Sex Ratios, Schooling, and the Marriage Market in India*

S Anukriti†

April 2013

Abstract

This paper analyzes the effects of sex ratio imbalances on pre-marital investments and marital outcomes in India. Changes in the availability of pre-natal sex-selection technology differentially altered the mating pool of individuals born in different states, cohorts, and endogamous social groups. Increases in the male-female sex ratio at birth are associated with lower educational attainment, age at marriage, and labor force participation rates, and an increase in spouse's age for women relative to men.

Keywords: India, Sex Ratios, Marriage Market, Schooling

*I am grateful to Pierre-André Chiappori, Janet Currie, Lena Edlund, Todd Kumler, Bentley MacLeod, Cristian Pop-Eleches, Bernard Salanié, Lesley Turner, and Till von Wachter for helpful feedback. I also thank the Columbia University Economics Department for funding access to the National Sample Survey data.

†Department of Economics, Boston College. anukriti@bc.edu.

1 Introduction

In the past few decades, increased availability of ultrasound technology has led to a sharp rise in the male-female sex ratio at birth in several Asian countries. This paper examines the differential effect of imbalances in the sex ratio at birth on pre-marital investments in schooling, marital matching, and post-marital outcomes for men and women in India.

There are many channels through which fluctuations in the sex ratio at birth can affect pre-marital investments in schooling. First, scarcity of women in the population may change marriage and labor market returns from education. Consequently, forward-looking parents may strategically alter investments in their children's schooling as the sex-imbalance grows. In the context of second-generation American immigrants, Lafortune (2011) finds that worse marriage market conditions caused men and women (in highly endogamous groups) to increase investment in education. Second, girls born after prenatal sex-selection becomes feasible are also likely to be more "wanted" than those who are born before, if parents substitute postnatal discrimination with sex-selective abortions. Hu and Schlosser (2011) find a negative association between the sex ratio at birth and the prevalence of malnutrition among girls in India. Similarly, Lin et al. (2010) find that legalization of abortion in Taiwan decreased excess female child mortality. Third, greater sex-selection may differentially affect the family size girls are born into, relative to boys (Jensen (2005)). This is because fertility and sex ratios are jointly determined and greater sex-selection may be associated with lower fertility (Anukriti, 2013). Fourth, access to sex-selection may increase the likelihood that girls are born in families with lower socio-economic status if mating is non-random and parents value grandchildren (Edlund (1999)), thereby worsening the outcomes for females.

Fluctuations in the sex ratio at birth may also affect the likelihood and the timing of marriage as well as the match quality through changes in the size and the characteristics of the mating pool. Abramitzky et al. (2011), Angrist (2002), and many other papers find that an imbalance in the marriage market sex ratio improves marriage outcomes for the scarcer sex. Post-marital outcomes, such as labor force participation, may also change. The labor market effects could be a direct result of the adjustments in educational attainment, but may also be driven by changes in the distribution of intra-household bargaining power resulting from altered matching outcomes. Most prior literature on sex ratios and marriage markets focuses on societies and time periods where social norms about marriage are quite different from the Indian context. This is the first paper to examine the consequences of biased sex ratios on marital outcomes in a developing country with highly endogamous marriage markets.

I use repeated cross-sectional data from the National Sample Survey (NSS) and the National Family Health Survey (NFHS) of India and a triple difference-in-difference type estimation strategy similar to the one adopted by Hu and Schlosser (2011). India exhibits substantial variation in the sex ratio at birth over time and across states. Bhalotra and Cochrane (2010) and Hu and Schlosser (2011) show that the variation in sex ratios at birth over time can be explained by differential access to sex-selection technology. In the absence of reliable annual births data, I construct retrospective sex ratios at birth by pooling the NSS rounds. To the extent that several outcomes of interest in this paper are likely to operate through the marriage market, I also construct sex ratios that vary across endogamous caste-religious groups, in addition to the state and the year of birth.

My results show that an increase in the sex ratio at birth is associated with a decrease in educational attainment for women relative to men. This finding is reflected in reduced school attendance rates for 5-14 year old girls as well as in primary and secondary school completion rates for 15-60 year old women. These results are strongest for lower-caste Hindu families who are less likely to practice sex-selection and, hence, more likely to have daughters in the first place. Moreover, age at marriage decreases (increases) for women (men) and they are more likely to marry older men as sex ratios increase. Since men typically marry younger women, this also implies an increase in the spousal age gap. Women born in higher sex ratio at birth environment are also less likely to participate in the labor force, relative to men. The labor market effects are mainly driven by married women. Together, these results suggest that increases in the sex ratio at birth lead to "improvements" in women's position on the marriage market. However, to what extent decreases in educational attainment, age at marriage, and labor force participation are welfare-improving in the long-run is unclear.

