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

Tridip Ray† Arka Roy Chaudhuri‡ Komal Sahai±

February 18, 2020

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. One standard deviation increase in the years of education of the husband's mother is associated with a 10.16% increase in the probability of inter caste marriage over 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.

JEL codes: I21, J12, J16, Z13, D10

*We are grateful to Farzana Afridi, Sourav Bhattacharya, Diane Coffey, Sabyasachi Das, Ashwini Deshpande, Mukesh Eswaran, Kanika Mahajan, Ajit Mishra, Anirban Mitra, Anirban Mukherjee, Abhiroop Mukhopadhyay, Bharat Ramaswami, Ranjan Ray, Soham Sahoo, Sudipta Sarangi, E. Somanathan, Dean Spears and seminar participants at various places this paper was presented for their valuable comments and suggestions. We especially thank Sonalde Desai for her suggestions and for clearing several doubts regarding the IHDS data set. We also thank Laura Schechter and two anonymous referees of this journal for helpful suggestions. All remaining errors are ours.

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

‡Shiv Nadar Univeristy; arka.chaudhuri@snu.edu.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 (Davis 1941; Bisin and Verdier 2000). In the Indian context too, endogamy is central to the institution of caste.[1] Indian castes are largely endogamous groups and violations of caste endogamy are often punished by social ostracism (Chowdhry; Kaur 2010; Bidner and Eswaran 2015). It 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[2]. In this paper we study the relationship of caste endogamy with education, taking into account the nature of the Indian marriage market where marriages arranged by parents and close relatives is largely the norm.

Two aspects of the institution of caste highlight the importance of inter caste marriages. First, the caste system has been shown to be discriminatory (Shah 1985; Thorat and Newman 2007), and detrimental to democracy (Jeffrey 2002; Munshi 2017), social mobility (Munshi and Rosenzweig 2006; Munshi 2017), trade (Anderson 2011) and environment (Gadgil and Rao 1994). Second, caste endogamy being the pillar of caste system (Bidner and Eswaran 2015), inter caste marriages can directly weaken the foundations of caste system. In addition, though not directly for caste, there exists evidence of positive impact of intermarriages. For example, in-

[1] A huge body of literature has been developed to understand the origin, nature and contemporary aspects of the caste system in India. While it is too vast to be summarized here, see Srinivas (1962), Beteille (1971) and Dumont (1980), for some seminal works in this area. For excellent surveys of the literature, see, for example, Vaid (2014), Munshi (2017), Mosse (2018).

[2] Authors' calculations from the data set used for the study, the second round of the Indian Human Development Survey.

termarriages between natives and different immigrant ethnicities are associated with higher immigrant wages (Meng and Gregory 2005) and higher female labour supply (Gevrek et al. 2013; Wong 2014) in the context of Australia, Canada and the USA, respectively. Kalmijn (2010) shows strong evidence that interracial marriages have integrative effects on the offsprings for the case of the Netherlands. Positive effects of inter-ethnic marriages have also been shown on the social, cultural and economic integration of the children in England, Germany, the Netherlands and Sweden by Kalmijn (2015) and in two American cities by Stephan and Stephan (1991). Finally, extreme manifestation of endogamy in the form of consanguineous marriages may even be inefficient from the perspective of democracy and it may promote corruption and nepotism (Luke and Munshi 2006; Schulz 2019; Carl 2017; Akbari et al. 2019a,b).

A number of factors may influence the marriage choice of an individual. Since we are interested in looking at inter caste marriages in the particular context of weakening the institution of caste, we explore how education is associated with the probability of an inter caste marriage. Dr. Bhim Rao Ambedkar, the chief architect of the Constitution of India and one of the tallest leaders of the disadvantaged castes, was of the view that education could free the marginalized sections of the society and set them on a path of upward mobility (Velaskar 2012; Moon and Narke 2014a,b; Zene 2018). This spirit is incorporated as the primary focus of all education policies of India (National Policy on Education 1968, 1986; Right of Children to Free and Compulsory Education Act 2009; Joshee 2008; Mander and Prasad 2014). In addition, Dr. Ambedkar proposed that inter caste marriages will directly weaken the caste system (Ambedkar 1936). The same idea has been propounded by the Indian judiciary as well as policymakers (`https://www.timesnownews.com/mirror-now/in-focus/article/inter-caste-marriages-should-be-encouraged-for-`

uprooting-caste-system.-madras-high-court/440195; Ambedkar Scheme for Social Integration through Inter-Caste Marriages 2016).

In this paper, we aim to establish a link between education and inter caste marriages since education can not only mitigate deeply held prejudices, educational institutes can also serve as platforms for social mingling, especially since inclusive education has been a primary focus of the Indian education policy (National Policy on Education 1968, 1986; Right of Children to Free and Compulsory Education Act 2009; Joshee 2008; Mander and Prasad 2014).

A large part of the literature on out-marriages focuses on its relationship with the education of individuals. Qian (1997) and Fryer (2007) find a positive relationship between educational attainment and the likelihood of an interracial marriage in the US. 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 education and the likelihood of interracial marriages for whites.

Studies on exogamy in South Asia have been relatively scarce and primarily based on localized samples (Dugar et al. 2012; Banerjee et al. 2013; Allendorf and Thornton 2015; Ahuja and Ostermann 2016). To the best of our knowledge our paper is the first to make a systematic attempt at understanding the relationship of education with inter caste marriages in India using a nationally representative data set. But, at the outset, we recognize that we have to pay due attention to the fact that marriage markets in India work very differently as 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. In our data set (second

round of the Indian Human Development Survey, IHDS-II), 73% of marriages were reported to have been arranged by parents and almost 70% of the women said that they met their husbands only on the day of their wedding/*gauna*[3]. 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 even over time, the movement has not been towards "Western-style marriage, in which young people choose their own spouses" (Allendorf and Pandian 2016). The shift is rather towards increased say of women within the purview of "arranged marriages"[4] (Banerji et al. 2013; Allendorf and Pandian 2016).

The wide prevalence of the arranged marriage institution in the Indian marriage markets strongly suggests that any analysis of marriages in India must consider parental attributes along with individual ones. To justify this approach, 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 West, 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.

To examine our null results, we attempt at disentangling two potentially opposing effects of education identified by Furtado (2012). The first is the 'cultural adaptability effect' through which education makes members of different groups more aware

[3] *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.

[4] 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).

of and adaptable to the culture of each other and hence, may increase the incidence of intermarriage. The second one, the 'assortative matching effect'[5], however, may work in either direction. In a group with average education level below (above) the average education level of the relevant population, a more educated individual will marry out (marry in) and education will have a positive (negative) effect on exogamy for that group. The net effect can go in either direction and one may observe a positive, a negative or no relationship between education and exogamy depending on a particular group's characteristics. We adapt the methodology suggested by Furtado (2012) to the Indian context and our original findings are reaffirmed. None of the channels have any statistically significant association with the probability of inter caste marriages in India.

