

# Affirmative Action, Caste and the Marriage Market

Kunjai Desai

13th August, 2016

## Abstract

This study investigates the role of a strong social norm - caste - on the marriage market in India and how it responds to economic changes. A jobs-based affirmative action policy, the implementation of the Mandal Commission report in December 1993, differentially affects incomes across cohorts for treated castes. This is used to separately identify the effects of caste and income on marital partner choice and the distribution of the match surplus using a difference-in-differences strategy. The affirmative action policy is found to have an asymmetric effect across genders - male members of the treated group are found to attract more desirable spouses and get a higher share of the marital match surplus while there is no similar effect for women. A structural model of the marriage market based on Choo & Siow(2006) is used to investigate the marital welfare effects of the policy while controlling for selection effects. Similar results are found. These findings suggest that the marriage market and intra household decision making plays a significant role in determining the distribution of welfare changes from economic shocks.

1

---

<sup>1</sup>Department of Economics, Columbia University. Contact: kd2380@columbia.edu. I am grateful to Suresh Naidu, Pierre-Andre Chiappori and Bernard Salanie for guidance, and to Cristian Pop-Eleches, Eric Verhoogen, Jonas Hjort, Supreet Kaur, Ritam Chaurey, Prabhat Barnwal, Ashna Arora and Siddharth Hari and participants of the Development Colloquium at Columbia University for comments. All errors are my own.

**Preliminary work. Please do not cite or circulate without permission.**

# 1 Introduction

Caste is an old and persistent institution of Indian society. The origins of the caste system were occupational and can be seen in caste names to this day. Despite being social in nature, it is found to have an impact on a large set of economic decisions. Apart from occupation choice, caste is also found to affect public goods allocations(Banerjee and Somanathan(2006)), capital allocations(Banerjee and Munshi(2004)), migration decisions(Munshi and Roszentsweig(2016)) and marital partner choice in India(Banerjee et al(2009), Dugar et al(2012) and Rosenzweig and Stark(1989)) .

There is mixed evidence on whether the institution of caste has been weakening in recent times. Some evidence shows that the role of caste in determining education and job outcomes, though still significant, has been decreasing over time(Hnatkowska, Lahiri and Paul(2012), Munshi and Rosenzweig(2006)). However, caste has also become highly salient in Indian politics since the 1980s(and since the 1960s in southern India), which saw the rise of exclusively caste based political parties competing in elections. These parties have, moreover, been highly successful and have replaced the national parties as the main political force in a large part of the country, including the populous states of Bihar, Uttar Pradesh and Tamil Nadu.

The marriage market is the mechanism through which the caste system propagates. It is therefore, the key to understanding if the social institution will persist alongside economic development. Even today, a vast majority of marriages still take place within one's own group(NFHS-III). This is not just an equilibrium outcome. A recent opinion poll shows 74% of people were opposed to inter caste marriages(Banerjee et al(2009)).

Besides its role in propagation of caste, one would be interested in the marriage market for several other reasons. The traditional literature on the marriage market has focused on developed western countries and studies how people sort based on attributes like education, income, BMI(Chiappori et al(2012)), attitudes towards risk(Dupuy and Galichon(2014)) and age(Choo and Siow(2006)). Tastes for caste themselves are another attribute that would affect matching patterns. These might have implications on how people sort themselves based on education or other economic attributes. For example, consider the case where there

are two communities. Community A has highly educated men and uneducated women while community B has highly educated women but uneducated men. Now, if communities have no preference for marrying within each other and sort assortatively solely on education, the highly educated men and women pair off and the uneducated men and women pair off with each other. However, if communities have a strong social norm that prevents them from intermarrying, then in equilibrium, the only marriages that are observed are between one uneducated member and one highly educated one. This example is extreme but captures the intuition of the impact of a social norm for endogamy. One impact that this might have is on inequality. Another depends on the nature of the household production function. If household production is supermodular in education of spouses, there is an aggregate surplus loss from pairing individuals with different levels of education. A third effect relates solely to human capital investments. If one considers part of the returns of human capital investments to be the partner that one can attract on the marriage market, like in Chiappori, Iyigun and Weiss(2009) and Lafortune(2013), tastes for caste might explain differential rates for returns to education across different communities.

This study seeks to investigate how income changes affect the marriage market in India. This is done by exploiting a large affirmative action programme implemented in India during the 1990s. This followed the implementation of the Mandal Commission report in December 1993. The report recommended reserving 27% of all government jobs(including in public sector undertakings) for a group of socially and economically backward castes, the Other Backward Castes(OBCs). These were different from the historically most backward groups - the Dalits and the adivasis, that had their own reservations in government jobs since the 1950s under the Scheduled Caste(SC) and Scheduled Tribes(ST) categories. Government jobs in India pay a significant wage premium over comparable private sector jobs of approximately \$1000, which is 60% of India's current per capita GDP. As a result, these jobs are highly sought after and entry is extremely competitive. A specific quota for OBCs would exogenously increase their incomes. In addition, government jobs in India have strict age caps. For example, one needs to be under 33 years old to be eligible to apply for the Indian Administrative Services if one belongs to one of the economically and socially backward groups and 27 years old otherwise. This restricts the impact of the affirmative action program to cohorts under the

age of 33 when the program went into effect. All members of the OBC group that were older than 33 in 1994 would not be able to benefit from the program. This program design results in a difference in differences setting, OBCs would see a differing cohort profile of government employment and hence, income, compared to the other caste groups. This is observed in Figure 1.

The marriage market effects might take three different forms, (1) the probability of getting married, (2) the choice of partner and (3) the distribution of the marital surplus among partners. While there is some work studying these questions individually, this is the first study to my knowledge tying all three together in India. For example, using responses to matrimonial advertisements, Banerjee et al(2011) show that grooms are willing to trade off the difference between a self described “very beautiful” woman and a “decent looking” one to obtain brides in the same group as their’s while potential brides are willing to trade off the difference between a master’s degree and no education in a groom to marry within the same caste. The authors also find that these caste preferences don’t create a friction(or have a zero price) in the marriage market because male and female populations are largely balanced within different caste groups with respect to characteristics that the other gender find attractive. Dugar et al(2012) use a random assignment of caste to marital profiles to obtain similar results, that there is a tradeoff between income and caste of prospective spouses.

Both of these prior studies differ from this one in several other ways. First, they observe responses to matrimonial advertisements and construct preference measures using a revealed preference approach but do not observe the full set of equilibrium matches, which leaves them vulnerable to strategic manipulation. The approach here is different, preference parameters are obtained by observing the equilibrium set of matches. Second, both the earlier studies are conducted in a very specific setting - among highly educated people in a major metropolitan city. The samples studied are not representative of India, and the outcomes of matching exercises are highly sensitive to the studied population. The sample used in this study is nationally representative and has a substantial rural portion. Third, both of these studies focus on preferences at a single cross section in time. By using historic marriage data, this study investigates how preferences interact with changing population characteristics to affect

the final equilibrium match. This also allows for the identification of elasticities of endogenous quantities, for example the rate of intercaste marriage, with respect to incomes.

The distribution of the match surplus with income is well studied in the intrahousehold bargaining literature. Bobonis(2009) is an example, where randomized grants to women are shown to increase their household bargaining power. The relative contribution of this study to the intrahousehold decision making literature is to take into account the fact that couples are formed endogenously, which might affect the decision making power within households.

A structural model of the Indian marriage market is also constructed. This approach helps identify a shadow price for agents with different caste identities and education levels, giving an idea of how the marital returns to caste affect returns on human capital investment. One can also disentangle the source of the changes in observed surpluses and if they are driven by changes in the household surplus generation function or by changes in identity based occupational frictions. The source of variation has important policy consequences, the latter would call for focusing on the labour market(like jobs and education quotas) while the former is harder to solve.

The framework used in the structural model, that of Choo and Siow(2006) is motivated by the data and institutions particular to the Indian marriage market. Matching patterns are observed - who marries whom and who stays single and intra couple allocations. One must make an assumption on the nature of the household production process. Utility within households is assumed to be perfectly transferable. The widespread prevalence of pre marital transfers - dowries - in India makes this assumption highly plausible. Even if intrahousehold utility is non-transferable(or even imperfectly transferable) ex post, pre marital transfers allow for individuals to be compensated for post marriage expected utility differences. Information on these transfers is, however, extremely limited.

The affirmative action policy is found to have asymmetric effects. Treated men of the OBC community, that were more likely to work outside the household due to social norms, observe a positive marital return. While their probability of getting married does not change, they match with more attractive spouses in terms of education , height and BMI. Further, they capture a higher share of the resulting household surplus as seen in fertility decisions. Jewelry consumption and the practice of purdah do not change. There is no differential marriage

market effect on OBC women. Their spouses are not different as a result of treatment. Their workforce participation rate also drops. These findings suggest that the distributional welfare effects of affirmative action policies need to be accounted for when doing a cost-benefit analysis.

The rest of the paper is organized as follows. Section 2 discusses the data, Section 3 the reduced form empirical strategy. Section 4 discusses reduced form results. Section 5 presents the structural model and Section 6 concludes.

## 2 Data

The ideal dataset would have 6 characteristics. These are (1) spouses would be identifiable within households, (2) the year of marriage or age at marriage for each couple would be known, (3) the caste or caste group for each member of the couple should be observable, (4) the consumption of each member should be observable, (5) information on parental background for each member should be observable, and (6) incomes and dowries at the time of marriage should be observable. No dataset satisfies all of these characteristics. Therefore, several datasets are used, each of these has one or more drawbacks. Results from these are aggregated to get an idea of the relevant factors.

Data on education, wages and employer type are obtained from the 10th Schedule of the 55th, 61st and 68th “thick” rounds of the National Sample Survey. These are the largest repeated cross sections that are available for India. These are also the earliest surveys that recorded the OBC caste group, since 1999. The drawback of this dataset is that it only records the caste group of the household head, not individual members and only records the relationship of people to the household head so that it is often not possible to match all spouses within a household to each other.

The 3rd round of the National Family Health Survey(NFHS) is used to study spousal attributes and estimate the decision making power within couples. This is the Demographic and Health Survey(DHS) counterpart for India. It is a nationally representative survey of 109,041 households that was conducted in 2006. Information on age, years of education and

marital status is collected for each household member along with the caste of the household head. More detailed information is collected for married couples where the age of the wife is 15-49 years and where that of the husband is 18-55 years. Crucially, the broad caste group of both husband and wife are recorded in addition to their age, years of education, year of marriage, current marital status and physical attributes like height and weight. There are 39,257 such couples. Information on decision making responsibilities within the household is collected. Also present is the ideal number of children that each partner desires and the actual number of children that a couple has. The drawback of this dataset is that it does not report on economic variables like labor force participation and does not report employer types.