Section 2 describe the data and the empirical strategy. Section 3 presents the results. Section 4 concludes the paper.

2 Empirical Strategy

2.1 Data and Construction of Sex Ratios at Birth

The data used in this paper come from two sources - the NSS Employment Unemployment Survey and the NFHS. I use six quinquennial rounds from the NSS (38^{th}, 43^{rd}, 50^{th}, 55^{th}, 60^{th}, and 66^{th}) conducted in 1983, 1987-88, 1993-94, 1999-00, 2004-05, and 2009-10. These are nationwide, repeated cross-sectional, household surveys, representative at the state level. They include information on household characteristics like size, caste, religion, monthly per capita expenditure, and detailed demographic particulars for each household member, including age, sex, marital status, educational attainment, school attendance, and labor force participation status. However, the NSS does not report age at marriage and spousal characteristics for married individuals. Therefore, in addition to the NSS, I use three rounds of the NFHS that were conducted in 1992-93, 1998-99, and 2005-06 to examine the effect on marital outcomes.

Since the Census of India takes place once every ten years and birth registration is not widespread, the annual number of male and female births have to be estimated using survey data. I use data from the NSS to construct retrospective estimates of the sex ratio at birth. First, I deduce the year of birth for an individual using information on the reported age at the time of the survey.[1] Next, I pool individual data from all six rounds of the NSS and

[1] The NSS does not provide the year of birth.

collapse them at the appropriate level to construct estimates of the total number of males and females in each cell.[2] I use the 7-year moving average of the ratio of the number of male births to female births as a measure of the sex ratio at birth. However, all results are robust to using alternate smoothing procedures.

For most of the paper, the sex ratio at birth is defined and constructed at the state-year level (similar to Hu and Schlosser (2011)). Although the NSS can be used to construct district-level estimates, I use state as the relevant geographic level for the following reasons. First, state is the more appropriate geographic dimension for defining the marriage market (which is likely to be an important explanatory channel for my findings) in India. Second, both data sources used in this paper do not report an individual's place of birth. Consequently, I assume that the place of birth is the same as the place where the interview was conducted. This assumption is especially likely to be violated for married women due to the practice of patrilocal exogamy.[3] The measurement error would, however, be smaller for state-level sex ratios since it is more likely that a woman is married in the same state than in the same district. In general, inter-state migration is quite low in India - for example, less than 4% of the population in 2001 had moved across states in the last ten years according to the Census data.

Indian marriages are also highly endogamous, both in terms of caste (*jati*) and religion. According to the 2005 India Human Development Survey, only 4.4% of women were married to a spouse from a different caste. Using responses to matrimonial advertisements in a Bengali newspaper, **?** find evidence in favor of a strong preference for in-caste marriage - for example, the bride's family is willing to trade-off the difference between no education and a master's degree in the prospective husband to avoid marrying outside their caste. In addition to this "horizontal" preference for same-caste marriages, inter-subcaste marriages are governed by strict rules of hierarchy. Although caste is primarily a Hindu phenomenon, the notion of

[2]The relevant sampling weights are used when data is collapsed to the cell-level.
[3]Most migrants in India are women relocating at the time of marriage.

caste-based hierarchy remains well-preserved among many other religious groups in India. In the 2009 National Sample Survey, 31% of Sikh households identified themselves as belonging to a Scheduled Caste (SC).[4] Lastly, inter-religious marriages are far less common than inter-caste marriages.

Therefore, in addition to the state and the year of birth, I also define sex ratios at birth by caste-religious groups. Since detailed subcaste-level data is not available, I divide the sample into twelve broadly-defined "endogamous" groups using a classification system recognized in the Indian Constitution. The Constitution of India recognizes certain castes, Scheduled Castes (SC), and tribes, Scheduled Tribes (ST), as historically disadvantaged. In addition, the Government of India has classified a large group of castes and communities as "Other Backward Classes" to reflect their social and economic disadvantage. I interact these caste groups with the four largest religious groups in India to create twelve endogamous categories - General caste Hindus, SC Hindus, OBC Hindus, ST Hindus, General caste Muslims, OBC Muslims, General caste Sikhs, SC Sikhs, OBC Sikhs, General caste Christians, OBC Christians, and ST Christians. Some categories, such as ST Muslims and ST Sikhs, are excluded because very few individuals belong to these groups. For the construction of sex ratios by these endogamous groups as well as for subsequent regression analyses based on the NSS data, the sample is restricted to the 55th, 61st, and 66th rounds of the NSS because the earlier rounds (38th to 50th) do not distinguish between OBC and general caste households.

