large. One standard deviation increase in husband's mother's years of education

leads to a 10.6% increase (over the sample mean) in the probability of a couple's marriage being an inter caste one. In other words, increase in years of education of the husband's mother by 10 years would lead to an increase in the probability of inter caste marriage by 1.86 percentage points which is equivalent to approximately 36 percent of the sample mean.

5.2.3 Robustness checks

In this section we further test the robustness of our results that the husband's mother's education has a positive and statistically significant association with the likelihood of an inter caste marriage while the education of the spouses themselves do not have any statistically significant association.

In the first set of robustness checks, we want to see if the results we obtain in the previous sections are specific to the sample or they are robust to variations in the sample. Table 6 shows the results for four such samples. Since caste system, as mentioned above, is theoretically a Hinduism phenomenon, in the first column in Table 6, we look at the sample of only Hindus and drop all those households who report their religions to be Buddhism, Jainism or Sikhism. In column 2, we expand the sample to include all religions in the major states because, as mentioned earlier, in our data set households belonging to all religions have reported their castes. In the next two columns, we expand the sample further to include all religions in all states and to all religions and all castes (including the STs) in all states, respectively[22]. The results reported in all the columns are qualitatively similar to those in the main regression. Thus, our results obtained in the previous subsections are robust to all the above mentioned selections applied to our sample: the education of the spouses

[22]We included only 20 major states in our main regression. Also, we excluded STs from the main results.

themselves do not matter whereas that of the husband's mother has a positive and statistically significant association with the likelihood of an inter caste marriage[23].

It can be argued that many women may continue their education after marriage too, and this could potentially contaminate the results since such women will actually have a lower amount of education at the time of their marriage as compared to what is measured by the data. To address this concern, we remove such women from our sample and check if the results still hold. The results are reported in column 1 in Table 7. It can be seen that the results are qualitatively the same even for this sample where we remove these observations[24].

One may argue that our results are driven by the households where the bride had more decision making power in her marriage. This may bias the coefficient on the education of the husband's mother upwards as a greater decision making power of the brides in their marriages may be positively correlated with both husband's mother having more education as well as with the probability of an inter caste marriage. Therefore, we look at the sample of only parents-arranged marriages, or simply arranged marriages as they are commonly known. We define arranged marriages as marriages in which the eligible woman's response to the question "Who chose your husband" was either "Parents/other relative alone" or "Others". The results for this sample are shown in the second column in Table 7. It can be seen that even here own education of the spouses has no association but the education of the husband's

[23]Apart from these samples, we ran the regressions for the following other combinations of religions, castes and states: four main religions, main states, including STs; all religions, main states, including STs; and four main religions, all states, excluding STs. Our results are robust to all these sample specifications. These results, not reported here, are available upon request.

[24]It should be noted that the inclusion of women who continued education after marriage should have, in fact, overestimated the results, making own education statistically significant even when actually it is not. The statistical insignificance of the education variables of the spouses' education coefficients, even when these observations are not dropped, only strengthens the claim that own education does not predict the likelihood of inter caste marriages.

mother has a positive and statistically significant association with the probability of an inter caste marriage.

In the last column of Table 7, we add to our district fixed effects and year of marriage fixed effects, another set of fixed effects: the interaction of district and year of marriage fixed effects. Thus, in addition to controlling for any unobservables at the level of the district and at the level of the year of marriage, we control for all unobservables at the level of a particular district-year as well. The coefficient of husband's mother's education is still positive and statistically significant as can be seen from column 3. Also, spouses' own education does not show any association. Thus we conclude that our results are robust to variations in the sample, to the addition of a number of controls as well as to the addition of a number of fixed effects.

6 Discussion

Our analysis of the relationship between education and the age-old practice of inter caste marriages in India highlights the importance of recognizing the arranged marriages institution in Indian marriage markets. We first establish the interesting result that the education levels of the individual spouses themselves do not have any statistically significant association with the probability of their marriage being an inter caste one. This is in sharp contrast to the findings in the existing literature on interracial and inter ethnic marriages in the Western countries where the marriage market works very differently with decisive roles played by the individuals themselves. We complete our analysis by establishing that the education level of the husband's mother has a positive, statistically significant and quite large association with the likelihood of an inter caste marriage. This reinforces the idea that parents

are indeed the main players in the marriage market in India.

The second part of our findings is nuanced in the following two ways. First, only the education of the husband's mother predicts inter caste marriage, but not that of his father. Second, education of the wife's parents are not associated with the likelihood of an inter caste marriage. In this section we briefly discuss the potential channels for these differential associations that we observe.