Our null results can mask important heterogenity across caste groups. According to the Status Exchange theory (Davis 1941; Merton 1941; Kalmijn 1998; Gullickson 2006; Fu and Heaton 2008)[6], in an inter caste marriage, the upper caste individual will typically be able to exchange her/his caste status for a higher level of education of a spouse from a lower caste as compared to the level of education of the spouse she/he would be matched to in an *intra* caste marriage. As a result, the marginal effect of an increase in education will be higher for a lower caste individual compared to a higher caste individual in the inter caste marriage market because education can be exchanged for caste status. We check for such heterogenity but find, very similar to Banerjee et al. (2013), no evidence of status exchange taking place.

[5]The term assortative matching refers to a positive correlation between the attributes of the husband and the wife. In our case, for example, the attribute is education.

[6]The status exchange theory broadly postulates that an intermarriage, especially between two groups which are unequally ranked in the social hierarchy, often involves an exchange of characteristics between the two parties such that both stand to gain from the union. Typically one party exchanges its social status for some other trait of the spouse, like beauty or education. Thus, more educated blacks would marry less educated whites because they would gain from a higher social status of their spouse (Gullickson 2006).

Having established the irrelevance of the spouses' own education, we next explore whether parental education is associated with the likelihood of an inter caste marriage. We add the education levels of the parents of both the spouses to our set of explanatory variables. Here we find that the level of education of the husband's mother has a positive and statistically significant association with the likelihood of an inter caste marriage. One standard deviation increase in the years of education of the husband's mother is associated with a 10.16% increase in the probability of inter caste marriage over the sample mean. The result is very robust to variations in the sample, to the addition of a number of controls as well as fixed effects, to alternate model specification and to omitted variable bias (Oster 2019). 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 a predictive power on the likelihood of inter caste marriage. Given our dataset, we are unable to empirically establish a precise channel for this finding. However, we posit some potential channels based on theoretical arguments from the existing literature and provide some suggestive evidence for our proposed mechanism.

The rest of the paper is organized as follows. In section 2, we describe the data. The descriptive analysis in Section 3 prepares the contextual background and provides the descriptive statistics. Sections 4, 5 and 6 detail the regression analysis, specifying the empirical strategy and discussing the main results and robustness checks, respectively. Section 7 gives a brief discussion of the possible channels behind the results and section 8 concludes.

2 Data

We use data from the latest round of the Indian Human Development Survey (IHDS II). 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.[7] 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. The household schedule 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.[8]

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 schedule: "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". This question accurately reflects whether a marriage is inter caste or not since the mar-

[7]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.

[8]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 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).

riage is recognized by the responding woman as inter caste and hence is "ultimately closer to the lived reality of an inter-caste marriage".[9,10]

Our main independent variables of interest are the years of education of the spouses and their respective parents. They range from 0 (illiterate) to 16 (above graduate) years. Our set of control variables include the caste and the urban or rural location (according to Census 2011) of the husband's household at the time of the survey. We include assets (index created by IHDS) and annual per capita income (in INR) 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. Finally, we control for the age at marriage of the bride and the comparative economic status of the two families at the time of their marriage.

We use three rounds of the Employment and Unemployment Survey of the National Sample Survey of India (NSS) conducted in 2004-05, 2009-10 and 2011-12 to construct average and caste-wise average years of education of females in the marriageable age group (12 to 35 years) for each district at the time of marriage.[11,12] We also calculate the proportion of population belonging to the same caste as that of a

[9]According to The Hindu (New Delhi, 13 November 2014) (Rukmini 2014), 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".

[10]Although in the English version the word caste is used, the question actually uses the word "*jati*" in Hindi (and its equivalents in all the other eleven languages in which the survey was administered), and not caste. This takes care of the fact that the finer *jati* level is relevant for marriages in India and not the caste level, which is often synonymous with the broad administrative categories in India.

[11]The marriageable age group is constructed by looking at the distribution of age at marriage of the eligible women in the IHDS sample where 96.8% of women report their age at marriage to be from 12 to 35 years.

[12]The nature of the NSS data and the fact that inter-district migration due to marriage is very low in India (Desai and Andrist 2010; Stopnitzky) helps us in constructing these variables at the time of marriage and not just for the NSS survey years.

husband in our sample in his district of residence using these NSS data sets. These variables are used to separate the opposing effects of education, namely, cultural adaptability and assortative matching effects.

3 Descriptive Analysis

Our specific aim in this paper is to look at the relationship between inter caste marriages and education. We set the stage by looking 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.

Figure 1 plots the rate of inter caste marriages by the year of marriage.[13] Even in the face of industrialization and urbanization in India, an upward trend is not visible over the last four decades: the rate of inter caste marriages has hovered around 5% since 1970 to 2012.[14] The average for 2000-2012 is marginally higher than 1971-80 and 1981-90, but is not statistically different from the decade of 1990-2000.

In Table A1 we look at the distribution of inter caste marriages by various characteristics of the husbands' households. 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.[15] However, the rate of exogamy for *Brahmins* is not statistically different from

[13]In IHDS II the year of marriage variable has 30.66% missing values. We, therefore, construct our own variable for the year of marriage using the year of birth of the eligible woman respondent and her age at marriage.

[14]The Modernization theory in Sociology explains the process of transition of a nation from a traditional political structure to a democratic one via causal chains of 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, which "... likely signals declines in the importance of ethnicity/caste, religion ..." (Allendorf and Pandian 2016).

[15]Refer to the Online Appendix for a description of the social and administrative categorizations

any other caste categories. The only significant differences are between the rates of OFCs and OBCs, and OFCs and SCs.[16]

The second panel of Table A1 shows that the rates of inter caste marriages are not statistically significantly different between urban and rural households. 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). Within the rural sector, developed villages have a higher rate, while less developed villages have a lower rate of inter caste marriages (5.72% and 4.86% respectively). Thus more urbanized areas do not necessarily have a higher rate of out-marriages in India.

The next two panels of Table A1 show the rate of inter caste marriages by asset and annual per capita income quartiles of the households respectively. In both cases the rate goes down as we move up the distribution (poorest to the richest): the rate of inter caste marriages is significantly higher in the first quartile than that in the fourth quartile. The last panel of Table A1 shows that no difference is observed in the rate of inter caste marriages 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.

In Table A2 we look at the decision making process at the time of marriage. The second column of Table A2 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 relatives) chose their

of the caste system in India.

[16] A reported inter caste marriage may not necessarily involve two broad administrative caste categories.

husbands, and in fact almost 70% of them met their husbands only on the day of their wedding/*gauna*. Only a quarter of the women had met their husbands or had seen their photos before marriage; even fewer had talked to their husbands before getting married to them (third panel of Table A2).

Even among the subset of only inter caste marriages, almost 63% of them are arranged by parents/other relatives only. Interestingly, even here 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 central in any analysis of marriages in India.

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.[17] Figure 2 shows that this rate is not statistically significantly different among the different educational categories of the spouses themselves.

From 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[18]. However, the rate of inter caste marriages appears to be significantly higher at higher educational categories of the mothers of the spouses[19]. This corroborates well with the earlier observation that parental attributes should be important in the analysis of marriages

[17]These categories are constructed by dividing the years of education into five bins: Illiterate (0 years), Up to Primary (1 to 5 years), Up to Secondary (6 to 10 years), Up to Bachelors (11 to 15 years) and Above Bachelors (more than 15 years).