Data on matched spouses is also obtained from the 2011-12 round of the India Human Development Survey(IHDS). This is the the second round of a nationally representative panel survey, the first round of which was in 2004-05. It consists of 42,152 households. While spouses are matched and economic variables are reported, this datasets drawbacks are that only the caste of the household head and household level consumption is reported for jewellery and the practice of the custom of purdah. The year of marriage is also not reported for most couples.

The 1999 round of the Rural Economic Development Survey(REDS), a nationally representative survey of rural areas, is used to obtain information on transfers at the time of marriage, parental background and the caste of both spouses. The drawback of this study is that it does not report an employer type and it's small sample size so that the treatment group consists of only 18 observations.

Since the object of investigation is the set of stable marital matches, cohorts where either the husband is of age less than 31 or the wife less than age 26 are dropped in all cases. There are very few marriages above these ages for both genders across the study period(more than 95% of marriages had occurred by 3 years before each cutoff). Similarly, couples where either member is greater than 60 years old are dropped since this would introduce a survival bias and workforce participation falls dramatically at age 60. Additionally, members of couples that are divorced, separated, widowed or married more than once are also dropped since the decision to get divorced, separated, or remarried could be strategic. This decision is

difficult to model in the theoretical framework used and hence outside the scope of this study. Fortunately, this is a small part of the sample.

Education attainment is divided into four categories -

- Illiterate - 0 years of education
- Below Primary - 1 to 4 years of education
- Primary - 5-9 years of education
- Secondary and Above - 10+ years of education

Two definitions of treatment are used depending on if the year of marriage is observable. If it is, then treatment comprises of marriages that occurred in or after 1994 and consisted of at least one member that belonged to the OBC category. If the year of marriage is not observed, then the treated group consists of OBC cohorts that were less than 33 years old in 1994.

### 3 Empirical Strategy

The reduced form results are obtained using a difference in differences strategy. OBCs were eligible for quotas in government jobs starting in December 1993. At the same time, government jobs in India have an upper age limit for entry. For example, to be eligible to apply for the Indian Administrative Services, candidates must be no more than 30 years old. Some classes of candidates have relaxed age limits. For example, OBCs are eligible up to the age of 33. The phasing in of quotas for OBCs as well as an age limit for eligibility would result in an increased number of younger OBCs taking up government jobs if these jobs are attractive relative to their outside option. Therefore, there should be an increase in the proportion of OBCs in government employment for cohorts less than 33 years old in December 1993 relative to other caste groups. The impact on education is measured in a similar manner.

Treatment is defined in one of two ways. If the year of marriage is observable, then the treated group is OBC individuals that were single in 1993 and less than 33 years of age then.

This definition allows for the reservations to have made an impact through the marriage market. If the year of marriage is not observable, then the treated group is defined as OBC individuals that are less than 30 years old in 1993. Further, individuals that are between 30 and 33 years of age are excluded in these specifications. In the absence of observable year of marriage this definition allows for the marriage market to have an effect since a significant portion of individuals below the age of 30 are married. Excluding the ages 30-33 reduces measurement error for treatment. This exclusion does not have a significant effect on any of the reported results.

The specification used to measure if there is any effect on the attractiveness of the spouse uses the first definition of treatment and is

$$\begin{aligned} Spouse\ Attribute_{icrs} = & \alpha + \beta_{1cs} + \beta_{2r} + \beta_{3r}Time + \beta_{4r}Education \\ & + \beta_5\delta(year\ of\ marraige_i) + \beta_6Height_i \\ & + \beta_7BMI_i + \beta_8OBC\ X\ After + \epsilon_{icrs} \end{aligned}$$

where, the outcome is an attribute of the spouse of individual  $i$ , of birth cohort  $c$ , community  $r$  and state  $s$ .  $\delta(year\ of\ marraige_i)$  is a dummy for the year of marraige. Caste specific fixed effects and linear time trends are controlled for in addition to state and birth cohort. Other controls are individual attributes like height, BMI and an interaction of education and caste. Parental background is controlled for in some specifications and is measured using family landholdings. The coefficient of interest in this specification is  $\beta_8$ , that measures the differential change for OBCs across the phasing in of reservations.

Impacts on the household consumption of jewelery and the practice of purdah are measured using the second definition of treatment. The specification used is

$$\begin{aligned} Outcome_{ij} = & \alpha + State_{ij}\ X\ Birth\ Cohort_i + State_{ij}\ X\ Birth\ Cohort_j \\ & + Caste_i + OBC_i\ X\ \delta(Age_i < 45) + Log(income_{ij}) \\ & + Education_i\ X\ Education_j + \epsilon_{ij} \end{aligned}$$

where  $i$  is the husband and  $j$  is the wife. The only caste observed is that of the household

head, who is usually either the husband, his father or his brother. Observations where these are not the case are dropped. Treatment is defined as the male being a member of the OBC community and being less than the age of 45, ie, he was young enough to be under the threshold of age eligibility for a government job in 1994. The caste of the wife is not used since it is not observed in the IHDS dataset. The coefficient of interest is that of  $OBC_i X \delta (Age_i < 45)$ .

The specification used to study the impact on female labor supply is

$$\begin{aligned} Wife\ Currently\ Working_{ij} = & \alpha + State_{ij} X Birth\ Cohort_i + State_{ij} X Birth\ Cohort_j \\ & + Year\ of\ Marriage_{ij} + Education_i X Caste_i \\ & + Education_j X Caste_j + Caste_i X Time + Caste_j X Time \\ & + After_{ij} + OBC_i X After_{ij} + OBC_j X After_{ij} + \epsilon_{ij} \end{aligned}$$

where  $i$  is the husband and  $j$  is the wife. The coefficient of interest is that of  $OBC_i X After_{ij}$  and  $OBC_j X After_{ij}$ .

The other measure of the distribution of the intrahousehold surplus is the fertility choice of couples. Along with the actual number of children for each couple, the NFHS reports the ideal number of children desired by the husband and the wife. If children are assumed to be a public good and parental preferences are assumed to be single peaked at their ideal number of children, the actual number of children would be a weighted average of the number of children of the husband and the wife, where the weights equal the Pareto weights assigned to the utilities of each individual. The specification used is

$$\begin{aligned} y_{ij} = & \alpha + State_{ij} X Birth\ Cohort_i + State_{ij} X Birth\ Cohort_j + Year\ of\ Marriage_{ij} \\ & + Education_i X Caste_i + Education_j X Caste_j + Caste_i X Time \\ & + Caste_j X Time + After_{ij} + OBC_i X After_{ij} + OBC_j X After_{ij} \\ & + Ideal\ Number\ Husband_i + Ideal\ Number\ Wife_j \\ & + Ideal\ Number\ Husband_i X After + Ideal\ Number\ Wife_j X After \end{aligned}$$

$$\begin{aligned}
&+Ideal\ Number\ Husband_i\ X\ After_{ij}\ X\ OBC_i \\
&+Ideal\ Number\ Wife_j\ X\ After_{ij}\ X\ OBC_j \\
&+Interactions_{ij} + \epsilon_{ij}
\end{aligned}$$

where,  $y_{ij}$  is the number of children, total and of each gender, that the couple  $ij$  actually has, *Ideal Number* is the ideal number of children that each spouse reports and *Interactions* are the interactions of the ideal number of children for each spouse interacted with education, age, physical attributes and caste of both members of the couple. This is to rule out the differential effect that changes in these observables might have on fertility choice. The coefficient of interest is that on *Ideal Number Husband<sub>i</sub> X After<sub>ij</sub> X OBC<sub>i</sub>* and *Ideal Number Wife<sub>j</sub> X After<sub>ij</sub> X OBC<sub>j</sub>*. If the result of fertility is part of the intrahousehold bargaining process, then these coefficients should sum to 0 since they represent changes in the relative Pareto weights assigned to the husband and the wife and these always sum to 1. This is a testable hypothesis.

## 4 Results

### 4.1 Government Jobs

Figures 1-3 show the cohort-wise probability that a person is employed by the government relative to the state average for each community group for each year. This is plotted for the years 1999-00, 2004-05 and 2011-12. Sub-figure (a) shows this for STs, (b) for SCs, (c) for OBCs and (d) for the others.

Figures 1(a)-(d) show that the probability of being employed by the government is lower for younger workers. This is true for each community and reflects the shrinking size of the government relative to the private sector in India. However, the rate of decrease in the probability of government employment by age changes differentially over time for each group. There is no significant change in the rate of decrease for STs, SCs and others. However, the rate of decrease lowers considerably for OBCs from 1999-00 to the later years. This means

that OBCs from younger cohorts are more relatively likely to be employed by the government in 2004-05 than other groups, which matches the prediction if the policy is implemented.

Figure 2(a)-(d), which shows the probability of government employment for men, matches the findings of Figure 1. Figures 3(a)-(d) show the probability of being employed by the government for women of each community. These show that there is no differential change in the probability of being employed by the government for women. The labor market effects of the affirmative action program are concentrated among men.

## 4.2 Education

Figures 4, 5 and 6 show the average number of years of education for each cohort in 1999-00, 2004-05 and 2011-12 for each community relative to the state average for that year. There are two main takeaways from these figures. First, figures 4(a)-(d) show that there is convergence in education across groups. The average number of years of education is increasing for the disadvantaged communities, SCs and STs, and is relatively flat for OBCs and others. Further, there is no differential change in rate of change in the average years of education for each community.

Second, figures 5(a)-(d) and 6(a)-(d) show that there is no differential change for men or women of different communities, except for women in the others category(Figure 10(d)). The rate of increase in education among women in this group slowed down considerably from 1999-00 to the later years. This suggests that the labor market affirmative action policy had no effect on educational attainment.

## 4.3 Marriage Market Effects

### 4.3.1 Marriage rates

Figures 7 and 8 show the probability of getting married for each caste group by education level for men and women, respectively. There does not appear to be a differential change in the probability of getting married for the treatment group in either men or women for

any education level. This suggests that any marriage market effect is not on the extensive margin.

### **4.3.2 Are Spouses More Attractive?**

The effects of being a part of the community that receives affirmative action are investigated.