To understand the first result we put together three stylized facts. Firstly, a growing body of literature finds evidence that a more educated woman has an increased bargaining power and decision making power in a household[25]. It also makes the household more gender balanced. Secondly, it is also well documented, especially in the context of developing countries, that a mother is more responsive to the needs of her child, as compared to the father[26]. Provided with the resources, a mother is more likely to utilize them in the best possible interest of her children. A father, on the other hand, is more likely to spend it on various adult consumption goods like tobacco, liquor and so on. Finally, from our own analysis and from the literature cited in previous sections, we know that a marriage decision in India is undertaken by the family in a majority of the cases, irrespective of whether the marriage is endogamous or not. Combining these three stylized facts we try to understand the first aspect of our finding as follows. Given that we are looking at marriages ex-post, they must be revealed preferred to be the optimal matches from all the potential matches available. An intra caste match could, then, be a constrained optimum if the father, driven by the prestige or reputation of the family and being less sensitive to the best outcome for the son, insists on the intra caste constraint. An inter caste marriage

[25]See, for example, Beegle et al. 2001; Doss 2013; Thomas 1994.

[26]See, for example, Duflo 2000; Duflo and Udry 2004; Friedberg and Webb 2006; Haddad and Hoddinott 1995; Lundberg et al. 1997; Phipps and Burton 1998; Thomas 1990.

is more likely to occur when an educated mother can overcome this constraint and implement the best outcome for the son, empowered by her increased bargaining and decision making authority in the family.

Consider next the second aspect of our finding that only the education of the husband's mother has a statistically significant association, but not that of the wife's parents. We argue that this asymmetry between the two families arises from the fact that in any inter caste marriage the bride's family bears more stigma or costs than the groom's family. Some theoretical backing for this is provided by the analytical model in Bidner and Eswaran (2015) where stability of the endogamy equilibrium requires that the punishment for deviation from the equilibrium should be greater for a female and her family as compared to her male counterpart[27]. While we could not find any empirical work on this asymmetry that arises in equilibrium, much of the anecdotal evidence involving "honour" killings in India validates our assertion[28]. Honour killing is killing someone in the name of family honour with the belief that the act will redeem the reputation of the family. It is often committed in cases where a couple marries against the wishes of the family, especially across caste lines. The fact that the crime is generally perpetrated by the bride's family, in which either or both of the spouses are killed, suggests that these families correctly expect to face

[27]This result is obtained without assuming any ex-ante asymmetry in the preferences of the two families. It is a result of the technical complementarity between the inputs of the husband and the wife and patrilocality where the wife adopts the occupation of the husband. This means that a deviation is simply more beneficial for a female and it imposes a cost on the male member of the group who remains unmarried. The result holds true for the upper most castes. For it to hold for the lowest castes, it needs to be combined with patriarchy. Patriarchy implies that the cost to a male member of not leaving behind a progeny is much higher than that to a female. Therefore, the cost of a potential mismatch where a man is left unmarried is higher for the group. As a result, the punishment of violation is also higher for the females (Bidner and Eswaran 2015).

[28]The Tribune, Chandigarh (03 July 2007): "*Honour killing rocks state, again*" (Manoj Babli honour killing Case); Times of India, New Delhi (20 November 2011): "*Parents held for 'honour' killing of 21-year-old Delhi University girl*"; The Indian Express, Ludhiana (09 May 2016): "*'Honour killing': Man kills daughter over relationship*"; Aljazeera (07 December 2016): "*India sees huge spike in 'honour' killings*".

the greater burden of the stigma of an inter caste marriage.

Our argument here is that education may not have enough mitigating effect on the stigma of an inter caste marriage for the bride's family which bears these costs disproportionately. Similar to the groom's father, the bride's father's education is not associated with the likelihood of inter caste marriage. However, unlike the case of the groom's mother, the education of bride's mother also has no association. This difference may be due to the fact that unlike the groom's family, the bride's family bears a significant cost of an inter caste marriage. In other words, education works through giving more voice to the mother in the household to implement the best outcome for her child, if the stigma or social costs of an inter caste marriage is not too high.

7 Conclusion

We look at the relationship between education and the practice of caste endogamy, which is the defining and one of the most resilient features of the caste system in India. Using a nationally representative data set, the second round of the Indian Human Development Survey, we report novel and interesting findings. The rate of inter caste marriages in India is only 5.82% even in 2011, and there has been no secular increase in this rate over the previous four decades. In keeping with the existing literature, descriptive analysis of our data set shows that in the Indian marriage market families, rather than individuals, are the primary decision makers. An overwhelming 73% of marriages are arranged by parents, and spouses have very little contact with each other before marriage. Interestingly, this pattern holds true for inter caste marriages as well.