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

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

in India where the institution of arranged marriages plays a dominant role. In what follows, we further explore along these directions in a regression analysis of the relationship between inter caste marriages and education.

4 Empirical Framework

Our observations in the previous section suggest that marriages in India are arranged primarily by parents with minimal say of the individuals themselves. Thus we must pay due attention to parental education along with the education of individual spouses.

We, therefore, 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:

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

Here $ICmarriage_{id}$ 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 and *wife's* edu_{id} is that attained by the wife.

In the next step 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.

In both equations (1) and (2), X_{id} is a vector of couple and household level control variables, namely, administrative 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).

Marriages in India occur overwhelmingly within the district (Desai and Andrist 2010; Stopnitzky). Therefore, we include district fixed effects, δ_d, to control for any time invariant unobserved factors at the level of a district. We also include year of marriage fixed effects, τ_t, to control for all unobservables across districts in the year a couple got married.

In our data set, households belonging to all religions report 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 which 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 of the Constitution of India. We also 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)[20] For our analysis we consider the 20 major states of India.[21] Our final sample consists of 25,070 couples of which 1079 couples have inter caste marriages. Standard errors are clustered at the Primary Sampling Unit (PSU) level. Table A3 provides the summary statistics for all the variables used in the regressions. All calculations use the survey weight of the eligible woman.

5 Results

5.1 Inter caste marriages and own education

Table A4 reports our first set of results. The first two columns report results from the estimation of equation (1). In column 1, the regression coefficients from the parsimonious specification with only caste controls and the education levels of the spouses show 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 the full set of our control variables. The addition of these controls has no effect on the coefficients of the spouses' own education – they remain statistically insignificant. This result stands in sharp contrast to the findings in the existing literature on out-marriages in the Western countries where individual's own education shows up as a predictor

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

[21]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 states of North-East, Goa and Jammu and Kashmir.

of one's marriage being within or outside one's race or ethnicity.

To investigate our null results, in the next three columns, we test whether any of the mechanisms of the effect of education as described in Furtado (2012) come out to be statistically significant. We adapt the model suggested by Furtado (2012) to the Indian context:

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

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.[22] The first term on the RHS, husband's years of education, captures the cultural adaptability effect of education. If the analysis of Furtado (2012) holds for our sample, this coefficient should be positive: an increase in education makes an individual more accepting and adaptable to the culture of other castes. The next term, $avg\ \ FemEdu_{cd}$ is the average education level of females in the marriageable age group (12 to 35 years) in the husband's caste in his district and $avg\ \ FemEdu_d$ is the average education level of *all* females in the marriageable age group in his district.[23]

The coefficient π_2 measures the assortative matching effect of education, which is captured by the interaction term of husband's years of education with the deviation of average education of females within his caste in the district from the average female education in the entire district.[24] The expected sign of π_2 is negative if the

[22]Since we do not know the caste of the wife in a couple, our sample consists of only husbands for this set of regressions.

[23] Both the variables, $avg\ \ FemEdu_{cd}$ and $avg\ \ FemEdu_d$, have been calculated at the district level and correspond to the relevant couple's year of marriage.

[24]The coefficient π_1 captures the main effect of this deviation.

assortative matching effect of education is at work. 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.

We also include the proportion of female population in the marriageable age group of husband's caste in his district, $pop\ \ pr_{cd}$, which captures the enclave effect: likelihood that the individual will encounter a potential spouse of the same caste in his relevant region of search, which we take to be the district based on the literature (Desai and Andrist 2010; Stopnitzky).[25]

Column 3 of Table A4 contains results from a regression similar to that of the second column, but uses only the husband's education variable (and replaces district fixed effects with state fixed effects) to make it comparable to the regressions in the next two columns. This coefficient, capturing the cultural adaptability effect, is still statistically insignificant. In column 4, we add the assortative matching term. The estimated coefficient of this variable is statistically insignificant and it also does not affect the coefficient of husband's own education.[26] Finally in column 5, we add the enclave effect term. The addition of this control and its square term too have no effect on insignificance of the coefficient of the husband's education variable. The coefficients on the variables themselves are also statistically insignificant.

Thus, even after we explicitly take into account the potential channels, as analyzed in Furtado (2012), through which own education might have an effect, we find that neither of these channels predict the likelihood of an inter caste marriage.

As noted in the introduction, our null results might mask important heterogenity across caste groups. According to the status exchange theory, one party exchanges

[25]For this set of regressions, we include state fixed effects, Ψ_s, instead of district fixed effects because our regressors are district level variables.

[26]We also calculate the education difference term by excluding husband's own caste females from the district average and use this variable in our regressions. All our results remain the same.

its social status for some other trait of the spouse, like beauty or education. Hence we might have a positive association between education and exogamy for some caste groups and negative for some other groups leading to a net null association between education and exogamy for all caste groups taken together. We run another set of regressions to check this but we find no evidence of status exchange[27]. Our result is very similar to Banerjee et al. (2013) who also find almost non-existent preference for "marrying up" or exchanging other attributes for caste status.

5.2 Inter caste marriages and parental education

Now we move on to 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 A4. Column 2 reports 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[28]. The addition of these variables has little effect on the coefficient of the husband's mother's

[27]Refer to the Online Appendix for a detailed discussion of the status exchange theory, our empirical specification and the regression results.

[28]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. In column 4, we add current income and assets of the household, and whether the household was located in an urban or rural area.

education.

We conclude this section with the key finding that husband's mother's education positively predicts the likelihood of an inter caste marriage and that it is robust to the inclusion of a number of controls and fixed effects. A one standard deviation increase in husband's mother's years of education leads to a 10.16% increase (over the sample mean) in the probability of the couple's marriage being an inter caste one. To put this in perspective, we compare the effect size of education on exogamy between the Indian and US data. Based on calculations made from Furtado (2012), we find that a one standard deviation increase in education of the husbands in her sample leads to only a 7.08% increase in his likelihood of inter ethnic marriage. A similar calculation shows that a one standard deviation increase in the husband's mother's years of education (in our data) explains 46.85% of the increase in the rate of inter caste marriages from 1970 to 2012.

6 Robustness checks

We report four robustness checks in Table 6. In column 1, we remove the women who continued their education post marriage as this could potentially contaminate the results since these women will actually have a lower amount of education at the time of their marriage as compared to what is measured by the data. All our results are qualitatively the same even for this sample.

It is plausible that if women had a greater say in their marriages, it may bias the coefficient on the education of the husband's mother upwards. A greater decision making power of the brides in their marriages may be positively correlated with both higher education of her husband's mother as well as with the probability of an inter caste marriage. Therefore, in column 2, 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". It can be seen that even here own education of the spouses has no association but the education of the husband's mother has a positive and statistically significant association with the probability of an inter caste marriage.

In the third column of Table 6, we add another set of fixed effects to our controls – the interaction of district and year of marriage fixed effects, to control for any unobservables at the level of a particular district-year. 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.

Since our dependent variable is binary, we report the estimation results from a logistic regression in the final column of Table 6. As can be seen, all our results go through with the logistic specification. The marginal effect of the husband's mother's years of education variable is 0.0025, while individual education coefficients are statistically insignificant as before.