Table 1 shows how wife attributes vary with treatment, which is defined here as being a member of the OBC community with a marriage date after 1993. Such husbands are found to pair up with wives that have received on average 1.79 years more education, are 17cm taller and have a BMI higher by 5.89 units. The age difference is also larger, wives are younger by 2 years on average. These men are also 2% more likely to marry outside their community group, which is 25% of the 8% average rate of marrying outside one's community.

Table 2 shows spousal attributes for women. Women from the OBC community that were married after 1993 are not found to marry husbands that are more educated. Their husbands are 13cm taller on average but do not have higher BMI. They are also not significantly more likely to marry outside their community and the age difference to their spouses is lower by 1.5 years.

These findings suggest that OBC men become more attractive spouses over time.

### **4.3.3 Distribution of the Intra Household Surplus**

The outcomes investigated are the budget shares of jewellery, the practice of the custom "purdah", that is associated with a lower decision making power of women, female workforce participation and fertility choices made by households.

Table 3 reports that the consumption of jewellery and the practice of purdah did not change in treated households.

Table 6 reports results for female workforce participation. OBC women are 6% less likely to work outside the home if they were married after 1993. Further, wives of OBC men are 4-5.5% less likely to work outside the home in some specifications.

#### 4.3.4 Fertility Choices

Table 4 reports summary statistics on the ideal number of children for each caste group before and after 1993. In the prior period, women of all groups desired more children than men on average. This was true regardless of the gender of children. In the post treatment period, this relationship changes. Women and men desire similar number of children for SCs and STs. OBC women still desire more children than OBC men on average. General category men desire more children than women in the same group.

Table 5 reports regression results from the diff-in-diff strategy. It is found that the actual number of children was closer to that desired by the husband than the wife for if the couple is treated and the husband belongs to the OBC community by 0.6%. This translates to a fall of 0.03 children in the lifetime fertility of these couples. The findings are consistent with a redistribution of Pareto weight within a couple, the p-value for the difference of the change in the Pareto weight of the husband and wife caused by treatment is 0.753, and is therefore, statistically not different from 0.

The change in fertility is largely concentrated in the number of boys. The coefficient for boys is similar in magnitude to that of the total number of children. However, it is not statistically significant.

## 5 Structural Estimation

### 5.1 Theory

The theoretical framework is similar to the Chiappori, Salanie and Weiss(2015) multiple markets extension of Choo and Siow(2006).

#### 5.1.1 Preferences

Assume a population of men  $i \in N_I$  and a population of women  $j \in N_J$ . Every man  $i$  (woman  $j$ ) has a vector of characteristics  $I = (E_i, c_i, r_i) \in \mathbf{I}$  ( $J = (E_j, c_j, r_j) \in \mathbf{J}$ ), denoting education

$E$ , cohort  $c$  and caste  $r$ , respectively. Individual utilities consist of an economic and a non economic taste-based component.  $V_i(x_i)$  is the economic utility of staying single for man  $i$ . A similar notation holds for woman  $j$ .

The couple consisting of man  $i$  and woman  $j$  is assumed to have transferable utility and generates the economic utility  $H(i, j)$ . The economic surplus to marriage for the couple  $(i, j)$  is then given by  $H(i, j) - V_i(x_i) - V_j(x_j)$ .

Suppose the man has characteristics  $I$  and the woman has characteristics  $J$ . Then, the expected economic surplus of marriage for the couple is

$$S^{IJ} = E \{H(i, j) | I, J\} - E \{V_i(x_i) | I\} - E \{V_j(x_j) | J\}$$

. Expectations are used because future incomes are not yet realized but agents are assumed to perfectly know the distribution of these incomes based on currently observable characteristics.

Agents are also assumed to have marital preferences over partners that are not completely reflected by economic utility. As assumed by Choo and Siow(2006), preferences over partners are assumed to be individual specific and only depend on the partners characteristics, caste and education, and not individual identity. A man  $i$  with characteristics  $I$  has a vector of preferences  $a_i^I = (a_i^{i\phi}, (a_i^{IJ})_{J \in \mathbf{J}})$ , where,  $a_i^{i\phi}$  is the taste based utility of man  $i$  with characteristics  $I$  from staying single and  $a_i^{IJ}$  is the utility from marrying a woman with characteristics  $J$ . Note again that preferences are over spouse characteristics and not over particular spouse identity.

Similarly, woman  $j$  with characteristics  $J$  has a vector of preferences  $b_j^J = (b_j^{\phi J}, (b_j^{IJ})_{I \in \mathbf{I}})$ , where,  $b_j^{\phi J}$  is the utility of woman  $j$  with characteristics  $J$  from staying single and  $b_j^{IJ}$  is the utility from marrying a man with characteristics  $I$ .

It is assumed, as in Choo and Siow(2006), that these additively separable shocks are the only source of unobserved heterogeneity. The distribution of these shocks may, however, vary with the characteristics of the man or the woman. Define  $A^{IJ} = E(a_i^{IJ})$  and  $B^{IJ} = E(b_j^{IJ})$ . The total marital surplus generated by a couple is the sum of the expected surplus from consumption and the surplus from the idiosyncratic preferences of spouses. This is given by

$s_{ij} = S^{IJ} + (a_i^{IJ} - a_i^{I\phi}) + (b_j^{IJ} - b_j^{\phi J})$ . This can be rewritten as,

$$\begin{aligned} s_{ij} &= (S^{IJ} + (A^{IJ} - A^{I\phi}) + (B^{IJ} - B^{\phi J})) \\ &\quad + \left( (a_i^{IJ} - A^{IJ}) - (a_i^{I\phi} - A^{I\phi}) \right) \\ &\quad + \left( (b_j^{IJ} - B^{IJ}) - (b_j^{\phi J} - B^{\phi J}) \right) \end{aligned}$$

Define  $Z^{IJ} = (S^{IJ} + (A^{IJ} - A^{I\phi}) + (B^{IJ} - B^{\phi J}))$ , which is the expectation of the total surplus conditional on the observable characteristics of the husband and the wife. The component  $S^{IJ}$  is the conditional expected economic surplus, and  $(A^{IJ} - A^{I\phi}) + (B^{IJ} - B^{\phi J})$  denotes the conditional expected surplus from marital preferences. The terms  $\left( (a_i^{IJ} - A^{IJ}) - (a_i^{I\phi} - A^{I\phi}) \right)$  and  $\left( (b_j^{IJ} - B^{IJ}) - (b_j^{\phi J} - B^{\phi J}) \right)$  have a conditional expectation of 0 by construction. Define  $\alpha_i^{IJ} = a_i^{IJ} - A^{IJ}$  and  $\beta_j^{IJ} = b_j^{IJ} - B^{IJ}$ . Then, the total expected surplus from marriage can be written as,  $s_{ij} = Z^{IJ} + (\alpha_i^{IJ} - \alpha_i^{I\phi}) + (\beta_j^{IJ} - \beta_j^{\phi J})$

### 5.1.2 Matching

A matching consists of (i) a measure,  $d\mu$  on the set  $N_I \times N_J$  and (ii) a set of payoffs  $\{u_i : i \in N_I\}$  and  $\{v_j : j \in N_J\} : u_i + v_j = z_{ij} \forall (i, j) \in \text{supp}(d\mu)$ , ie, an allocation of who marries whom and how the surplus is split between members of a couple.

A matching is stable if (i)  $u_i \geq 0, v_j \geq 0$  and (ii)  $u_i + v_j \geq z_{ij} \forall (i, j) \in N_I \times N_J$ , ie, married agents would not prefer to be single and no two agents can be made strictly better off by matching with each other.

This form of stable matching has been well studied. One property is that matching is stable if and only if it maximizes the total surplus  $\int z d\mu$ . This implies that existence is guaranteed under weak assumptions. Additionally, the dual of this maximization problem generates a shadow price  $u_i$  and  $v_j$  for each man  $i$  and woman  $j$ , respectively, that coincides exactly with the payoffs of the matching problem.

In general, the stable matching is not unique for finite populations. The intuition is that  $u_i$  and  $v_j$  can be altered marginally without violating a finite set of inequality restrictions. As

populations become large, however, the degree of permissible alterations shrinks so that in the limit of a continuous population set, individual payoffs are exactly determined (Chiappori, McCann and Nesheim(2009)).

### 5.1.3 Characterization

Proposition 1 - For any stable matching,  $\exists U^{IJ}$  and  $V^{IJ} \forall I \in \mathbf{I}, J \in \mathbf{J} : U^{IJ} + V^{IJ} = Z^{IJ}$  and that satisfies the property for any matched couple  $(i, j)$  with characteristics  $(I, J)$  that  $u_i = U^{IJ} + (\alpha_i^{IJ} - \alpha_i^{I\phi})$  and  $v_j = V^{IJ} + (\beta_j^{IJ} - \beta_j^{\phi J})$ .

Proof - Assume that  $i$  and  $i'$  both have the same characteristics  $I$  and  $j$  and  $j'$  also have the same characteristics  $J$ . Then, from the definition of the stable match above,

$$u_i + v_j = Z^{IJ} + (\alpha_i^{IJ} - \alpha_i^{I\phi}) + (\beta_j^{IJ} - \beta_j^{\phi J})$$

$$u_i + v_{j'} \geq Z^{IJ} + (\alpha_i^{IJ} - \alpha_i^{I\phi}) + (\beta_{j'}^{IJ} - \beta_{j'}^{\phi J})$$

$$u_{i'} + v_j = Z^{IJ} + (\alpha_{i'}^{IJ} - \alpha_{i'}^{I\phi}) + (\beta_j^{IJ} - \beta_j^{\phi J})$$

$$u_{i'} + v_{j'} \geq Z^{IJ} + (\alpha_{i'}^{IJ} - \alpha_{i'}^{I\phi}) + (\beta_{j'}^{IJ} - \beta_{j'}^{\phi J})$$

Subtracting the 1st from the 2nd and 3rd from the 4th, we get,

$$(\beta_{j'}^{IJ} - \beta_{j'}^{\phi J}) - (\beta_j^{IJ} - \beta_j^{\phi J}) \leq v_{j'} - v_j \leq (\beta_{j'}^{IJ} - \beta_{j'}^{\phi J}) - (\beta_j^{IJ} - \beta_j^{\phi J})$$

which implies that  $v_j - (\beta_j^{IJ} - \beta_j^{\phi J}) = v_{j'} - (\beta_{j'}^{IJ} - \beta_{j'}^{\phi J})$  does not depend on  $j$ , but only on the characteristics  $J$ , so that  $v_j - (\beta_j^{IJ} - \beta_j^{\phi J}) = V^{IJ} \forall i \in I, j \in J$ . The proof for  $u_i$  is identical. This means that the equilibrium utility of woman  $j$  has two components, the deterministic component,  $V^{IJ}$  that depends only on the characteristics of the two spouses and the idiosyncratic component  $(\beta_j^{IJ} - \beta_j^{\phi J})$ . The systematic returns to singlehood,  $U^{I\phi}$  and  $V^{\phi J}$  are normalized to 0  $\forall I, J$ .