2.2 Descriptive Statistics

Table 3.3.1 presents the summary statistics for the NSS data disaggregated by rounds. The sample comprises 26 states and more than 3 million observations. Roughly 48% of the sample is female. To examine the effects on schooling, I create two sub-samples: 5-14 year old children and 15-60 year old adults. The year of birth ranges from 1969-2004 for the children's sample

[4] The Census of India also allows Sikhs and Buddhists to be classified as Scheduled Castes.

and from 1923-1994 for adults. Average age at the time of the survey is 23 years. About 38% of the individuals reside in a rural area, 78% are Hindus, and 15% belong to a scheduled caste. The sex ratio at birth (at the state-year level) has steadily increased over time - from 1.048 in the 38th round to 1.084 in the 66th round. Children in the 5-14 age-group comprise 12% of the sample. Among them, 69% of the girls and 81% of the boys report attending school at the time of the survey. The school attendance rates have increased for both sexes over time and the gap in schooling has decreased from 21% in 1983 to 2% in 2009-10. Among 15-60 year old individuals, females are more likely to be married and have lower educational attainment as reflected in the gap between literacy and the primary and the secondary school completion rates. They are also less likely to participate in the labor force, but conditional on being in the labor force, the difference in employment rates in not large.

Figure 3.3.1 plots the constructed (state-year) sex ratio at birth by year for India as a whole. Until the early 1980s, the sex ratio at birth was close to the natural rate, i.e. 1.05. Thereafter, the population has increasingly became more male-biased. Since the sex ratio has been constructed from retrospective population data, it is likely to suffer from measurement errors due to differential mortality and migration by sex. As a robustness check, Figure 3.3.1 also plots the sex ratio of 0-year old population as reported in the decennial Census of India. The Census methodology for constructing single year age population returns has changed significantly over time and is also likely to suffer from measurement error.[5] Nevertheless, it is reassuring that the NSS sex ratio at birth follows a similar trend as the Census ratio, although the former appears to be systematically larger than the Census estimates.

Figure 3.3.2 exhibits the substantial state-level variation in the sex ratio at birth (both levels and trends). The sex-imbalance in north and north-west India is, in general, much larger than south and east India. However, it is not the case that all states within the same

[5] For example, in the 1971 Census, the state-level tabulations were made on the basis of a 10 percent sample of individual slips for rural and a 20 percent sample for urban areas. In the 1981 Census, the numbers were calculated based on a 20 percent sample of enumeration blocks, while the 1991 Census used a 10 percent sample of individual slips. Since the 2001 Census, enumeration has been done on a 100 percent basis.

region exhibit the same pattern. For example, Himachal Pradesh has a much lower sex ratio at birth than its neighbors, Haryana and Punjab. Figure 3.3 describes the time variation in the sex ratio at birth by caste for Hindus. While all groups experienced a sharp increase during the early 1980s, the sex ratio appears to have returned to the normal level for the lower caste groups (SC and OBC). However, the upper or general castes continue to exhibit sex ratios significantly above 1.05.

2.3 Regression Specifications

The regression framework seeks to isolate the differential effect of sex ratio imbalances on pre-marital investments, marital matching, and post-marital outcomes of males and females. My identification strategy utilizes the (presumably exogenous) variation in the sex ratio at birth within states, over time, and across endogamous caste-religious groups caused by the variation in the timing and the prevalence of prenatal sex-selection.