We conduct another set of robustness checks to see if our results are robust to variations in the sample. We run our regressions for a Hindu-only sample, all-religions sample, all-castes (including STs) sample and all-states sample. We also use some other combinations: four main religions, main states, including STs; all religions, main states, including STs; and four main religions, all states, excluding STs. We find that our results are robust to all these sample variations. We report the first set of these regressions in the Online Appendix. The others are available upon request.

Although we do not claim any of our results to be causal, we still check if our results are being driven by unobservables. We examine the robustness of the result

to omitted variable bias using the bound analysis methodology developed by Oster (2019). Here again we deduce that the coefficient of the husband's mother's education variable is not contaminated by omitted variables bias. The details of the methodology and our results can be found in the Online Appendix.

Finally, since we test six simultaneous hypotheses (two education variables of the spouses and four of those of the parents), we also perform a series of multiple hypotheses corrections which control for Family-Wise Error Rate (FWER) as well as for False Discovery Rate (FDR). Our coefficient on the education of the husband's mother does not retain its statistical significance under these corrections. This is because given the structure of any of these tests, the lowest p value of the set is always corrected in a way similar to the Bonferroni correction (Farcomeni 2008) which is the most stringent correction (Abdi 2010; Fink et al. 2014; Streiner 2015). Since the lowest p value in our set is 0.033 (statistically significant at 5%), it is unable to retain significance under any of the available multiple hypotheses correction procedures.[29] While we present our results with this caveat, we nevertheless conclude this section with a reasonable confidence in the robustness of our results to variations in the sample, to the addition of a number of controls, to the addition of a number of fixed effects, to a change in the regression model and to correction for omitted variable bias.

7 Discussion

Although our results do not have a causal interpretation, they do point out some interesting features of the Indian marriage market. Our analysis of the relationship

[29]However, we would also like to point out that the idea of multiple hypothesis correction has its criticisms and the available methods might lead to too high rates of Type II error (Perneger 1998; Ruhm 2003; Nakagawa 2004; Kim et al. 2013).

between education and the age-old practice of caste endogamy 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. 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. All of our results survive a battery of robustness checks.

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 what follows we try to offer a plausible mechanism to explain our empirical findings regarding the heterogenity in the relationship between the parents' education and the probability of inter caste marriage with the caveat that we cannot offer any direct evidence because of a lack of data.

To understand the first result we put together three stylized facts. Firstly, a large body of literature finds evidence that a more educated woman has an increased decision making power in a household.[30] In our own dataset too, we find some support for this claim by looking at the responses to various questions under the "Gender Relations" section asked to the eligible women.[31] Interestingly, one of the questions directly asks about who has the most say in the decision to whom the respondent's

[30]See, for example, Thomas (1994); Beegle et al. (2001); Banerji (2008); Doss (2013); Banerji et al. (2013). Banerji (2008) and Banerji et al. (2013) use IHDS I to show that education is associated with greater autonomy in partner choice decision and it strongly improves the individual's involvement in parent arranged marriages.

[31]We find that the respondent woman's education is positively and statistically significantly associated with her likelihood of having the most say in seven out of the eight household decisions enquired in this section. The complete analysis can be found in the Online Appendix.

children should marry. We find that education of the respondent woman is positively and statistically significantly associated with the probability that she has the most say in this decision.

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. Provided with 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 and liquor.[32]

Finally, from our own analysis and from the literature cited in previous sections, we know that marriage decisions in India are taken by parents and other senior relatives and not by the prospective bride and groom.

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, the realized matches 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 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. This asymmetry between the two families might arise from the fact that in any inter caste marriage the bride's family bears more stigma or costs than the

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

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. 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[33]. 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 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.

[33]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*".

8 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, the education level of an individual does not predict the likelihood of his/her marriage being an inter caste one. In addition, we analyze if any of the possible channels suggested by Furtado (2012) is at work, but fail to find such evidence. We also see if there is any heterogenity in the relationship between education and exogamy across castes as suggested by status exchange theory, but do not find any such evidence. 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. Both 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. Our results also stand the scrutiny of a logistic regression model as well as omitted variable bias using the bound analysis (Oster 2019). 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.

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, rather it affects the marriage decision of one's offspring.

References

Abdi, H. (2010). Holms sequential Bonferroni procedure. *Encyclopedia of Research Design*, 1(8):1–8.

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.

Akbari, M., Bahrami-Rad, D., and Kimbrough, E. O. (2019a). Kinship, fractionalization and corruption. *Journal of Economic Behavior & Organization*, 166:493–528.

Akbari, M., Bahrami-Rad, D., Kimbrough, E. O., Romero, P., and Alhosseini, S. (2019b). An experimental study of kin and ethnic favoritism. *Available at SSRN 3442519*.

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.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of political economy*, 113(1):151–184.

Ambedkar, B. R. (1936). *Annihilation of Caste*. Bheema Patrika Publications, Jallandhar, India.

Ambedkar Scheme for Social Integration through Inter-Caste Marriages 2016. *Dr. Ambedkar Foundation Ministry of Social Justice and Empowerment*.

Anderson, S. (2011). Caste as an Impediment to Trade. *American Economic Journal: Applied Economics*, 3(1):239–63.

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. (2008). Is education associated with a transition towards autonomy in partner choice? A case study of India. Master's thesis, University of Maryland.

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), NCAER, New Delhi.*

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.

Bisin, A. and Verdier, T. (2000). Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988.

Carl, N. (2017). Ethnicity and electoral fraud in Britain. *Electoral Studies*, 50:128–136.

Chowdhry, P. Enforcing cultural codes: Gender and violence in northern India. *Economic and Political Weekly*, 32(19).

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

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, New Delhi.

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. *Working Paper No. 00-05, Department of Economics, MIT.*

Duflo, E. and Udry, C. (2004). Intrahousehold resource allocation in Côte d'Ivoire: Social norms, separate accounts and consumption choices. *NBER Working Paper No. 10498*.

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.

Farcomeni, A. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research*, 17(4):347–388.

Fink, G., McConnell, M., and Vollmer, S. (2014). Testing for heterogeneous treatment effects in experimental data: False discovery risks and correction procedures. *Journal of Development Effectiveness*, 6(1):44–57.

Friedberg, L. and Webb, A. (2006). Determinants and consequences of bargaining power in households. *NBER Working Paper No. 12367*.

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, X. and Heaton, T. B. (2008). Racial and educational homogamy: 1980 to 2000. *Sociological Perspectives*, 51(4):735–758.

Furtado, D. (2012). Human capital and interethnic marriage decisions. *Economic Inquiry*, 50(1):82–93.

Gadgil, M. and Rao, P. S. (1994). A system of positive incentives to conserve biodiversity. *Economic and Political Weekly*, 29(32):2103–2107.

Gevrek, E. Z., Gevrek, D., and Gupta, S. (2013). Culture, intermarriage, and immigrant women's labor supply. *International Migration*, 51(6):146–167.

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

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. (1995). The SES selectivity of interracially married Asians. *International Migration Review*, 29(2):469–491.

Jeffrey, C. (2002). Caste, Class, and Clientelism: A political economy of everyday corruption in rural North India. *Economic Geography*, 78(1):21–41.