Proposition 2 - A set of necessary and sufficient conditions for stability are that (i) for

any matched couple  $(i, j)$  with characteristics  $(I, J)$ ,  $(\alpha_i^{IJ} - \alpha_i^{IK}) \geq U^{IK} - U^{IJ} \forall K \in \mathbf{J}$  and  $(\beta_j^{IJ} - \beta_j^{KJ}) \geq V^{KJ} - V^{IJ} \forall K \in \mathbf{I}$ , (ii) for any single man  $i$  with characteristics  $I$ ,  $(\alpha_i^{IJ} - \alpha_i^{I\phi}) \leq -U^{IJ} \forall J$ , (iii) for any single woman  $j$  with characteristics  $J$ ,  $(\beta_j^{IJ} - \beta_j^{\phi J}) \leq -V^{IJ} \forall I$ .

Proof - Follows from Proposition 1.

Proposition 2 implies that stability can be characterized by a set of inequalities that apply to individual agents only and not couples.

$U^{IJ}$  can be interpreted as the expected utility if a random man with characteristics  $I$  is matched with a random woman with characteristics  $J$ . However, this is not the expected utility of those men that are matched with women with characteristics  $J$  in equilibrium, since those men will also tend to have a stronger than average preference for women with characteristics  $J$ . The expected marital return to having characteristics  $I$  for a man can be defined as  $\bar{u}^I = \max_{J \in \mathbf{J}, 0} (U^{IJ} + \alpha^{IJ})$ .

#### 5.1.4 Empirical Implementation

The empirical framework is related to that of Choo and Siow(2006), Chiappori, Salanie and Weiss(2015) and Salanie and Galichon(2012). The dataset used is the 2005-06 NFHS survey that contains matched information for couples.

The baseline model is that from Chiappori, Salanie and Weiss(2015), which extends the framework from Choo and Siow(2006). In the latter, the marital surplus from man  $i$  marrying woman  $j$  is given by  $s_{ij} = Z^{IJ} + (\alpha_i^{IJ} - \alpha_i^{I\phi}) + (\beta_j^{IJ} - \beta_j^{\phi J})$ . The insight of the former study is that one can assume that shocks can be assumed to be independent across cohorts and that restrictions on the nature of the variation of the surplus function  $Z^{IJ}$  can allow for testable restrictions. The surplus is allowed to vary as  $Z_c^{IJ} = Z^{IJ} + \zeta_c^I + \xi_c^J$ , where  $c$  is a cohort subscript. The terms  $\zeta$  and  $\xi$  are category specific drifts and represent how the production function of domestic production and therefore, gender roles may have changed over time. Examples might be attitudes towards fertility(number of children desired and how to avoid having more), investments in children, healthcare and others.

Assumption - The idiosyncratic preferences are distributed independently according to an extreme type I value distribution, with  $F(\epsilon) = \exp(-\exp(-\epsilon))$ . This along with Proposition 2, implies that for a man with characteristics  $I$ ,

$$\begin{aligned}\gamma_c^{IJ} &= Pr(\text{man with characteristics } I \text{ matches with woman with characteristics } J) \\ &= \frac{\exp(U_c^{IJ})}{1 + \sum_K \exp(U_c^{IK})}\end{aligned}$$

and

$$\gamma_c^{I\phi} = Pr(I \text{ single}) = \frac{1}{1 + \sum_K \exp(U_c^{IK})}$$

. Similarly, for a woman with characteristics  $J$ ,

$$\begin{aligned}\delta_c^{IJ} &= Pr(\text{woman with characteristics } J \text{ matches with man with characteristics } I) \\ &= \frac{\exp(V_c^{IJ})}{1 + \sum_K \exp(V_c^{KJ})}\end{aligned}$$

and

$$\delta_c^{\phi J} = Pr(J \text{ single}) = \frac{1}{1 + \sum_K \exp(V_c^{KJ})}$$

This system of equations can be inverted to give ,

$$U_c^{IJ} = \ln \left( \frac{\gamma_c^{IJ}}{1 - \sum_K \gamma_c^{IK}} \right)$$

and

$$V_c^{IJ} = \ln \left( \frac{\delta_c^{IJ}}{1 - \sum_K \delta_c^{KJ}} \right)$$

One requirement for existence is that  $\gamma_c^{I\phi} > 0$  and  $\delta_c^{\phi J} > 0$  for every  $I, J$  which implies that the denominator in the logarithms is positive.

The class specific expected utilities can also be evaluated,  $\bar{u}_c^I = E [\max_{J \in \mathbf{J}, 0} (U^{IJ} + \alpha^{IJ})] = \ln (\sum_K \exp (U_c^{IK}) + 1) = -\ln (\gamma_c^{I\phi})$  and  $\bar{v}_c^J = \ln (\sum_K \exp (V_c^{KJ}) + 1) = -\ln (\delta_c^{\phi J})$ . Therefore, the sufficient statistic for computing the expected utility is the probability of remaining single. This is a general property of multinomial logit models.

This baseline model is overidentified, there are  $(I \times J) \times T$  total observable quantities and  $(I - 1) \times (J - 1) + I \times (T - 1) + J \times (T - 1)$  quantities to be computed. In this setting,  $I = J = 16$  and  $T = 17$  so that there are 3872 degrees of freedom.

The regression equation implied by the model is,

$$\ln \left( \frac{\gamma_c^{IJ}}{1 - \sum_K \gamma_c^{IK}} \right) + \ln \left( \frac{\delta_c^{IJ}}{1 - \sum_K \delta_c^{KJ}} \right) = Z^{IJ} + \zeta_c^I + \xi_c^J + \epsilon_c^{IJ}$$

The coefficients on the RHS can be identified as the coefficients on the reduced form regression

$$\begin{aligned} \ln \left( \frac{\gamma_c^{IJ}}{1 - \sum_K \gamma_c^{IK}} \right) + \ln \left( \frac{\delta_c^{IJ}}{1 - \sum_K \delta_c^{KJ}} \right) &= Z^{IJ} (\text{caste}^I \times \text{education}^I \times \text{caste}^J \times \text{education}^J) + \\ &\zeta_c^I (\text{caste}^I \times \text{education}^I \times \text{cohort}_c) + \\ &\xi_c^J (\text{caste}^J \times \text{education}^J \times \text{cohort}_c) + \epsilon_c^{IJ} \end{aligned}$$

Match probabilities are calculated for each pair of types of spouses. These are calculated for each state by caste, birth cohort group and education level for the husband and wife. This results in a  $80 \times 80$  matrix for each state. An element  $ij$  of the matrix represents the probability that a husband of type  $i$  matches with a wife of type  $j$ . There is one such matrix for each state, giving  $6400 \times 21 = 134400$  observations for match probabilities. These are used to calculate the imputed marital surplus for a couple of type  $ij$ . Let  $S_{ij}$  de the surplus finction. The assumed surplus function is parametrized using the form -

$$\begin{aligned} S_{ijk} &= \alpha_k + \beta_1 \delta (\text{educ}_i \times \text{caste}_i) + \beta_2 \delta (\text{educ}_j \times \text{caste}_j) + \beta_3 \delta (\text{educ}_i \times \text{educ}_j) + \beta_4 \delta (\text{cohort}_i) \\ &+ \beta_5 \delta (\text{cohort}_j) + \beta_6 \text{treatment}_i + \beta_7 \text{treatment}_j + \beta_8 \text{intercaste}_{ij} + \epsilon_{ijk} \end{aligned}$$

where  $\delta$  is a category dummy operator. The surplus function has 3 important restrictions -

(1) individual idiosyncratic preferences are additively separable from the systematic marital surplus, (2) each individual contributes a certain amount to the marital surplus no matter who they are matched to and this varies by birth cohort and the interaction of education and caste, (3) part of the surplus is produced through separable complementarities - complementarities in education that do not depend on caste and caste complementarities only depend on marrying someone from the same group (all groups are equally dissimilar from each other). Treatment is defined as cohorts that would be eligible for the affirmative action policy, men and women that are under 45 years of age.

Similar regressions are run to find the return to men and women from marrying certain types of spouses.

$$U_{ijk} = \alpha_k + \beta_1 \delta (\text{educ}_i \times \text{caste}_i) + \beta_2 \delta (\text{educ}_j \times \text{caste}_j) + \beta_3 \delta (\text{educ}_i \times \text{educ}_j) + \beta_4 \delta (\text{cohort}_i) \\ + \beta_5 \delta (\text{cohort}_j) + \beta_6 \text{treatment}_i + \beta_7 \text{treatment}_j + \beta_8 \text{intercaste}_{ij} + \epsilon_{ijk}$$

$$V_{ijk} = \alpha_k + \beta_1 \delta (\text{educ}_i \times \text{caste}_i) + \beta_2 \delta (\text{educ}_j \times \text{caste}_j) + \beta_3 \delta (\text{educ}_i \times \text{educ}_j) + \beta_4 \delta (\text{cohort}_i) \\ + \beta_5 \delta (\text{cohort}_j) + \beta_6 \text{treatment}_i + \beta_7 \text{treatment}_j + \beta_8 \text{intercaste}_{ij} + \epsilon_{ijk}$$

where  $U_{ij}$  is the surplus that a random man of type  $i$  receives if matched with a random woman of type  $j$ .  $V_{ij}$  is the analogue for women.

Results from these regressions are reported in Table 7. It is found that the total marital surplus is much lower for intercaste couples. Most of the cost of the lower surplus is borne by men. Treatment does not have a significant effect on the total surplus. Male surplus does not change significantly either. Female surplus is found to decrease for OBC women of the treated cohorts. This suggests that the price of caste on the marriage market changes as a result of treatment.

The model can be used to make counterfactual simulations. This is ongoing.