Our regression analysis brings out two important results. First, contrary to the

findings of the existing literature on Western countries, the education level of an individual does not predict the likelihood of his/her marriage being an inter caste one. Our results are reinforced when we find the same to hold true even if we decompose the effects of education into its constituent channels – the 'cultural adaptability effect' and the 'assortative matching effect' – as identified in the literature (Furtado 2012). Second, complementing the observations from our descriptive analysis, we find that it is the education of the husband's mother that has a positive and statistically significant association with the likelihood of an inter caste marriage. Our results are robust to the inclusion of a host of control variables, a wide range of variations in the sample, *and* a varied set of fixed effects, which includes district fixed effects, year of marriage fixed effects as well as district and year of marriage interaction fixed effects. We posit that education works through giving more voice to the mother in the household to implement the best outcome for her child, if the stigma or cost of an inter caste marriage is not too large. Given that the bride's family disproportionately bears the stigma of an inter caste marriage, education of only the groom's mother has a positive association

Thus our analysis highlights the importance of recognizing the institution of arranged marriage in any analysis of Indian marriage markets. Taken together, the two aspects of our result indicate that once the arranged marriage set up is recognized, one can easily understand the result that education has no effect on the decision of one's own marriage, but only on the decision of the marriage of one's offspring.

References

Ahuja, A. and Ostermann, S. L. (2016). Crossing caste boundaries in the modern Indian marriage market. *Studies in Comparative International Development*, 51(3):365–387.

Allendorf, K. and Pandian, R. K. (2016). The decline of arranged marriage? Marital change and continuity in India. *Population and Development Review*, 42(3):435–464.

Allendorf, K. and Thornton, A. (2015). Caste and Choice: The influence of Developmental Idealism on Marriage Behavior. *American Journal of Sociology*, 121(1):243–287.

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.

Banerji, M., Martin, S., and Desai, S. (2013). Are the young and educated more likely to have 'love' than arranged marriage? A study of autonomy of partner choice in India. *Working Paper Series (pp. 1À43). New Delhi: NCAER.*

Batson, C. D., Qian, Z., and Lichter, D. T. (2006). Interracial and intraracial patterns of mate selection among Americas diverse Black populations. *Journal of Marriage and Family*, 68(3):658–672.

Beegle, K., Frankenberg, E., and Thomas, D. (2001). Bargaining power within couples and use of prenatal and delivery care in Indonesia. *Studies in Family Planning*, 32(2):130–146.

Beteille, A. (1971). Race, caste and ethnic identity. *International Social Science Journal*, 23(4):519–535.

Bidner, C. and Eswaran, M. (2015). A gender-based theory of the origin of the caste system of India. *Journal of Development Economics*, 114:142–158.

Blau, P. M., Blum, T. C., and Schwartz, J. E. (1982). Heterogeneity and intermarriage. *American Sociological Review*, pages 45–62.

Chowdhry, P. (1997). Enforcing cultural codes: Gender and violence in northern India. *Economic and Political Weekly*, pages 1019–1028.

Conte, E. (1979). Politics and marriage in South Kanem (Chad): A statistical presentation of endogamy from 1895 to 1975. *Cahiers de LOrstom*, pages 262–275.

Desai, S. and Andrist, L. (2010). Gender scripts and age at marriage in India. *Demography*, 47(3):667–687.

Desai, S., Dubey, A., Joshi, B., Sen, M., Shariff, A., and Vanneman, R. (2009). India human development survey: Design and data quality–Technical paper no. 1. Technical report.

Deshpande, A. (2011). *The grammar of caste: Economic discrimination in contemporary India*. Oxford University Press.

Doss, C. (2013). Intrahousehold bargaining and resource allocation in developing countries. *The World Bank Research Observer*, 28(1):52–78.

Duflo, E. (2000). Grandmothers and grandaughters: the effects of old age pension on child health in South Sfrica. Technical report, National Bureau of Economic Research.

Duflo, E. and Udry, C. (2004). Intrahousehold resource allocation in Côte d'Ivoire: Social norms, separate accounts and consumption choices. Technical report, National Bureau of Economic Research.

Dugar, S., Bhattacharya, H., and Reiley, D. (2012). Can't Buy Me Love? A field experiment exploring the trade-off between income and caste-status in an Indian matrimonial market. *Economic Inquiry*, 50(2):534–550.

Dumont, L. (1980). *Homo hierarchicus: The caste system and its implications*. University of Chicago Press.

Ellis, M., Wright, R., and Parks, V. (2006). The immigrant household and spatial assimilation: Partnership, nativity, and neighborhood location. *Urban Geography*, 27(1):1–19.

Friedberg, L. and Webb, A. (2006). Determinants and consequences of bargaining power in households. Technical report, National Bureau of Economic Research.

Fryer, R. G. (2007). Guess who's been coming to dinner? Trends in interracial marriage over the 20th century. *The Journal of Economic Perspectives*, 21(2):71–90.

Fu, V. K. (2001). Racial intermarriage pairings. *Demography*, 38(2):147–159.

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.

Gullickson, A. (2006). Education and black-white interracial marriage. *Demography*, 43(4):673–689.

Gullickson, A. and Torche, F. (2014). Patterns of racial and educational assortative mating in Brazil. *Demography*, 51(3):835–856.