Joshee, R. (2008). Citizenship education in India: From colonial subjugation to radical possibilities. *The Sage Handbook of Education for Citizenship and Democracy, California: Sage Publications*, pages 175–188.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*, 24(1):395–421.

Kalmijn, M. (2010). Consequences of racial intermarriage for children's social integration. *Sociological Perspectives*, 53(2):271–286.

Kalmijn, M. (2015). The children of intermarriage in four European countries: Implications for school achievement, social contacts, and cultural values. *The Annals of the American Academy of Political and Social Science*, 662(1):246–265.

Kaur, R. (2010). Khap panchayats, sex ratio and female agency. *Economic and Political Weekly*, 45(23):14–16.

Kim, C., Kim, M. K., Lee, C., Spector, J. M., and DeMeester, K. (2013). Teacher beliefs and technology integration. *Teaching and teacher education*, 29:76–85.

Luke, N. and Munshi, K. (2006). New roles for marriage in urban Africa: Kinship networks and the labor market in Kenya. *The Review of Economics and Statistics*, 88(2):264–282.

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*, 32(3):463–480.

Mander, H. and Prasad, G. (2014). India exclusion report 2013-14. *Books for Change, Bangalore, India. http://www. indianet. nl/pdf/IndiaExclusionReport2013-2014. pdf*.

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

Merton, R. K. (1941). Intermarriage and the social structure: Fact and theory. *Psychiatry*, 4(3):361–374.

Moon, V. and Narke, H. (2014a). *Dr. Babasaheb Ambedkar Writings and Speeches Vol. II*. Dr. Ambedkar Foundation, Ministry of Social Justice Empowerment, Govt. of India.

Moon, V. and Narke, H. (2014b). *Dr. Babasaheb Ambedkar Writings and Speeches Vol. III*. Dr. Ambedkar Foundation, Ministry of Social Justice Empowerment, Govt. of India.

Mosse, D. (2018). Caste and development: Contemporary perspectives on a structure of discrimination and advantage. *World Development*, 110:422–436.

Munshi, K. (2017). Caste and the Indian economy. *Cambridge Working Paper Series No. 1759*.

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4):1225–1252.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral ecology*, 15(6):1044–1045.

National Policy on Education, . (1968). *Government of India*.

National Policy on Education, . (1986). *Department of Education, Ministry of Human Resource Development, Government of India*.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ*, 316(7139):1236–1238.

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

Right of Children to Free and Compulsory Education Act, . (2009). *Legislative Department, Ministry of Law and Justice, India*.

Ruhm, C. J. (2003). Good times make you sick. *Journal of health economics*, 22(4):637–658.

Rukmini, S. (13 November, 2014). Just 5% of Indian marriages are inter caste: survey. *The Hindu. https://www.thehindu.com/data/just-5-per-cent-of-.indian-marriages-are-intercaste/article6591502.ece*.

Schulz, J. (2019). Kin-Networks and Institutional Development. *Available at SSRN: https://ssrn.com/abstract=2877828*.

Shah, G. (1985). Caste, class and reservation. *Economic and Political Weekly*, 20(3):132–136.

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

Stephan, W. G. and Stephan, C. W. (1991). Intermarriage: Effects on personality, adjustment, and intergroup relations in two samples of students. *Journal of Marriage and the Family*, 53(1):241–250.

Stopnitzky, Y. The bargaining power of missing women: Evidence from a sanitation campaign in India. *Available at SSRN 2031273.*

Streiner, D. L. (2015). Best (but oft-forgotten) practices: The multiple problems of multiplicity - whether and how to correct for many statistical tests. *The American Journal of Clinical Nutrition*, 102(4):721–728.

Thomas, D. (1990). Intra-household resource allocation: An inferential approach. *Journal of Human Resources*, 25(4):635–664.

Thomas, D. (1994). Like father, like son; like mother, like daughter: Parental resources and child height. *Journal of Human Resources*, 29(4):950–988.

Thorat, S. and Newman, K. S. (2007). Caste and economic discrimination: Causes, consequences and remedies. *Economic and Political Weekly*, 42(41):4121–4124.

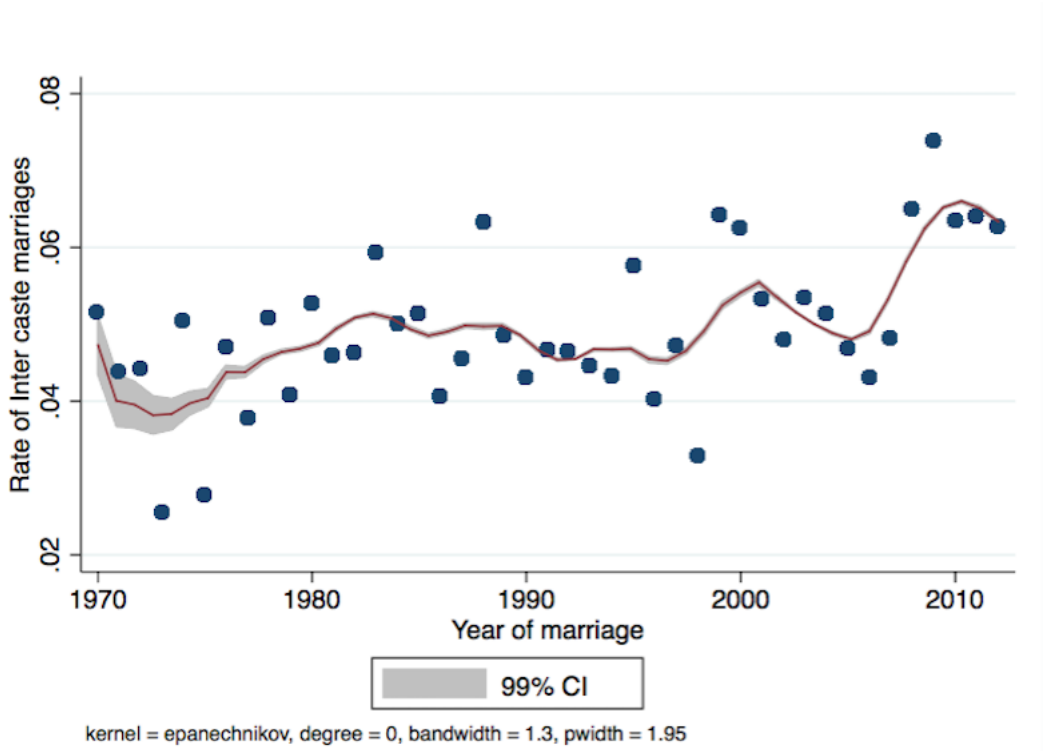
Vaid, D. (2014). Caste in contemporary India: Flexibility and persistence. *Annual Review of Sociology*, 40:391–410.

Velaskar, P. (2012). Education for liberation: Ambedkars thought and Dalit women's perspectives. *Contemporary Education Dialogue*, 9(2):245–271.

Wong, H.-P. C. (2014). The effects of Endogamous Marriage on Family Outcomes: Evidence from Exogenous Variation in Immigrant Flows During 1900-1930 in the United States. Working paper,, Department of Economics, West Virginia University.