For an individual i in state s born in year t and of age a at the time of the survey, my first specification is:

$$\begin{aligned} Y_{ista} = &\alpha + \beta(R_{st} * Female_i) + \gamma R_{st} + \theta_s + \sigma_t + \omega Female_i + X_i^{'}\delta + \psi_a \\ &+ \rho_s Female_i + \lambda_t Female_i + \pi_a Female_i + \phi_s t + \epsilon_{ista} \end{aligned} \tag{1}$$

where Y_{ista} is the outcome of interest, R_{st} is the sex ratio at birth for state s in year t, and $Female_i$ indicates that individual i is female. The vector X_i includes dummy variables for religion (Hindu, Muslim, Sikh, Christian), caste (SC and ST), and residence in a rural area. In addition, I include fixed effects for state (θ_s), year of birth (σ_t), and age at the time of the survey (ψ_a). Sex-specific time-invariant differences across states (for example, in the degree of son preference) are captured by the term $\rho_s Female_i$. Sex-specific year of birth effects ($\lambda_t Female_i$) control for differential all-India trends for males and females. Lastly, I also include state-specific time trends to take into account differential trends across states that are common for males and females. Robust standard errors are clustered at the state-level.

The coefficient γ captures the effect of changes in the sex ratio at birth on males while β measures the differential effect on females relative to males born in the same state and year. All results are robust to the replacement of the main effect of R_{st} and the state-specific time trends with state-year of birth fixed effects. This ensures that my findings are not driven by any unobservable state-time varying factors.

In addition, I also exploit variation in sex ratios by endogamous caste-religious groups to further isolate the effects operating through the marriage market channel. The second specification is as follows:

$$\begin{aligned} Y_{istca} = \alpha + \beta(R_{stc} * Female_i) + \gamma R_{stc} + \omega Female_i + \theta_s + \mu_t + \nu_c + \psi_a + X_i^{'}\delta \\ + \phi_{st} + \sigma_{ct} + \rho_s Female_i + \pi_t Female_i + \eta_c Female_i + \lambda_a Female_i + \epsilon_{istca} \end{aligned} \tag{2}$$

In specification (2), the sex ratio at birth, R_{stc}, has been computed separately for each state, year, and caste-religious group, but otherwise the methodology remains the same. The covariates vector, X_i, comprises an indicator for residence in a rural area. In addition to the fixed effects included in specification (1), now I also control for caste-religion fixed effects (ν_c), interaction fixed effects for state-year (ϕ_{st}), caste-year (σ_{ct}), and sex-specific caste-religious group fixed effects ($\eta_c Female_i$). Robust standard errors are clustered by state.

The effects on educational attainment are measured through four outcome variables. The first variable is an indicator that equals one if a child in the 5-14 age-group reports attending school. However, this variable does not adequately reflect completed education. Therefore, in addition, for adults (in the 15-60 age-group) I construct dummy variables for literacy, completion of primary schooling, and completion of secondary schooling. I also examine the effects on age at marriage, age and educational attainment of the spouse, and labor force participation rate for adults.

3 Results

3.1 Effects on Schooling

Table 3.3.2 presents the estimates for the effects of changes in the sex ratio at birth on educational outcomes using specification (1) and data from the NSS. In columns (1) and (2), the sample is restricted to children in the 5-14 year age-group and the outcome variable is an indicator that equals one if the child attends school, and zero otherwise. Columns (3)-(8) examine the impacts on the educational attainment of older individuals who are likely to have completed schooling by the time of the survey. The sample is restricted to the age-group 15-60 years. The three outcome variables are indicators for being literate, completion of primary schooling, and completion of secondary schooling. Columns (1), (3), (5), and (7) include fixed effects for state, year of birth, their interactions with the sex of the individual, and state-specific linear time trends. In addition, columns (2), (4), (6), and (8) also control for individual covariates (caste, religion, and residence in a rural area) and sex-specific fixed effects for age at the time of the survey.

The coefficient of the interaction between SRB and $Female$, β, is negative and highly significant in columns (1) and (2). This implies that girls born in states and years with a higher sex imbalance are significantly less likely to attend school when they are 5-14 years old, relative to boys born in the same state and year. The overall effect for girls (given by $\beta + \gamma$) is also negative suggesting that an increase in prenatal sex-selection over time is associated with a decrease in the likelihood that a girl attends school as compared to similar girls in the same state who are born in years with a lower sex ratio at birth. The interaction coefficient in column (2) implies that a one unit increase in the sex ratio at birth is associated with a 30 percentage point decrease in the likelihood that a girl attends school, relative to a boy. The average sex ratio at birth increased from 1.048 for individuals surveyed in the 38th round to 1.084 for individuals surveyed in the 66th round, thereby decreasing a girl's relative likelihood of attending school by 1 percentage point. For states like Punjab, where the sex ratio at birth increased by almost 20 percentage points from the 38th to the 66th round, this translates into a 6 percentage point decrease (for a sample mean of 0.774). For boys, the effect is positive, not significant, and smaller in magnitude than girls.