Haddad, L. and Hoddinott, J. (1995). Does female income share influence household expenditures? Evidence from Côte d'Ivoire. *Oxford Bulletin of Economics and Statistics*, 57(1):77–96.

Hwang, S.-S., Saenz, R., and Aguirre, B. E. (1994). Structural and individual determinants of outmarriage among Chinese-, Filipino-, and Japanese-Americans in California. *Sociological Inquiry*, 64(4):396–414.

Hwang, S.-S., Saenz, R., and Aguirre, B. E. (1995). The SES selectivity of interracially married Asians. *International Migration Review*, pages 469–491.

Iceland, J. and Nelson, K. A. (2010). The residential segregation of mixed-nativity married couples. *Demography*, 47(4):869–893.

Kalmijn, M. and Van Tubergen, F. (2010). A comparative perspective on intermarriage: Explaining differences among national-origin groups in the United States. *Demography*, 47(2):459–479.

Kaur, R. (2010). Khap panchayats, sex ratio and female agency. *Economic and Political Weekly*, pages 14–16.

Luke, C. and Luke, A. (1998). Interracial families: Difference within difference. *Ethnic and Racial Studies*, 21(4):728–754.

Lundberg, S. J., Pollak, R. A., and Wales, T. J. (1997). Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *Journal of Human Resources*, pages 463–480.

Meng, X. and Gregory, R. G. (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor Economics*, 23(1):135–174.

Phipps, S. A. and Burton, P. S. (1998). What's mine is yours? The influence of male and female incomes on patterns of household expenditure. *Economica*, 65(260):599–613.

Przeworski, A. and Limongi, F. (1997). Modernization: Theories and facts. *World Politics*, 49(02):155–183.

Qian, Z. (1997). Breaking the racial barriers: Variations in interracial marriage between 1980 and 1990. *Demography*, 34(2):263–276.

Qian, Z., Glick, J. E., and Batson, C. D. (2012). Crossing boundaries: Nativity, ethnicity, and mate selection. *Demography*, 49(2):651–675.

Qian, Z. and Lichter, D. T. (2001). Measuring marital assimilation: Intermarriage among natives and immigrants. *Social Science Research*, 30(2):289–312.

Saenz, R., Hwang, S.-S., Aguirre, B. E., and Anderson, R. N. (1995). Persistence and change in Asian identity among children of intermarried couples. *Sociological Perspectives*, 38(2):175–194.

Srinivas, M. N. (1962). *Caste in modern India and other essays.* Bombay and London: Asia Publ. House.

Stevens, G. and Swicegood, G. (1987). The linguistic context of ethnic endogamy. *American Sociological Review*, pages 73–82.

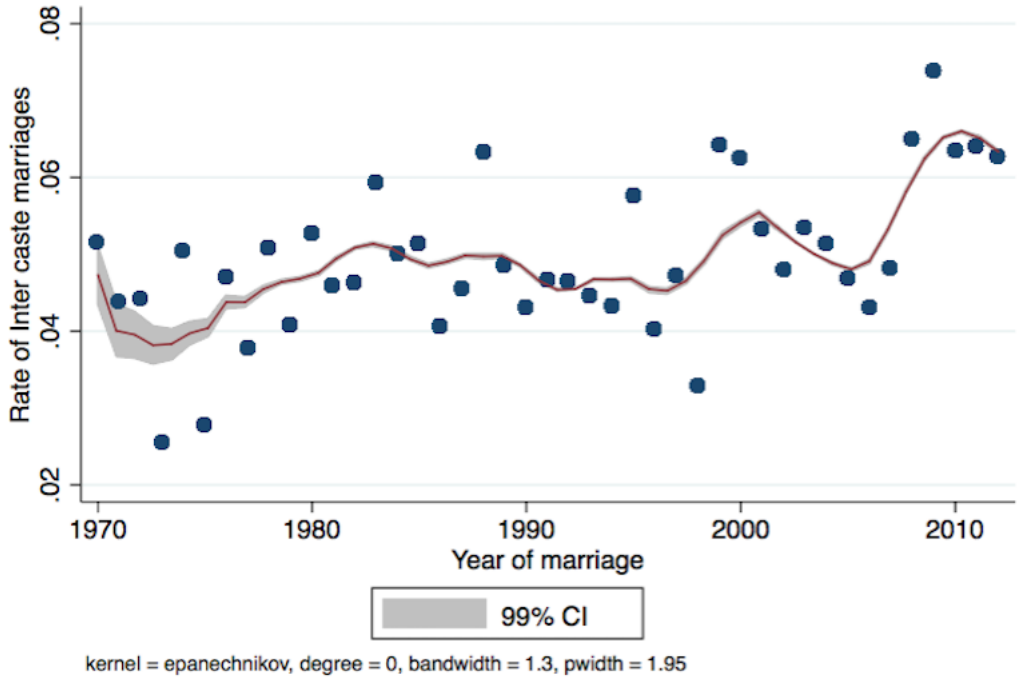
Thomas, D. (1990). Intra-household resource allocation: An inferential approach. *Journal of Human Resources*, pages 635–664.