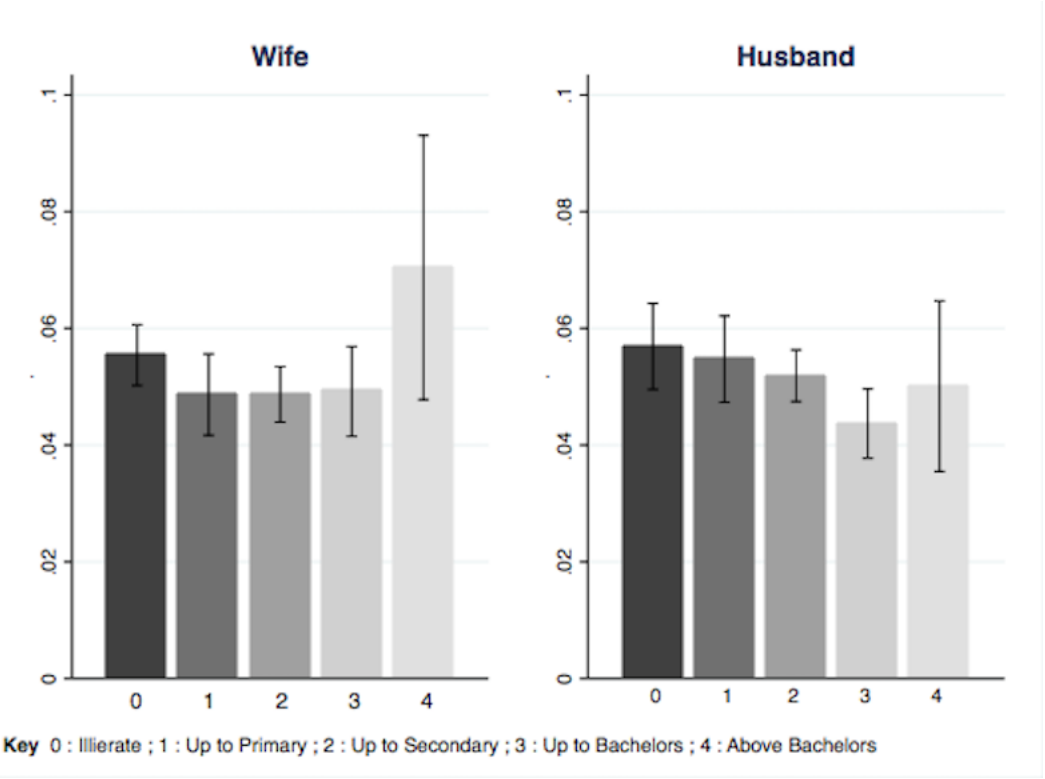
Zene, C. (2018). Justice for the Excluded and Education for Democracy in BR Ambedkar and A. Gramsci. *Rethinking Marxism*, 30(4):494–524.

Figure 1: Trend in the rate of inter caste marriages



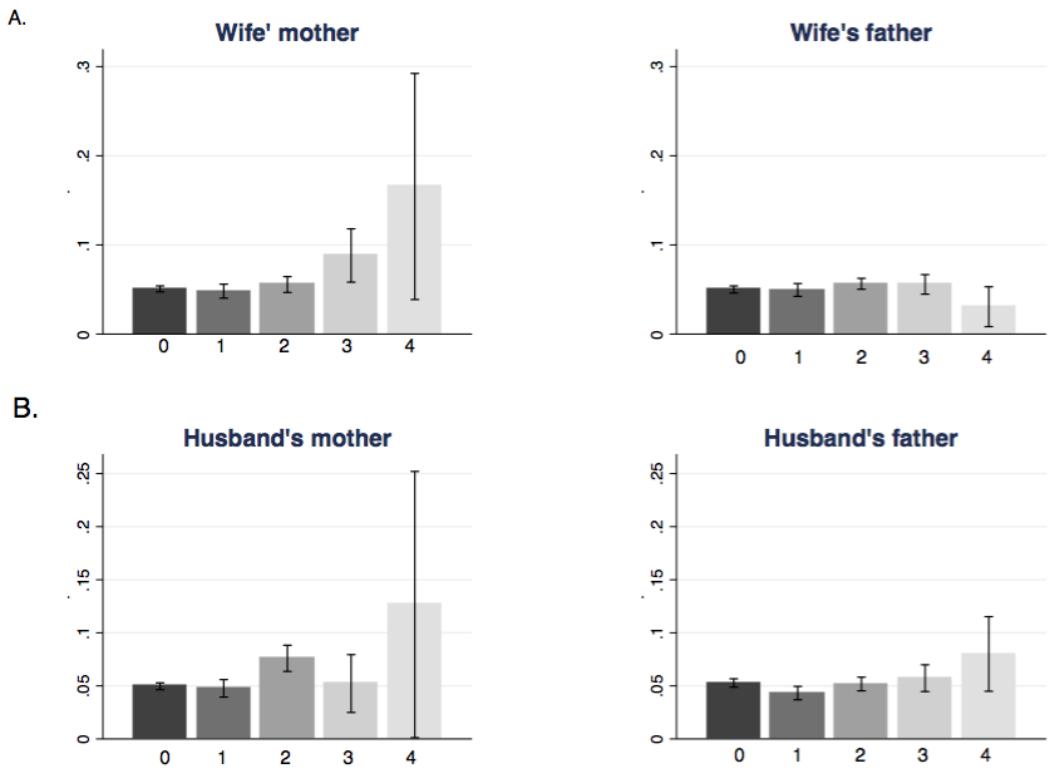
Note: The smooth line plots the local polynomial regression of the yearly percentage of inter caste marriages on the year of marriage. Data source is IHDS II.

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 (poorest)	5.89***
	(0.317)
Second quartile	5.48***
	(0.318)
Third quartile	5.01***
	(0.273)
Fourth quartile (richest)	4.01***
	(0.266)
Income quartiles	
First quartile (poorest)	5.08***
	(0.337)
Second quartile	5.58***
	(0.312)
Third quartile	4.07***
	(0.259)
Fourth quartile (richest)	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: Standard errors in parenthesis. Data source is IHDS II. The household here corresponds to the husband's household. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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

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
husband's edu	-0.000351	-0.000102	-0.000429	-0.000454	-0.000413
	(0.000543)	(0.000520)	(0.000767)	(0.000782)	(0.000780)
wife's edu	-0.000776	-0.000546			
	(0.000839)	(0.000814)			
(avg $FemEdu_{cd}-$ avg $FemEdu_d$)				-0.00184	-0.00176
				(0.00272)	(0.00276)
husband's edu* (avg $FemEdu_{cd}-$ $avg\ FemEdu_d$)				-0.000130	-0.000158
				(0.000226)	(0.000231)
pop pr					0.0969
					(0.0765)
pop pr sq					-0.159
					(0.100)
Controls I		√	√	√	√
Controls II		√	√	√	√
Caste controls	√	√	√	√	√
Year of marriage FE	√	√	√	√	√
District FE	√	√			
State FE			√	√	√
N	22476	22469	22470	22027	22027
R^2	0.221	0.222	0.033	0.034	0.035