Similarly, the coefficients of the interaction terms are negative and significant for columns (3) - (6) suggesting that 15-60 year old women born in states and years with higher sex ratios at birth are less likely to be literate and to complete primary schooling, relative to men born in the same state-year. There is no significant difference in the secondary schooling completion rate as shown by columns (7) and (8). The overall effects for females are also negative for literacy and primary school completion. Similar to columns (1) and (2), the effect on males is not significantly different from zero. The interaction coefficients in columns (4) and (6) translate into a 0.45 percentage point and 0.39 percentage point relative decrease, respectively, in the literacy and primary completion rates for women who experienced the average increase in the sex ratio at birth. For states like Punjab, the decreases are larger (2.6 and 2.29 percentage points respectively).The findings in Table 3.3.2 continue to hold when data from the NFHS are used instead of the NSS. The NFHS results are presented in Table 3.8.

Table 3.3 presents the results from specification (2). For each outcome variable, two sets of results are reported. In addition to fixed effects for state, year of birth, caste-religious group, age, rural dummy, and interaction fixed effects for sex-state, sex-caste, and sex-year of birth, columns (1), (3), (5), and (7) include fixed effects for state-year of birth and caste-year of birth, which are replaced by state-specific and caste-religious group specific linear time trends in the remaining columns. Like Table 3.3.2, the interaction terms are negative across all columns, but only significant in case of children's school attendance rates.

Unlike Hu and Schlosser (2011), who find that the malnutrition rate for girls decreases as the sex ratio at birth increases, I find that the female educational attainment worsens relative to males. This result cannot be explained by an improvement in the desirability of girls born after prenatal sex-selection becomes available. My findings are more consistent with the marriage market channel that suggests lower pre-marital investments by the scarcer sex (similar to Lafortune (2011)). The inferior educational outcomes for females can also be explained if girls are more likely to be born in low-status families after ultrasound technology becomes available. However, this second mechanism is not consistent with the findings of Hu and Schlosser (2011).

To explore the underlying channels further, next I split the sample by religious affiliation and re-estimate specification (1) for each group separately to examine heterogeneity in the effects on schooling. These results are presented in Table 3.4. The earlier results in Table 3.3.2 appear to be mainly driven by Hindus and Sikhs. In Panel A, the likelihood of attendance falls by 37 percentage points for Hindu girls relative to Hindu boys (whose attendance rates significantly increase) when SRB increases by one unit. Sikh children also show the same pattern of effects, but the coefficients are not significant, likely due to much smaller sample sizes. There is no significant effect for Muslims and Christians. Similarly, in Panel B, literacy and primary completion rates decrease significantly for Hindu and Sikh women, relative to men in these religious groups. These findings are consistent with prior evidence that the practice of prenatal sex-selection is more prevalent among Hindus and Sikhs. Next, in Table 3.5, I further divide the Hindu sample into various castes. Column (5) includes both general and OBC groups since earlier NSS rounds (38th to 50th) do not distinguish between OBC and general caste households. Columns (3) and (4) are based only on the 55th, 61st, and 66th rounds. Both Panels A and B show that the overall decrease in educational attainment of Hindu females in Table 3.4 is driven by lower-caste families. For upper-caste or general caste Hindu women, there is an *increase* in literacy as well as primary and secondary completion rates overall as well as relative to men. On the other hand, upper caste Hindu men experience worse schooling outcomes. The coefficients for upper-caste Hindu children in Panel A are insignificant, however.

These differential effects by caste cannot be explained by potential changes in the characteristics of families into which girls are born. It is possible that the desirability channel is at work for upper-caste families (who are more likely to practice sex-selection) while the marriage market channels are operative for the lower-caste families. It is also important to note that the effects on general caste Hindu women are different from those observed for general caste Hindu girls. However, the time-period covered by the adult sample (1923-1994) is not the same as that of the children's sample (1969-2004). As shown by Figure 3.3, the sex ratios for general castes and lower castes diverged around the mid-1990s. Bhalotra and Cochrane (2010) also point out that a structural break in the diffusion of ultrasound technology occurred in 1995. A better understanding of the underlying channels is beyond the scope of this paper.