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

Wong, L. Y. (2003). Why so only 5.5% of black men marry white women? *International Economic Review*, 44(3):803–826.

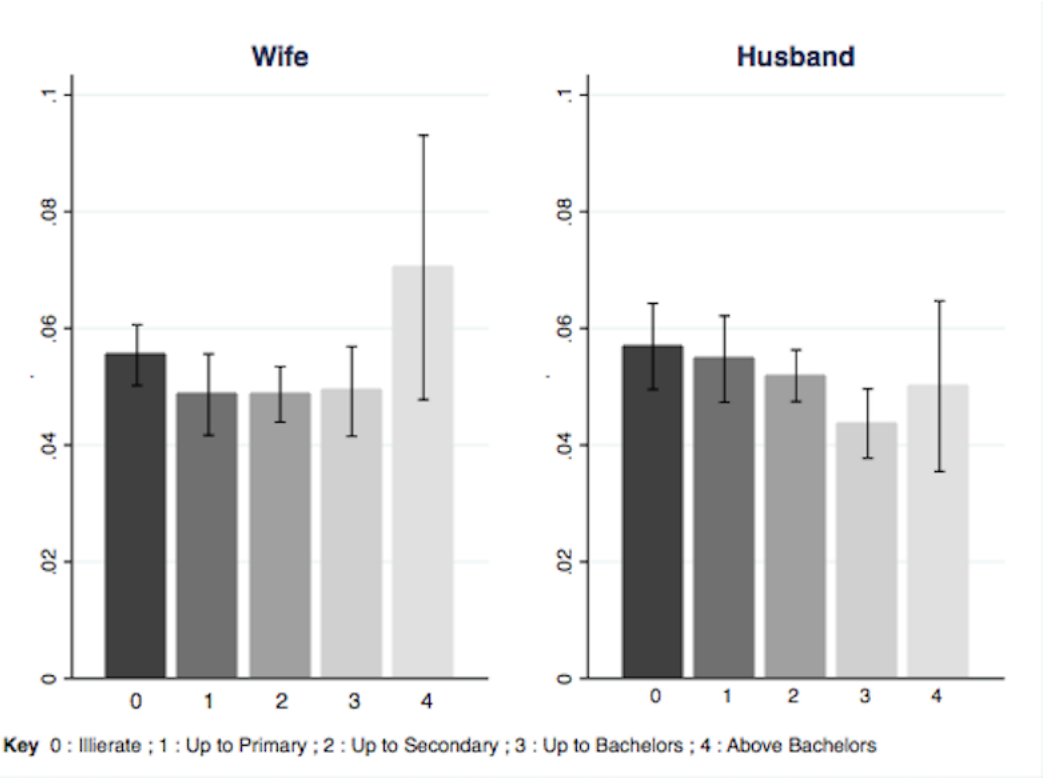
Zhang, Y. and Van Hook, J. (2009). Marital dissolution among interracial couples. *Journal of Marriage and Family*, 71(1):95–107.

Figure 1: Trend in the rate of inter caste marriages



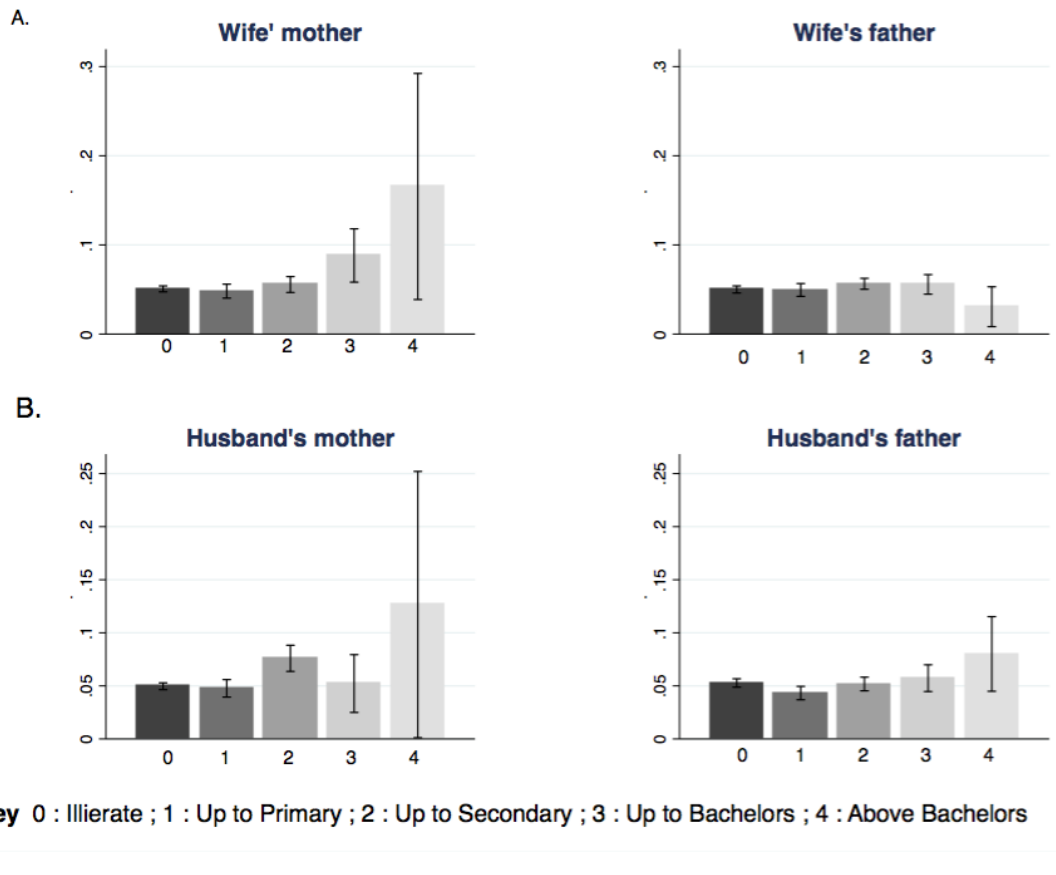
Note: The smooth line plots the local polynomial regression of the yearly percentage of inter caste marriages on the year of marriage. Data source is IHDS II.