Note: Linear probability results are reported. Data sources are IHDS-II and Schedule 10 of NSS Rounds 61 (2004-05), 66 (2009-10) and 68 (2011-12). Outcome is a dummy variable which takes value 1 if the marriage is inter caste, 0 otherwise. The term *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 $FemEdu_{cd}$ − avg $FemEdu_d$)* 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 and *husband's edu*(avg $FemEdu_{cd}$ − avg $FemEdu_d$)* 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
husband'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 sample of women and inclusion of interaction fixed effects

	(1) Completed education before marriage	(2) Only arranged marriages	(3) District*Year of marriage FE	(4) Logit
husband's edu	-0.000197	-0.000167	-0.000675	0.000441
	(0.000538)	(0.000639)	(0.000735)	(0.0145)
wife's edu	-0.000979	-0.000659	-0.000100	-0.0206
	(0.000875)	(0.000645)	(0.000829)	(0.0202)
husband's mother's edu	0.00226**	0.00210**	0.00220*	0.0471**
	(0.000951)	(0.000862)	(0.00123)	(0.0184)
husband's father's edu	-0.000878	-0.00110	-0.00103	-0.0208
	(0.000670)	(0.000704)	(0.000790)	(0.0168)
wife's mother's edu	0.000978	0.000313	0.000633	0.0204
	(0.000968)	(0.000736)	(0.00109)	(0.0200)
wife's father's edu	0.000264	0.000696	0.000550	0.0127
	(0.000539)	(0.000499)	(0.000684)	(0.0128)
Controls I	√	√	√	√
Controls II	√	√	√	√
Caste controls	√	√	√	√
Year of marriage FE	√	√	√	√
District FE	√	√	√	√
District*Year of marriage FE			√	
N	21269	16439	22244	16089
R^2	0.229	0.339	0.549	-

Note: Linear probability results are reported in columns 1 to 3. Logit results are reported in column 4. 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. Column 4 uses a logistic regression specification. 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.

Appendix

A *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 affirmative action policies were being formulated, *jatis* which were economically the weakest and were historically subjected to discrimination and de-

privation, the so called "untouchables", were identified in a government schedule as the target group for reservation policies (Deshpande, 2011). These *jatis* are referred to as the Scheduled Castes (SC). Another government schedule identified similarly placed tribes and tribal communities 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, Morarji Desai, recommended that the reservation policy be extended to a third group of *jatis* which were not former untouchables but were economically and educationally backward. These *jatis* are categorized as the Other Backward Classes or OBC. 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.

B Inter caste marriages and own education: Status exchange?

Status exchange theory argues that in a hierarchical society, out-marriage often involves an exchange of characteristics by the two parties. Typically one party exchanges its social status for some other trait of the spouse, like beauty or education. Thus, more educated blacks would marry less educated whites because they would gain from a higher social status of their spouse (Gullickson 2006). In the Indian context, since there are more than two castes, there can potentially be many types of higher caste-lower caste unions. Hence our null results might mask interesting

heterogenity across castes in the relationship between education and exogamy.

To test for the above possibility, we use the following specification:

$$\begin{aligned} ICmarriage_{id} = \gamma_0 &+ \beta_0.husband's\ caste_{id} + \beta_1.husband's\ edu_{id} \\ &+ \beta_2.husband's\ edu_{id}.BR_{id} + \beta_3.husband's\ edu_{id}.OBC_{id} \\ &+ \beta_4.husband's\ edu_{id}.SC_{id} + \alpha_0.wife's\ edu_{id} \\ &+ \alpha_1.wife's\ edu_{id}.BR_{id} + \alpha_2.wife's\ edu_{id}.OBC_{id} \\ &+ \alpha_1.wife's\ edu_{id}.SC_{id} + \sigma.X_{id} + \delta_d + \tau_t + \xi_{id}. \end{aligned} \tag{4}$$

Here husband's caste are caste dummies for Brahmin, OBC and SC.[34] The omitted category is OFC. If status exchange takes place, it implies that compared to OFC, the marginal effect of an increase in education will be higher for OBC and SC which are ranked lower than the OFC in the caste hierarchy, while it will be lower for Brahmins who are ranked above OFC. Thus we expect β_2 to be negative while β_3 and β_4 to be positive.

The opposite will hold true for the wife because we are interacting her education with her husband's caste. Therefore, the marginal effect of an increase in the education of the wife will be positive when the husband's caste is higher than OFC (the omitted category), whereas it will be negative when the husband's caste is lower than OFC. Thus we expect α_1 to be positive while α_2 and α_3 to be negative.

The results are reported in Table A1. The first column reports coefficients when year of marriage fixed effects and district fixed effects are not included. It can be seen that none of the coefficients are statistically significant here. It implies that education does not *differentially* improve the chances of an inter caste marriage for

[34] As mentioned earlier, we use only husband's caste in our specification since we do not know the caste of the wife in a couple.

any caste as compared to the OFCs, the omitted category. In columns 2 to 4, controls are successively added to the base specification. The results do not change: none of the coefficients are statistically significant in any of the columns.

C Robustness checks: Variations in religion and caste composition of samples

In this set of robustness checks, we test whether our results withstand variations in the sample which consists of only those households which have stated their religion as Hinduism, Buddhism, Jainism or Sikhism residing in the 20 major states in India, and excludes scheduled tribes. Table A2 shows the results for four such sample variations. Since caste system, as mentioned above, is theoretically a Hinduism phenomenon, in the first column in Table A2, 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. The results reported in all the columns are qualitatively similar to those in the main regression: the education of the spouses 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[35].

[35]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.

D Robustness check: Bound Analysis (Oster 2019)

We conduct the bound analysis to check the robustness of our coefficient on the education of the husband's mother variable to omitted variable bias. In this section we first describe the methodology briefly, and then report our results.

It is a common practice to infer about the robustness of a result to omitted variable bias by looking at coefficient movements upon the addition of controls. Oster (2019) argues that to use observables to estimate bias from unobservables, we must (a) invoke the assumption of related covariance, that is, we need to assume that the unobservables positively covary with the observables so that the observables are informative about the unobservables, *and* (b) scale the coefficient movements by movements in R^2. Building on Altonji et al. (2005) and using the assumption of related covariance, she explicitly links coefficient movements, R^2 movements and omitted variable bias. In particular, she assumes a proportional selection relationship between the observables and unobservables, and denotes this coefficient of proportionality by δ. Thus, δ essentially captures the relative strength of unobservable selection to observable selection. Using this assumption, one can calculate the bias adjusted value of the coefficient of interest, assuming a value for δ and a value for the R^2 in the hypothetical regression which controls for both observables and unobservables (R^2_{max}). If unobservables are as important as observables, then $\delta = 1$. Oster (2019) suggests that this is a reasonable upper bound for δ, that is, unobservables should not be more important than the observables in explaining the dependent variable. An upper bound for R^2_{max} is equal to 1 for the case when all of the variation in the dependent variable is explained by the observables and unobservables combined.

However, this may often not be the case and therefore, $R^2_{max} = min(\pi.R^2_{controlled}, 1)$ is a suggested function to arrive at an upper bound for R^2, where $R^2_{controlled}$ is the R^2 from the regression including all the observable controls and π is a multiplier[36]. This exercise will give a bias adjusted value of the coefficient of interest which can then be compared to the value of the coefficient in the controlled regression.