## 6 Counterfactual Simulations

Add section here

## 7 Conclusion

This paper investigates if income shocks affect the price of caste on the marriage market. It finds that OBC men are 10% more likely to be employed by the government as a result of a new affirmative action policy. This results in higher wages for OBCs. OBC women are not more likely to be employed by the government. Consistent with the hypothesis, this causes OBC men to be matched with more attractive spouses - they are more educated, taller and have a higher BMI. OBC men are also 2% more likely to marry outside their caste group. There is no such effect for OBC women. The distribution of the intrahousehold surplus is also affected. While the consumption of jewelry or the practice of purdah is not affected, the number of children is found to be closer to that desired by the husband if he is an OBC in the treatment cohorts. Female workforce participation is also found to decrease, suggesting it may drop further in India as incomes rise.

Finally, a structural model of the marriage market is set up. It finds that preferences for marrying within one's own caste are strong and not diminishing over time. Treatment decreases the expected marital surplus for OBC women while not affecting the male surplus, consistent with the reduced form findings. Using the model to perform counterfactual simulations is the subject of ongoing and future work.

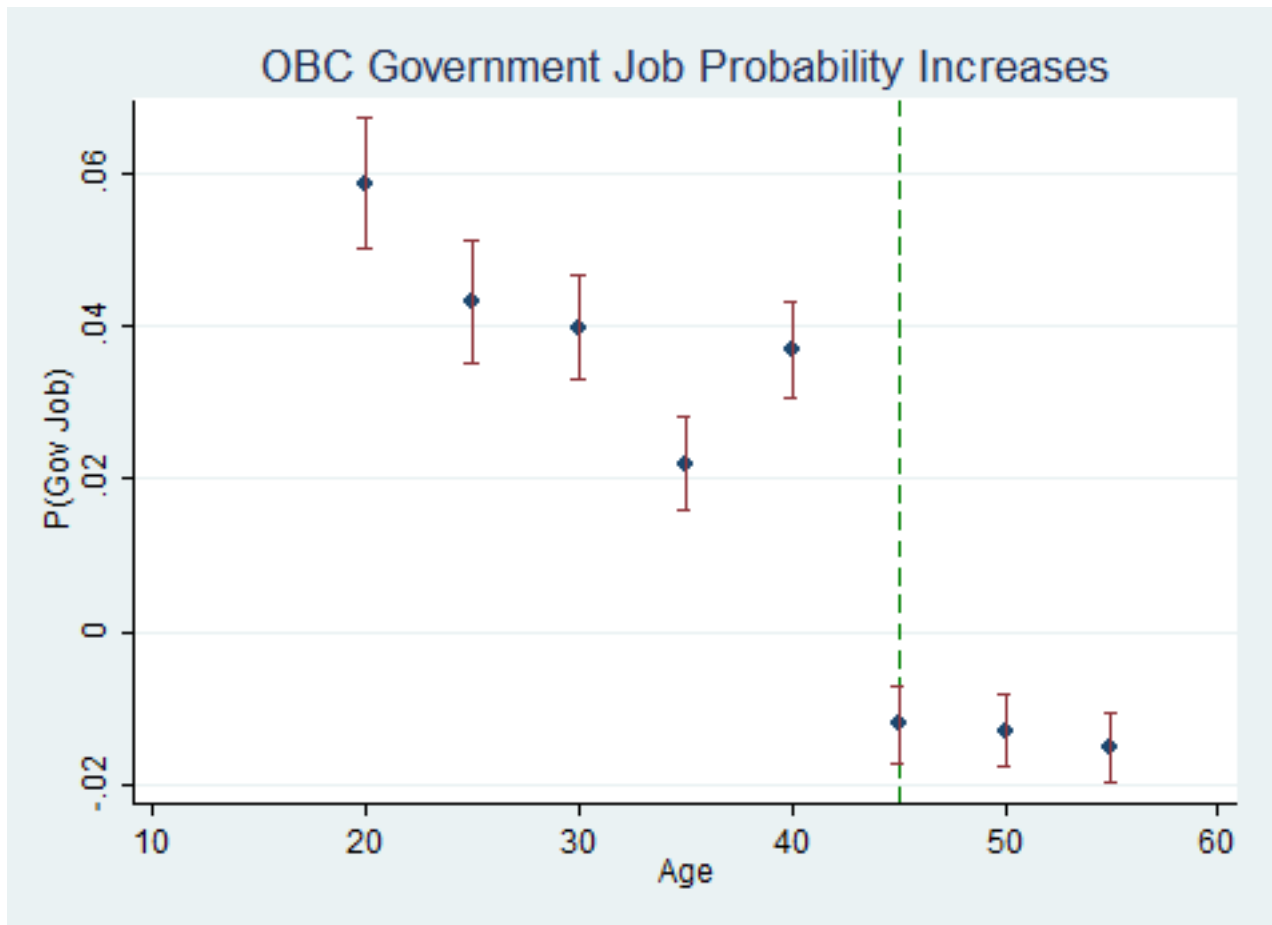
The findings suggest that the welfare distributional effects of affirmative action policies are not trivial and need to be taken into account. Simultaneously, social customs are found to respond to income changes, suggesting that the institution of caste might be weakened by economic development.

## 8 References

1. Anderson, Siwan (2007), "The Economics of Dowry and Bride Price", *Journal of Economic Perspectives* 21(4)
2. Banerjee, Duflo, Ghatak and Lafortune(2009) "Marry for What? Caste and Mate Selection in India",
3. Banerjee and Somanathan(2006) "The Political Economy of Public Goods:Some Evidence from India", *Journal of Development Economics* 82(2)
4. Banerjee and Munshi(2004) "How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur", *Review of Economic Studies* 71
5. Bergstrom and Cornes(1983), "Independence of Allocative Efficiency from Distribution in the Theory of Public Goods", *Econometrica* 51
6. Bertrand, Kamemnica and Pan, forthcoming, "Gender Identity and Relative Income within Households", *Quarterly Journal of Economics*
7. Bobonis, Gustavo(2009), "Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment", *Journal of Political Economy*, 117(3)
8. Burtless, Gary(1999), "Effects of Growing Wage Disparities and Changing Family Composition on the US Income Distribution", *European Economic Review* 43(4)
9. Chiappori, Iygun and Weiss(2009), "Investment in Schooling and the Marriage Market", *American Economic Review* 99(5)
10. Chiappori, Oreffice and Quintana-Domeque(2012), "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market", *Journal of Political Economy* Vol. 120(4)
11. Chiappori, Salanie and Weiss(2015), "Partner Choice and the Marital College Premium : Analyzing Marital Patterns Over Several Decades"

12. Choo and Siow(2006), “Who Marries Whom and Why”, *Journal of Political Economy* 114(1)
13. Dugar, Bhattacharya and Reiley(2011), “Can’t Buy Me Love? A Field Experiment Exploring the Tradeoff Between Income and Caste-Status in an Indian Matrimonial Market”, *Economic Inquiry* 50(2)
14. Duraisamy, P.(2002), “Changes in Returns to Education in India, 1983-94 : by Gender, Age-Cohort and Location”, *Economics of Education Review* 21(6)
15. Dupuy and Galichon(2014), “Personality traits and the marriage market”, *Journal of Political Economy*, 122 (6)
16. Fox(2010), “Identification in Matching Games”, *Quantitative Economics* 1
17. Galichon and Salanie(2013), “Cupid’s Invisible Hand : Social Surplus and Identification in Matching Models”
18. Hnatkovska, Lahiri and Paul(2012), “Castes and Labour Mobility”, *American Economic Journal : Applied Economics* 4(2)
19. Laforune, Jeanne(2013), “Making Yourself Attractive:Pre-Marital Investments and The Returns to Education in the Marriage Market”, *American Economic Journal : Applied Economics* 5(2)
20. Menzel, Konrad(2015), “Large Matching Markets as Two Sided Demand Systems”, *Econometrica* 83(3)
21. Munshi and Rosenzweig(2006), “Traditional Institutions Meet the Modern World:Caste, Gender and Schooling Choice in a Globalizing Economy”, *American Economic Review* 96(4)
22. Munshi and Rosenzweig(2016), "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap", *American Economic Review* 106(1)
23. Rosenzweig and Stark(1989), “Consumption smoothing, migration, and marriage: Evidence from rural India”, *Journal of Political Economy* 97(4)