Figure 2: Rate of inter caste marriages and education of the spouses



Note: 95% confidence intervals indicated. Data source is IHDS II. The y axis stands for the rate of inter caste marriages. The left panel plots the rate of inter caste marriages by education of the wife while the right panel plots it by the education of the husband.

Figure 3: Rate of inter caste marriages and education of the parents



Note: 95% confidence intervals indicated. Data source is IHDS II. The y axis stands for the rate of inter caste marriages. Panel A plots the rate of inter caste marriages by the education of the wives' parents. Panel B plots the rate by the education of the husbands' parents.

Table 1: Rate of inter caste marriages by household characteristics

Caste	**Rate of Inter caste marriage**
Brahmins	6.30***
	(0.656)
Other Forward Castes	6.20***
	(0.341)
Other Backward Castes	4.80***
	(0.216)
Scheduled Castes	4.76***
	(0.269)
Type of Residence	
Urban	4.99***
	(0.246)
Rural	5.24***
	(0.184)
Asset quartiles	
First quartile	5.89***
	(0.317)
Second quartile	5.48***
	(0.318)
Third quartile	5.01***
	(0.273)
Fourth quartile	4.01***
	(0.266)
Income quartiles	
First quartile	5.08***
	(0.337)
Second quartile	5.58***
	(0.312)
Third quartile	4.07***
	(0.259)
Fourth quartile	4.89***
	(0.273)
Comparative Economic Status of wife's family (at the time of marriage)	
Same	4.98***
	(0.169)
Better	5.92***
	(0.387)
Worse	5.20***
	(0.480)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis. Data source is IHDS II.

Table 2: Decision making at the time of marriage

Who chose the husband	**All marriages**	**Inter caste marriages**
	(percent)	(percent)
Respondent herself	3.91***	15.01***
	(0.122)	(1.1.5)
Respondent and parents/other relative	22.70***	21.68***
	(0.286)	(1.33)
Parents/other relative alone	73.01***	62.83***
	(0.300)	(1.56)
Others	0.29***	0.49***
	(0.036)	(0.443)
Knew husband for how long before marriage		
On wedding/gauna day only	69.69***	66.5***
	(0.313)	(1.52)
Less than a month	13.33***	12.3***
	(0.232)	(1.06)
More than one month but less than one year	7.43***	5.82***
	(0.180)	(0.775)
More than one year	3.64***	11.7***
	(0.128)	(1.04)
Since childhood	5.46***	3.44***
	(0.155)	(0.588)
Met husband before marriage	23.43***	32.8***
	(0.287)	(1.52)
Saw photo of husband before marriage	26.72***	30.8***
	(0.301)	(1.49)
Talked to husband before marriage	15.64***	22.1***
	(0.246)	(1.34)
Chatted over email with husband before marriage	1.69***	3.45***
	(0.0856)	(0.591)
Living immediately after marriage		
With parents	99.2***	98.07***
	(0.062)	(0.445)
Alone	0.82***	1.93***
	(0.0615)	(0.443)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis. Data source is IHDS II.