Assuming an appropriate bounding value for R^2, or R^2_{max}, one can also calculate the value of δ which renders the coefficient of interest zero. A value of δ greater than 1 would suggest a robust coefficient.

We carry out the bound analysis in both the ways as suggested by Oster (2019). The results are reported in Table A3. Using similar terminology, δ captures the relative importance of the unobservables with respect to the observables, and β is the coefficient of our variable of interest, that is, the years of education of the husband's mother. We first report the bias adjusted β under the assumption δ=1 and an upper bound for R^2_{max}[37]. We use the function $R^2_{max} = min(\pi.R^2_{controlled}, 1)$ and set $\pi = 1.3$ (Oster 2019). This translates to $R^2_{max} = 0.2912$ and the corresponding $\beta = 0.00159$. The bias adjusted coefficient has the same direction as that reported in our analysis up till now. We then calculate the value for δ if β were to equal zero, with the same assumption on R^2_{max}. As can be seen from the table, the value of δ comes out to be equal to 2.65. This suggests that for β to actually be statistically insignificantly different from zero, the unobservables must be almost three times as important as the observables. Since this seems unlikely to be the case, we deduce that our result is robust.

[36]Oster (2019) applies this adjustment to a host of studies, both randomized and non-randomized, as well as to constructed data to see whether the results survive. Taking randomized experiment results as the benchmark, she suggests $\pi = 1.3$ as the cutoff value at which at least 90% of the randomized results survive. We use this value of π for our tests.

[37]The term R^2max is the R^2 in the hypothetical regression which controls for both observables and unobservables.

Finally, we also report the 95% confidence interval for the controlled β and check if the "identified set" (bounded on one side by the uncontrolled regression coefficient and on the other side by the bias adjusted coefficient of interest) lies within the confidence interval. The last two rows of Table A3 show that the identified set indeed falls within the 95% confidence interval of the controlled β. This lends further proof that the coefficient of the husband's mother's education variable is not contaminated by omitted variables bias.

E Education and Female Autonomy

In this section we discuss the link between the decision making power of women and their education level using the responses to various questions under the "Gender Relations" section in the eligible woman's questionnaire in IHDS II. We run the following linear probability model:

$$Autonomy_{id} = \beta_0 + \beta_1.own\ \ edu_{id} + \theta.X_{id} + \delta_d + \tau_t + \varepsilon_{id}. \tag{5}$$

Here $Autonomy_{id}$ is a dummy variable which takes value 1 when an eligible woman i in district d has the *most say* in a particular household decision. The "Gender Relations" section asks eight such questions. The variable of interest is *own edu* which is the years of education of the respondent woman. A host of control variables are included in X. Apart from the standard control variables used in this research – caste category of the household, comparative economic status of the woman's natal family and her husband's family at the time of their marriage, per capita income and asset index of the household and its rural or urban location, X also includes the age of the respondent as well as the years of education of her husband and both parents

in law. We include district fixed effects, δ_d, and year of marriage fixed effects, τ_t, in all the regressions. The results are reported in the two panels of Table A4.

In Panel A, we can see that education of the eligible woman is positively correlated to her having the most say in whether to buy an expensive item (column 2), how many children to have (column 3) and what to do if she falls sick (column 4). We see the same positive association in Panel B in decisions on whether to buy land or property (column 1), how much money to spend on a social function (column 2), what to do if her child falls sick (column 3) and, most importantly for our analysis, to whom should her children marry (column 4).

Table A1: Own education: Status exchange?

	(1) ICmarriage	(2) ICmarriage	(3) ICmarriage	(4) ICmarriage
husband's edu (β_1)	-0.000367	-0.000109	-0.0000707	0.000107
	(0.00237)	(0.00183)	(0.00182)	(0.00187)
husband's edu*BR (β_2)	0.00457	0.00211	0.00210	0.00213
	(0.00388)	(0.00236)	(0.00236)	(0.00235)
husband's edu*OBC (β_3)	-0.0000651	-0.000123	-0.000159	-0.0000818
	(0.00289)	(0.00237)	(0.00239)	(0.00237)
husband's edu*SC (β_4)	-0.00157	-0.000915	-0.000943	-0.000884
	(0.00265)	(0.00207)	(0.00207)	(0.00207)
wife's edu (α_0)	-0.00117	-0.00202	-0.00192	-0.00180
	(0.00304)	(0.00246)	(0.00250)	(0.00242)
wife's edu*BR (α_1)	0.00109	0.00104	0.00109	0.000953
	(0.00378)	(0.00293)	(0.00294)	(0.00291)
wife's edu*OBC (α_2)	0.0000876	0.00152	0.00149	0.00154
	(0.00323)	(0.00230)	(0.00232)	(0.00232)
wife's edu*SC (α_3)	0.00286	0.00168	0.00162	0.00170
	(0.00318)	(0.00241)	(0.00241)	(0.00241)
Controls I			√	√
Controls II				√
Caste controls	√	√	√	√
Year of marriage FE		√	√	√
District FE		√	√	√
N	22476	22476	22476	22476
R^2	0.002	0.221	0.221	0.222

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. The Greek letters specified in paranthesis in the first column against the variable names correspond to the notations used in equation 4. Column 1 does not include year of marriage fixed effects and district fixed effects. 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 A2: Robustness checks: Variations in religion 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
husband'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 A3: Bound analysis

Uncontrolled β	0.00221
(R^2)	(0.001)
Controlled β	0.00186
(R^2)	(0.224)
β **for** $\delta = 1$, $R^2_{max} = 0.2912$	0.00159
δ **for** $\beta = 0$, $R^2_{max} = 0.2912$	2.65
Identified set	[0.00159, 0.00186]
95% Confidence interval	[.0001462, 0.003575]

Note: Bound analysis results are reported. Here $R^2_{max} = min(\pi.R^2_{controlled}, 1)$, $\pi = 1.3$. The Uncontrolled regression controls only for the education of the husband's mother, the controlled regression includes the full set of education variables and control variables. Data source is IHDS II.

Table A4: Education and female autonomy

Panel A				
	(1) What to cook on a daily basis	(2) To buy an expensive item	(3) How many children should she have	(4) What to do if she falls sick
own edu	-0.00121	0.00207***	0.00489***	0.00172*
	(0.00134)	(0.000667)	(0.00118)	(0.000912)
Panel B				
	(1) To buy land or property	(2) How much money to spend on a social function	(3) What to do if her child falls sick	(4) To whom her children should marry
own edu	0.00181***	0.00342***	0.00505***	0.00168**
	(0.000590)	(0.000864)	(0.00109)	(0.000760)
Controls I	√	√	√	√
Controls II	√	√	√	√
Caste controls	√	√	√	√
Year of marriage FE	√	√	√	√
District FE	√	√	√	√

Note: Linear probability results are reported. Data source is IHDS-II. Outcome is a dummy variable which takes value 1 if the eligible woman has the most say in that household decision. Controls I consists of age of the respondent and years of education of her husband, her mother in law and her father in law. Controls II consists of the economic status of the wife's natal family as compared to the husband's family at the time of marriage, 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 parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions weighted by survey weight of the individual.