Figure 1: Government Jobs



9 Figures

10 Tables

Figure 2: Government Jobs - Male

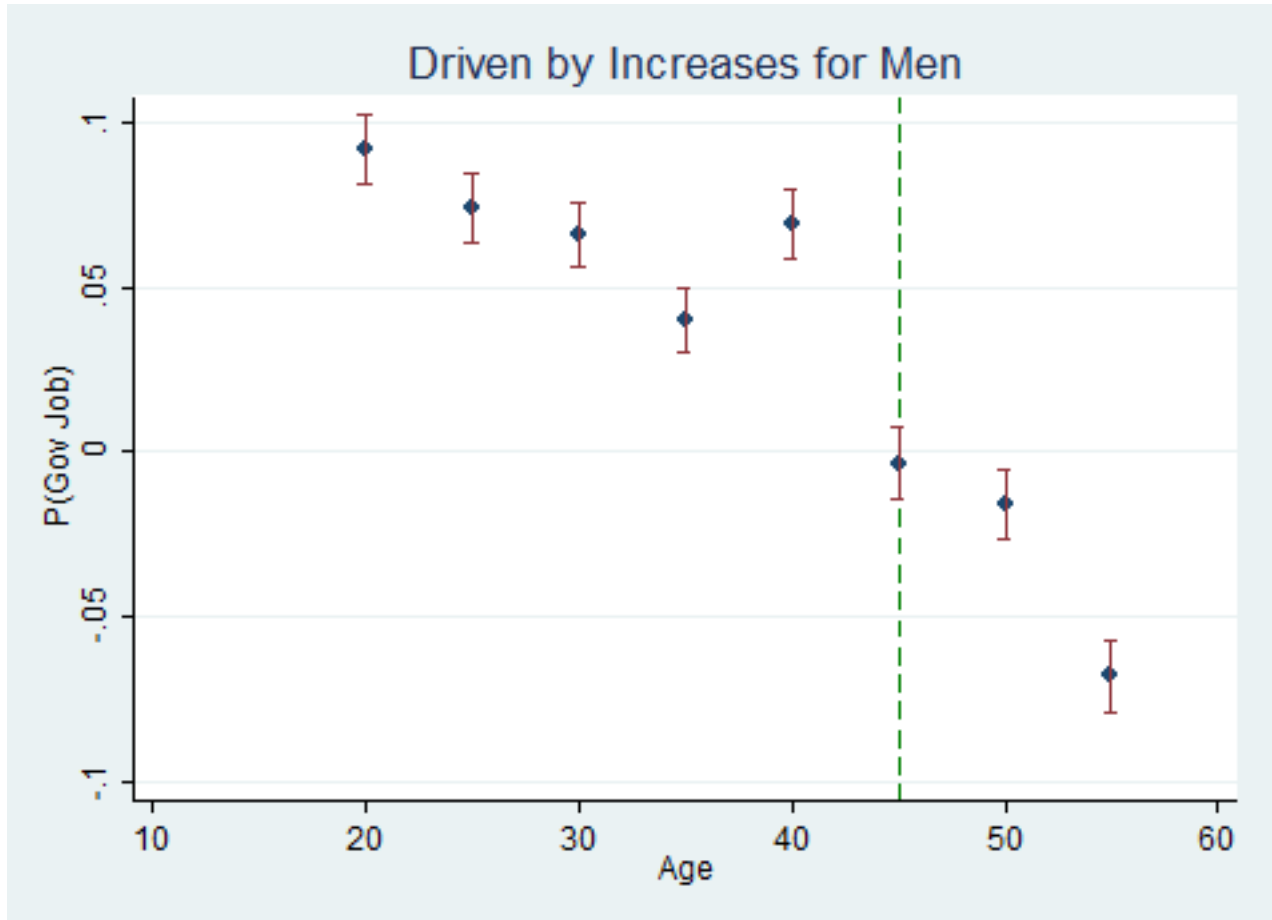


Figure 3: Government Jobs - Female

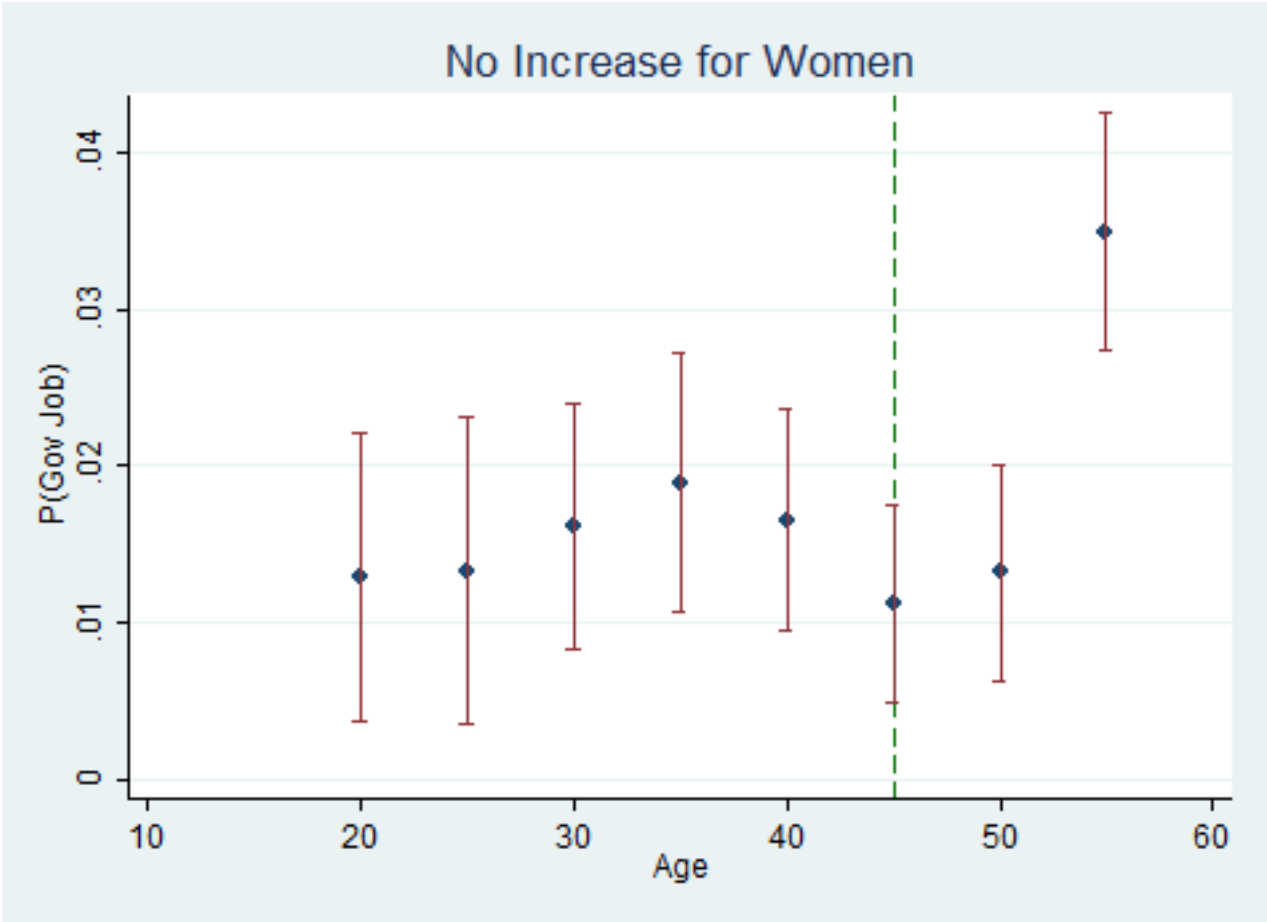


Figure 4: Education

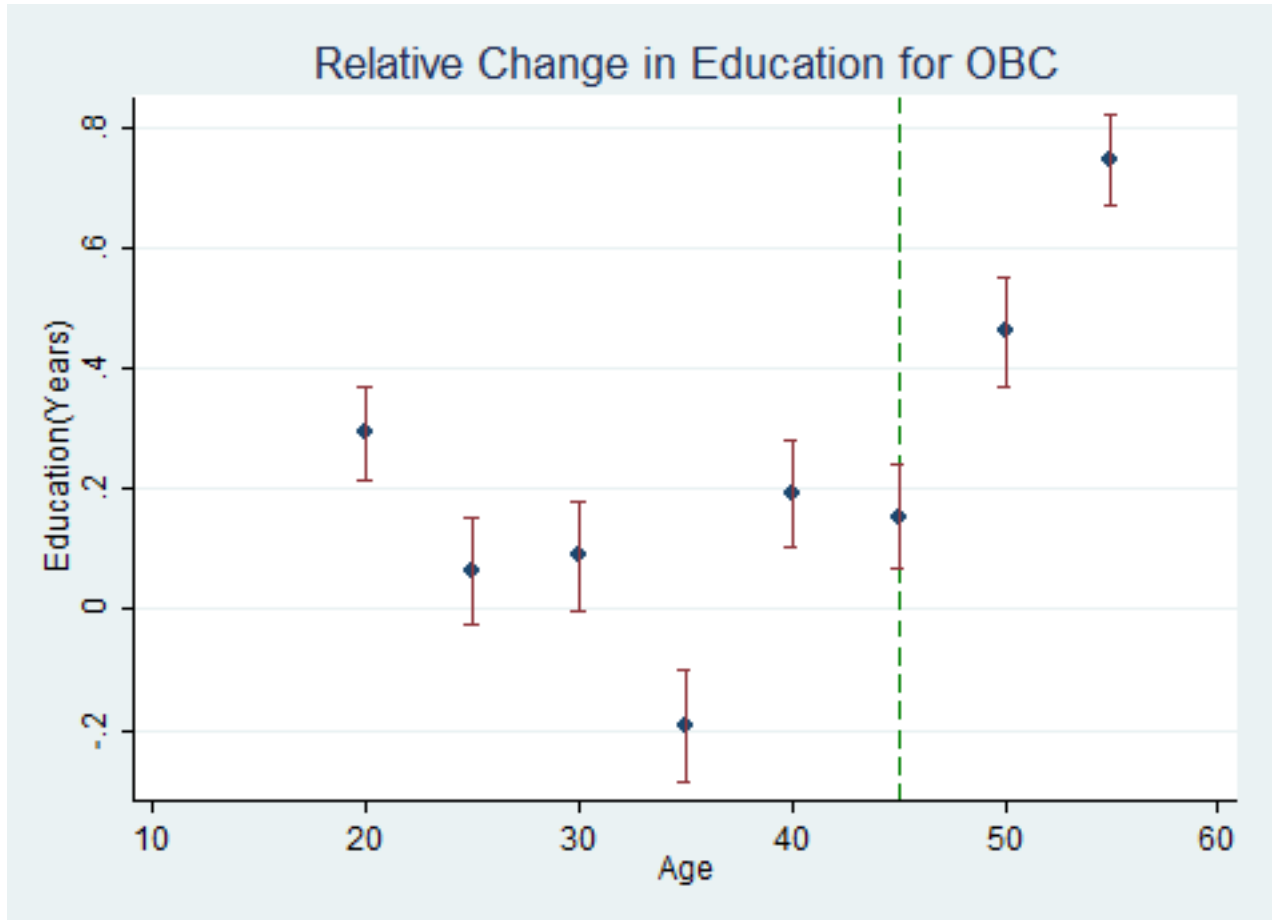


Figure 5: Education - Males

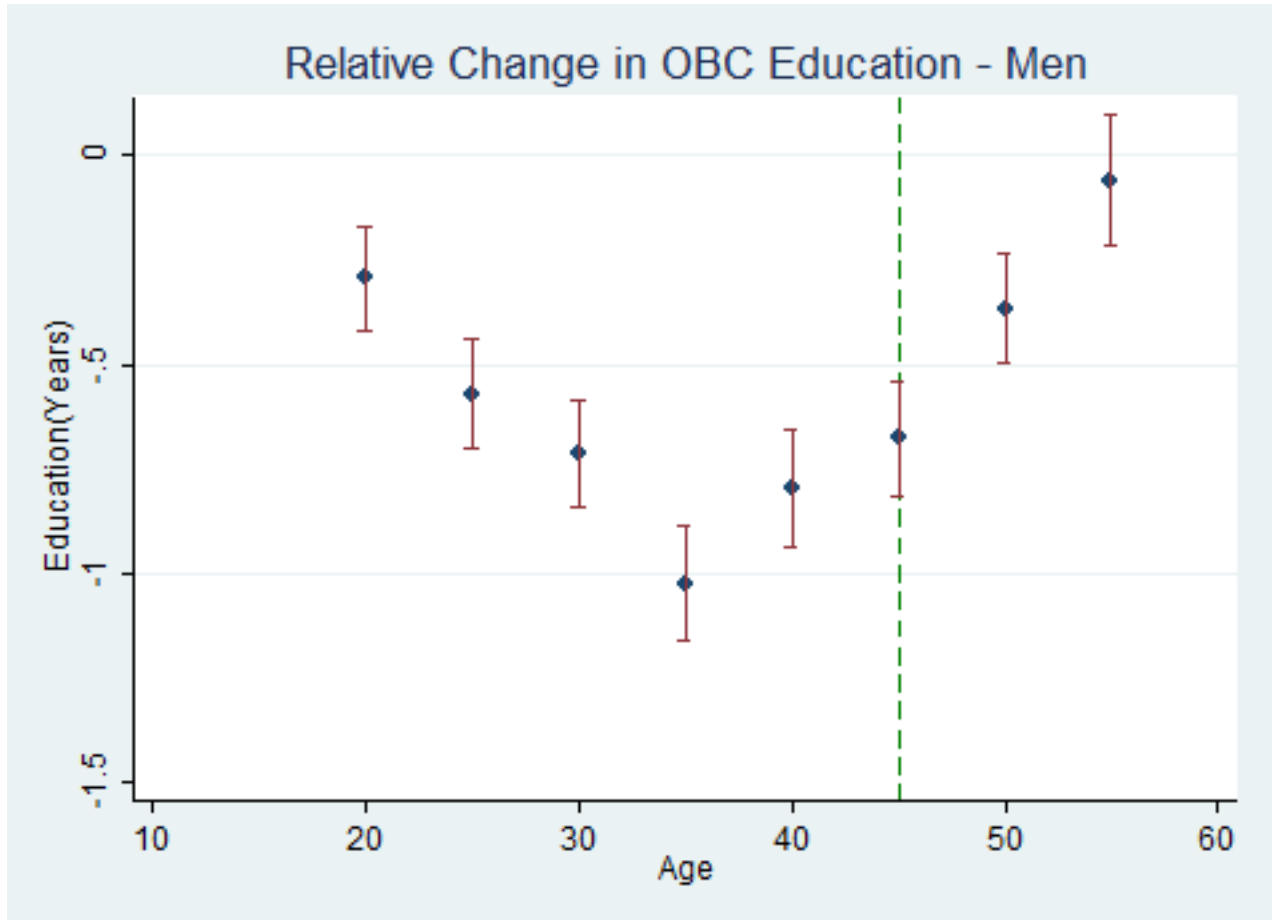


Figure 6: Education - Females

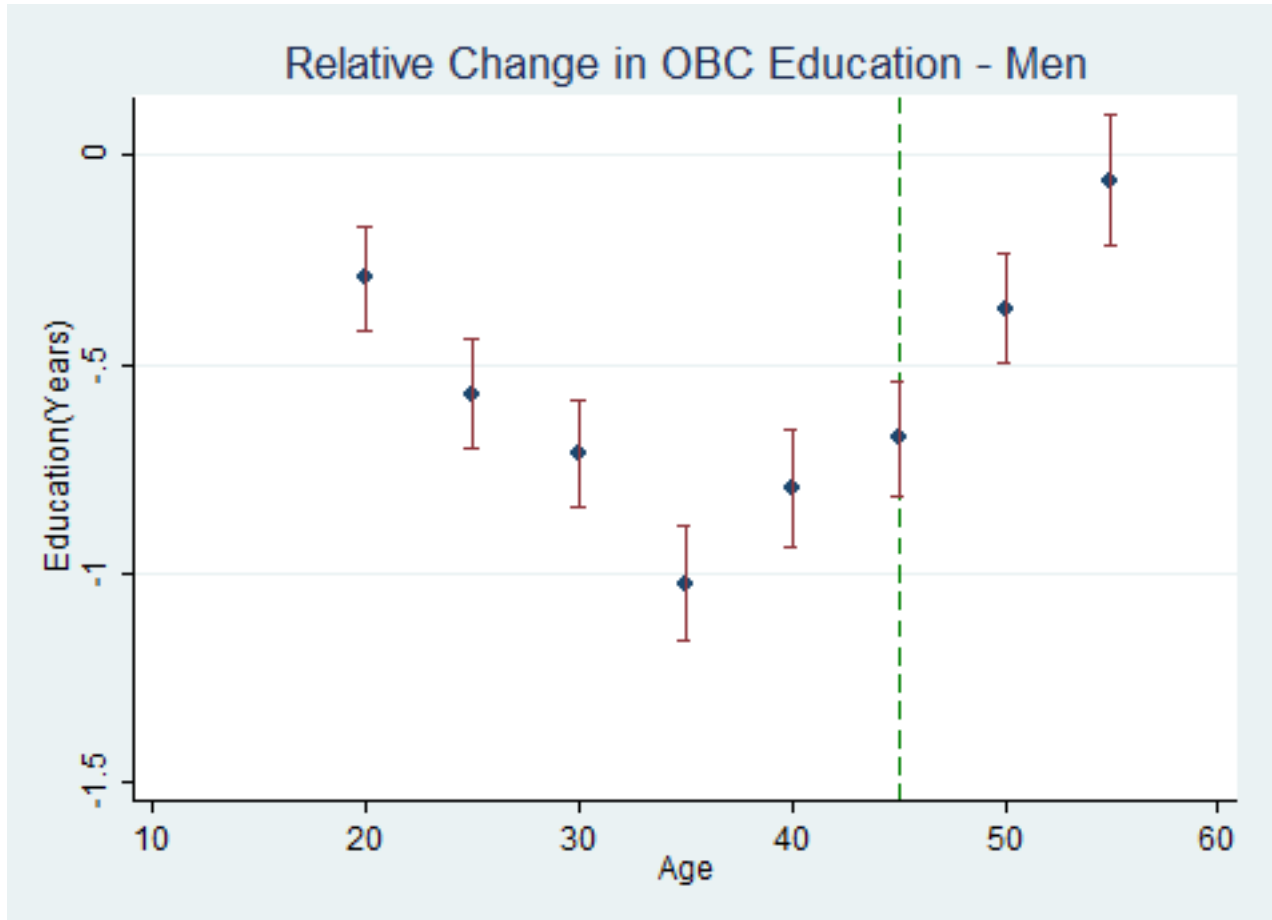


Figure 7: Marriage Rates - Male

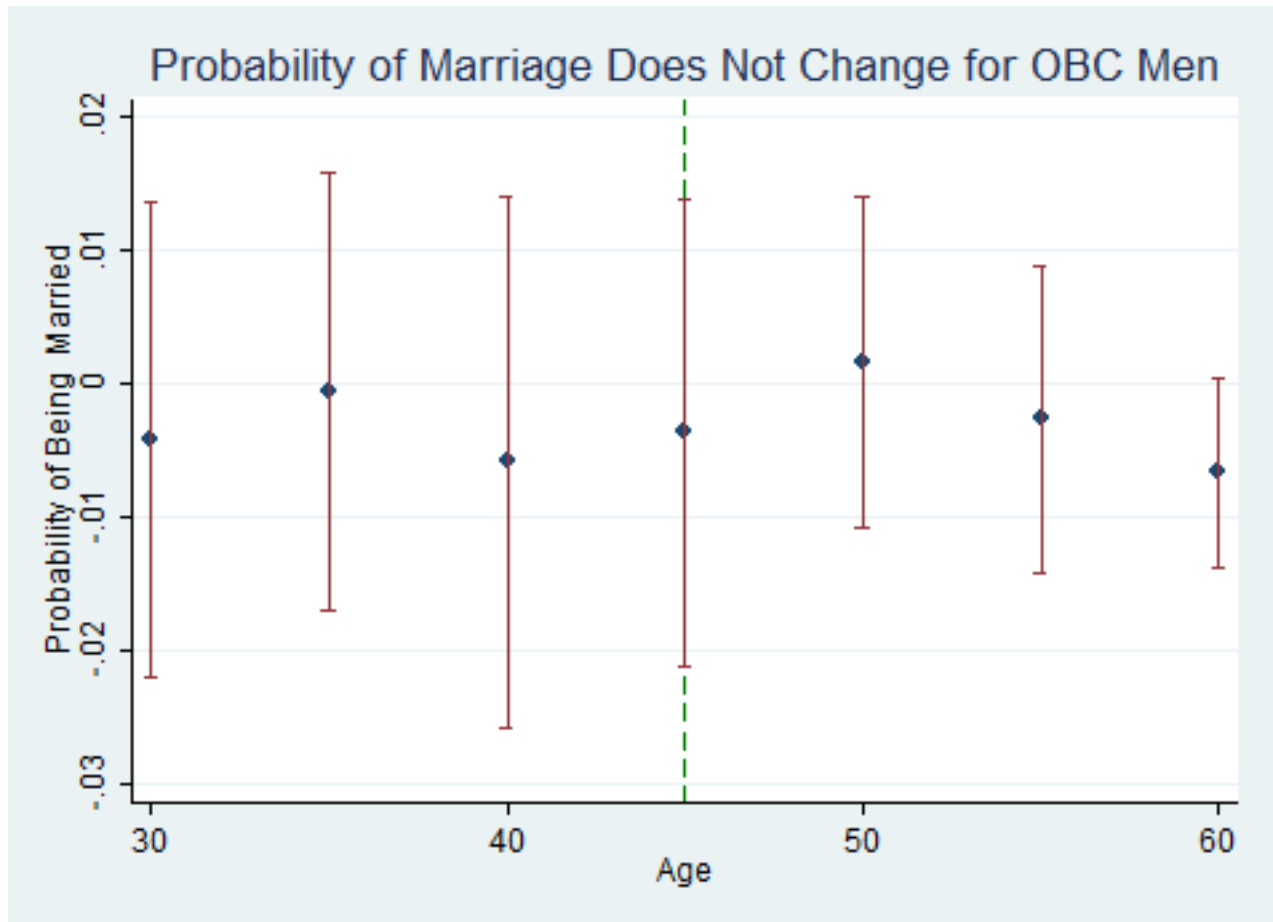


Figure 8: Marriage Rates - Female

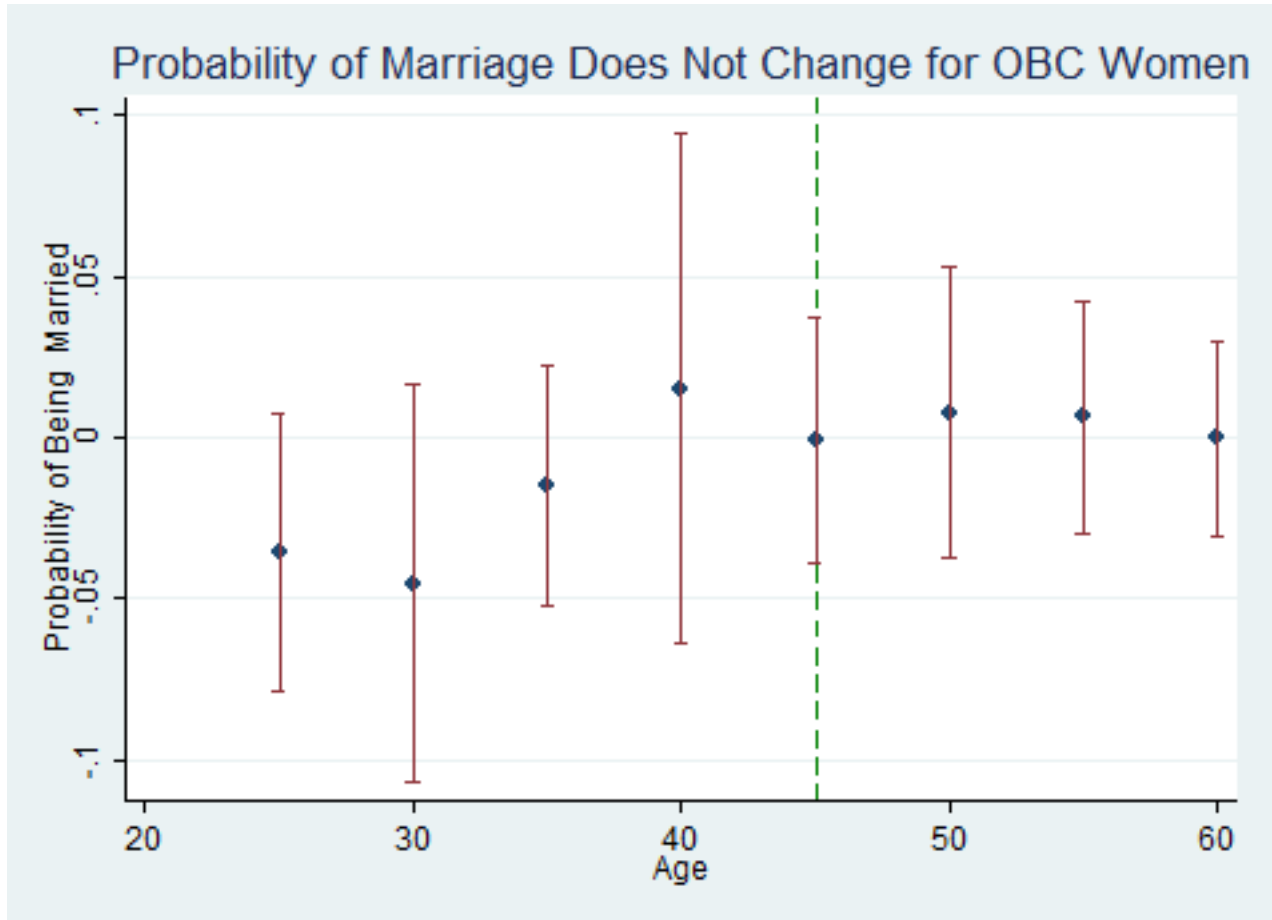


Table 1: Wife Attributes

Outcome -	Education	Height	BMI	Age Diff	Intercaste
Education	0.535*** (30.88)	6.167** (2.015)	11.22*** (5.779)	0.0259*** (4.035)	-0.00228*** (-4.536)
Height	2.82e-05 (1.412)	0.117*** (10.66)	0.00814 (1.601)	-7.66e-05*** (-3.696)	-4.04e-06** (-2.477)
BMI	0.000103*** (3.647)	0.202*** (9.349)	0.280*** (17.54)	7.85e-05** (2.545)	5.70e-06** (2.382)
OBC X Married After 1993	1.790*** (14.47)	170.8*** (2.876)	58.86* (1.731)	2.080*** (18.16)	0.0210* (1.946)
Constant	0.235 (0.894)	278.3** (2.129)	1,385*** (15.99)	1.269*** (5.515)	0.137*** (6.160)
Obs	23,212	23,212	23,212	23,212	23,212
R <sup>2</sup>	0.526	0.096	0.141	0.194	0.035

t-statistics in parentheses, standard errors are clustered at the state birth cohort level \*\*\*  
p<0.01, \*\* p<0.05, \* p<0.1

Table 2: Husband Attributes

Outcome -	Education	Height	BMI	Age Diff	Intercaste
Education	0.693*** (91.67)	28.79*** (6.942)	19.85*** (7.334)	-0.0141** (-2.036)	-0.00293*** (-4.606)
Height	-2.53e-05 (-1.234)	0.166*** (13.36)	0.0414*** (4.853)	-5.62e-05*** (-2.768)	-3.53e-06** (-2.194)
BMI	3.72e-05 (1.211)	0.281*** (10.29)	0.364*** (15.57)	7.48e-05*** (2.640)	5.21e-06** (2.404)
OBC X Married After 1993	0.154 (1.208)	139.9** (2.004)	26.76 (0.604)	-1.537*** (-12.16)	0.0146 (1.118)
Constant	3.758*** (14.55)	-26.88 (-0.163)	1,333*** (12.92)	7.794*** (25.08)	0.113*** (5.938)
Obs	23,212	23,212	23,212	23,212	23,212
R <sup>2</sup>	0.435	0.128	0.155	0.156	0.033