Table 3: Summary statistics

S.No	Variable	Mean	Standard Deviation
1	Inter caste marriage (binary variable)	0.0516	0.22
2	Wife's edu (years)	5.51	4.95
3	Husband's edu (years)	7.43	4.82
4	Husband's mother's edu (years)	1.26	2.82
5	Husband's father's edu (years)	3.33	4.42
6	Wife's mother's edu (years)	1.63	3.18
7	Wife's father's edu (years)	3.81	4.68
8	Age at marriage (Wife) (years)	17.61	3.55
9	Annual income per capita (INR)	25882.61	46471.64
10	Assets (Index)	15.76	6.46
11	Urban (binary variable)	0.3352	0.47

Table 4: Inter caste marriages and own education

	(1) ICmarriage	(2) ICmarriage	(3) ICmarriage	(4) ICmarriage	(5) ICmarriage
husabnd's edu	-0.000351	-0.000102	-0.000409	-0.000433	0.000431
	(0.000543)	(0.000520)	(0.000735)	(0.000735)	(0.00145)
wife's edu	-0.000776	-0.000546			
	(0.000839)	(0.000814)			
pop pr				0.0713	0.0718
				(0.0694)	(0.0706)
pop pr sq				-0.127	-0.127
				(0.0891)	(0.0905)
husband's edu* (avg $edu_{cd} - avg\ edu_d$)					4.46×10^{-8}
					(6.40×10^{-8})
(avg $edu_{cd} - avg\ edu_d$)					-0.000000266
					(0.000000956)
Controls I		√			
Controls II		√	√	√	√
Caste controls	√	√	√	√	√
Year of marriage FE	√	√	√	√	√
District FE	√	√			
State FE			√	√	√
N	22476	22469	22470	22158	22146
R^2	0.221	0.222	0.033	0.034	0.034

Note: Linear probability results are reported. Data sources are IHDS-II and Schedule 10 of NSS Round 68 (2011-12).Outcome is a dummy variable which takes value 1 if the marriage is inter caste, 0 otherwise. *pop pr* is the proportion of population that belongs to the same caste as husband's caste and captures the potential enclave effect of education. *(avg $edu_{cd} - avg\ edu_d$)* is the difference between the average education of females in the marriageable age in the husband's caste in his district and that of all females in the marriageable age in the husband's district. *husband's edu*(avg $edu_{cd} - avg\ edu_d$)* is the interaction between the education difference term and husband's own education which captures the potential assortative matching effect of education. Controls I consists of age at marriage of the wife and economic status of the wife's natal family as compared to the husband's family at the time of marriage. Controls II consists of per capita annual income of the husband's family, its assets and its rural or urban location at the time of the survey. Robust standard errors clustered at the primary sampling unit level are in paranthesis. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$. Regressions weighted by survey weight of the eligible woman.

Table 5: Inter caste marriages and parental education

	(1) ICmarriage	(2) ICmarriage	(3) ICmarriage	(4) ICmarriage
husabnd's edu	-0.000351	-0.000364	-0.000364	-0.0000839
	(0.000543)	(0.000545)	(0.000549)	(0.000530)
wife's edu	-0.000776	-0.00117	-0.00110	-0.000886
	(0.000839)	(0.000831)	(0.000849)	(0.000820)
husband's mother's edu		0.00181**	0.00186**	0.00186**
		(0.000889)	(0.000889)	(0.000874)
husband's father's edu		-0.000953	-0.000932	-0.000842
		(0.000626)	(0.000635)	(0.000632)
wife's mother's edu		0.00105	0.00109	0.00104
		(0.000927)	(0.000929)	(0.000917)
wife's father's edu		0.000284	0.000274	0.000327
		(0.000524)	(0.000526)	(0.000514)
Controls I			√	√
Controls II				√
Caste controls	√	√	√	√
Year of marriage FE	√	√	√	√
District FE	√	√	√	√
N	22476	22251	22251	22244
R^2	0.221	0.223	0.223	0.224

Note: Linear probability results are reported. Data source is IHDS-II. Outcome is a dummy variable which takes value 1 if the marriage is inter caste, 0 otherwise. Controls I consists of age at marriage of the wife and economic status of the wife's natal family as compared to the husband's family at the time of marriage. Controls II consists of per capita annual income of the husband's family, its assets and its rural or urban location at the time of the survey. Robust standard errors clustered at the primary sampling unit level are in paranthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions weighted by survey weight of the eligible woman.

Table 6: Robustness checks: Variations in the religious and caste composition of the samples

	(1) Only Hindus	(2) All religions, main states	(3) All religions, all states	(4) All religions, all castes, all states
husabnd's edu	-0.0000800	-0.000126	0.0000413	0.0000288
	(0.000542)	(0.000470)	(0.000474)	(0.000451)
wife's edu	-0.000913	-0.00106	-0.000986	-0.000760
	(0.000848)	(0.000726)	(0.000716)	(0.000705)
husband's mother's edu	0.00199**	0.00191**	0.00169**	0.00146*
	(0.000928)	(0.000833)	(0.000816)	(0.000836)
husband's father's edu	-0.000978	-0.000727	-0.000575	-0.000569
	(0.000637)	(0.000582)	(0.000573)	(0.000558)
wife's mother's edu	0.000917	0.000457	0.000674	0.000773
	(0.000955)	(0.000839)	(0.000827)	(0.000810)
wife's father's edu	0.000356	0.000501	0.000388	0.000278
	(0.000527)	(0.000476)	(0.000473)	(0.000469)
Controls I	√	√	√	√
Controls II	√	√	√	√
Caste controls	√	√	√	√
Year of marriage FE	√	√	√	√
District FE	√	√	√	√
N	21309	25693	26707	29030
R^2	0.226	0.198	0.230	0.220

Note: Linear probability results are reported. Outcome is a dummy variable which takes value 1 if the marriage is inter caste, 0 otherwise. Data source is IHDS II. Column 1 uses the sample of only Hindus in the main states. Column 2 uses the sample of all religions, excluding STs, in the main states. Column 3 uses the sample of all religions, excluding STs in all states. Column 4 includes all religions, all castes including STs in all states. Robust standard errors clustered at the primary sampling unit level are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions weighted by survey weight of the eligible woman.

Table 7: Robustness checks: Variations in the sample of women and inclusion of interaction fixed effects

	(1) Completed education before marriage	(2) Only arranged marriages	(3) District*Year of marriage FE
husabnd's edu	-0.000197	-0.000167	-0.000675
	(0.000538)	(0.000639)	(0.000735)
wife's edu	-0.000979	-0.000659	-0.000100
	(0.000875)	(0.000645)	(0.000829)
husband's mother's edu	0.00226**	0.00210**	0.00220*
	(0.000951)	(0.000862)	(0.00123)
husband's father's edu	-0.000878	-0.00110	-0.00103
	(0.000670)	(0.000704)	(0.000790)
wife's mother's edu	0.000978	0.000313	0.000633
	(0.000968)	(0.000736)	(0.00109)
wife's father's edu	0.000264	0.000696	0.000550
	(0.000539)	(0.000499)	(0.000684)
Controls I	√	√	√
Controls II	√	√	√
Caste controls	√	√	√
Year of marriage FE	√	√	√
District FE	√	√	√
District*Year of marriage FE			√
N	21269	16439	22244
R^2	0.229	0.339	0.549

Note: Linear probability results are reported. Outcome is a dummy variable which takes value 1 if the marriage is inter caste, 0 otherwise. Data source is IHDS II. Column 1 uses the sample of only those women who had completed their education before they got married. Column 2 uses the sample of only arranged marriages defined as in text. Column 3 adds interaction of district and year of marriage fixed effects to the set of district fixed effects and year of marriage fixed effects. Robust standard errors clustered at the primary sampling unit level are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions weighted by survey weight of the eligible woman.

8 Appendix

8.1 *Varna*, *jati*, and Caste categories

According to Deshpande (2011), in the ancient Hindu society, the institution of caste was divided into initially four and later five mutually exclusive *varnas* which were hereditary, endogamous and occupation specific. They were called *Brahmins* (priests and teachers), *Kshatriyas* (warriors and the royalty), *Vaishyas* (traders, merchants and money lenders) and *Shudras* (peasants and other menial and lowly job workers). The fifth category were the *Atishudras* who did the most polluting and menial jobs. These were the formal untouchables. The *varnas* are theoretically ranked according to the following hierarchy: *Brahmins* at the top, followed by *Kshtriyas*, *Vaishyas* and then *Shudras*. The *Atishudras* were the lowliest of the low and were in fact called the *avarnas* or without a *varna*. In other words, they were excluded from the caste system.

The building blocks of the contemporary social code are *jatis*, which are subcategories of the *varnas*. However, there does not exist a one-to-one mapping of a *jati* to a *varna*. There is a lot of fluidity and ambiguity involved in their categorization due to the numerous, and in most cases, unverifiable, claims of *varna* affiliations made by the more than 3000 *jatis* in India (Deshpande 2011).

The caste categories used in this paper are, on the other hand, administrative categories. When the Affirmative Action policies were being formulated, jatis which were economically the weakest and were historically subjected to discrimination and deprivation were identified in a government schedule as the target group for reser-

vation policies (Deshpande, 2011). They are referred to as the Scheduled Castes or SC. Another schedule identified similarly placed tribes for the reservation policy and they are referred to as the Scheduled Tribes or ST.

The Mandal Commission, appointed in 1979 by the then prime minister of India, Morarjee Desai, identified a third group of *jatis* which were not former untouchables but were economically and educationally backward. These *jatis* were categorized as the Other Backward Classes or OBC and were included in the reservation policy of the country. The residual category is often called the general category or the "Others" to mean all the castes that are not included in the Scheduled Castes (SC), Scheduled Tribes (ST) or Other Backward Classes (OBC). The IHDS is the unique data set which divides the "Others" category further into *Brahmins* and Other Forward Castes (OFC) to separate the group at the very top of the caste hierarchy.