t-statistics in parentheses, standard errors are clustered at the state birth cohort level \*\*\*  
p<0.01, \*\* p<0.05, \* p<0.1

Table 3: Consumption Within The Household

Outcome -	Log(Tobacco+Alcohol)	Male Hours Watching TV	Female Hours Wacthing TV
OBC*Age<45	-0.013 (0.10)	-0.114 (-0.80)	-0.519** (-2.03)
Log(Income)	0.206*** (6.31)	0.149*** (6.59)	0.223*** (7.69)
Constant	3.472*** (4.54)	0.622 (0.95)	1.31 (1.45)
Obs	4,045	5,725	5,679

t-statistics in parentheses, standard errors are clustered at the state birth cohort level \*\*\*

p<0.01, \*\* p<0.05, \* p<0.1

State cohort fixed effects and household characteristics controlled for

Table 4: Ideal Fertility

(a) 1993 and Before						
Number of -	Children		Boys		Girls	
	Women	Men	Women	Men	Women	Men
SC	4.531 (0.288)	4.208 (0.259)	3.315 (0.291)	2.938 (0.265)	2.969 (0.291)	2.562 (0.264)
ST	5.111 (0.506)	4.811 (0.447)	3.782 (0.514)	3.492 (0.448)	3.414 (0.512)	3.104 (0.452)
OBC	4.424 (0.227)	3.915 (0.177)	3.193 (0.230)	2.623 (0.181)	2.907 (0.229)	2.324 (0.180)
Gen	4.531 (0.269)	4.562 (0.330)	3.327 (0.272)	3.289 (0.333)	3.101 (0.271)	3.020 (0.332)

(b) After 1993						
Number of -	Children		Boys		Girls	
	Women	Men	Women	Men	Women	Men
SC	2.627 (0.219)	2.696 (0.240)	1.615 (0.293)	1.605 (0.299)	1.459 (0.291)	1.365 (0.296)
ST	3.204 (0.626)	3.195 (.507)	1.971 (0.632)	1.871 (.512)	1.752 (0.637)	1.637 (.512)
OBC	2.858 (0.238)	2.600 (0.178)	1.642 (0.240)	1.322 (0.179)	1.532 (0.240)	1.152 (0.179)
Gen	2.461 (0.210)	3.183 (0.323)	1.309 (0.213)	1.956 (0.327)	1.194 (0.213)	1.833 (0.327)

Table 5: Fertility Choices

Number of -	Children	Boys	Girls
Ideal Number - Husband	0.00911*** (0.00139)	0.00533*** (0.000887)	0.00299*** (0.000960)
Ideal Number X Treatment	0.00654*** (0.00251)	0.00681 (0.00436)	-0.000313 (0.00253)
Ideal Number - Wife	0.00541*** (0.000990)	0.00172*** (0.000649)	0.00307*** (0.000795)
Ideal Number X Treatment	-0.00479 (0.00319)	-0.00294 (0.00180)	-0.00106 (0.00269)
Constant	15.53*** (3.644)	47.98*** (2.546)	-32.74*** (2.767)
Obs	23,212	23,212	23,212
R <sup>2</sup>	0.367	0.233	0.160

Standard errors in parentheses are clustered at the state birth cohort level \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

State cohort fixed effects included, controls for couple characteristics as well as caste group fixed effects and time trends

Table 6: Female Labour Supply

Dependent Variable : Wife Currently Working						
OBC Husband X Married After 1993	-0.0455*		0.0230	-0.0535*		-0.00424
	(-1.892)		(0.533)	(-1.920)		(-0.0869)
OBC Wife X Married After 1993		-0.0655***	-0.0846**		-0.0642**	-0.0606
		(-2.757)	(-2.005)		(-2.338)	(-1.269)
Constant	14.08***	13.93***	13.97***	1.559	1.601	1.602
	(5.023)	(4.966)	(4.979)	(1.051)	(1.080)	(1.079)
Observations	22,882	22,882	22,882	12,003	12,003	12,003
$R^2$	0.091	0.091	0.091	0.095	0.095	0.095
State X Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Caste Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Caste Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Only Marriages Between 1987 and 1999	No	No	No	Yes	Yes	Yes

t-statistics in parentheses, standard errors are clustered at the state birth cohort level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Covariates controlled for are husband and wife's birth cohort, year of marriage, education, height percentiles and BMIs.

Table 7: Marital Surplus

	Total Surplus	Male Surplus	Female Surplus
Intercaste	-0.005*** (-7.20)	-0.004*** (-8.35)	-0.001* (-1.74)
Treatment Male	-0.002 (-1.45)	-0.001 (-0.61)	-0.001 (-1.44)
Treatment Female	-0.001 (-0.60)	0.000 (0.51)	-0.000 (-0.34)
Treatment Male X Treatment Female	0.005** (2.42)	0.004** (2.15)	0.001 (1.42)
Observations	134400	134400	134400
$R^2$	0.001	0.001	0.001

t-statistics in parentheses, standard errors are bootstrapped. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 8: Counterfactual - No Caste

P(Marriage) - Woman	Illiterate	Primary	Secondary	Higher
Man				
	Illiterate			
	Primary			
	Secondary			
	Higher			