3.2 Effects on Marriage

In this section, I examine the effects of being born in a higher sex ratio environment on the timing of marriage as well as the age and the educational attainment of the spouse, conditional on being married at the time of the survey. These regressions are based on data from the NFHS since the NSS neither reports the year of marriage nor provides information about spousal characteristics. Moreover, the first two rounds of the NFHS only interviewed women, whereas NFHS-3 interviewed both men and women. Since my empirical strategy compares men and women, I restrict the sample to NFHS-3 in this section. Table 3.6 presents the estimates for specifications (1) and (2) with age at marriage as the outcome variable. Columns (1) and (2) use the entire NFHS-3 sample. In order to further examine heterogeneity in effects by caste and religion, columns (3) - (6) split the sample by religion, and columns (7) - (10) further split the Hindu sample by caste, for specification (1). Overall, the age at marriage increases for men and decreases for women. This result is consistent with an improvement in (worsening of) women's (men's) position on the marriage market due to their relative scarcity (abundance). In the first two columns, the coefficient of SRB is positive and significant. The interaction term is negative but loses significance at conventional levels in column (2). Hindus and Sikhs exhibit a similar pattern as column (1) and among Hindus, the interaction coefficient is significant only for the SC sample. These findings are consistent with the results in Tables 4 and 5. The coefficients in column (1) approximately translate into a 2-year increase and a 1-year decrease in the age at marriage for men and women, respectively for a unit increase in the sex ratio at birth.

In Table 3.7, the outcome variables are the age and the years of education of the spouse. Columns (1) and (2) show that an increase in the sex ratio of birth is associated with men marrying younger and less educated women (except in column (4)), although the coefficients are not significant. For women, the age of the husband increases as women become more scarce but the interaction coefficient is only significant when the sex ratio at birth is calculated at the state-year level (column (1)). Since men typically marry younger women, the age effects in Tables 6 and 7 suggest that an increase in the sex ratio at birth is associated with a larger spousal age gap.

3.3 Effects on Labor Force Participation

Lastly, I compare the effects of changes in the sex ratio at birth on labor force participation rates of men and women using the NSS data. Column (1) in Table 3.8 shows that women are relatively less likely to participate in the labor force when the sex ratio at birth increases. Men, on the other hand, are more likely to be in the labor force. These effects could be driven by the effects on educational attainment shown earlier. It is also possible that women's improved position on the marriage market increases their bargaining power within household leading to reduced labor supply (and increased leisure) for them. In columns (2) - (5), I find that this result holds for Hindu men and women. Muslim women also experience a relative decrease in labor force participation, but unlike Hindus, the effect on men is not significant. Among Hindus, all caste groups exhibit the same pattern, but the effects are the largest for SC individuals. Endogamous sex ratios at birth yield qualitatively similar results, but the coefficients are not significant and are available upon request.

4 Conclusion

This paper examines how imbalanced sex ratios differentially affect pre-marital investments in schooling, marital matching, and labor force participation rates of Indian men and women. I exploit the variation in the sex ratio at birth caused by changes in the availability of technology for prenatal sex-selection along with differential exposure due to endogamous marriage markets to estimate these effects. My results suggest that a relative scarcity of women improves their position in the marriage market as reflected in decreased pre-marital investments in education, lower age at marriage, older spouses, and lower labor force participation rates, relative to men. On the other hand, there is an increase in the educational attainment (albeit insignificant), age at marriage, and the labor force participation rate, and a decrease in wife's age for men, suggesting that the more abundant sex invests more pre-maritally in order to attract a match and also fares worse in terms of the outcomes on the marriage market as well as post-marriage.

However, this paper is not without its limitations. First, in constructing the sex ratio at birth, I have ignored the fact that, typically, men marry younger women. To the extent that expectations about the marriage market are an important explanatory mechanism for my findings, it may be better to construct sex ratios by taking into account the average spousal age gap. Second, in the absence of better data, the sex ratios have been constructed using retrospective birth data. Any sex-specific differentials in mortality and migration rates are likely to introduce measurement error in the key variable of interest. In future work, I plan to construct more robust measures of marriage market sex ratios to address these concerns. Lastly, although the results presented in this paper paint a consistent picture, it is not a complete or perfectly causal picture. A better understanding of the effects of large-scale population imbalances also requires an examination of the effects on other outcomes and margins of adjustment, such as dowries, fertility, and expenditure patterns, among others.

References

Abramitzky, R., A. Delavande, and L. Vasconcelos (2011): "Marrying Up: The Role of Sex Ratio in Assortative Matching," *American Economic Journal: Applied Economics*, 3, 124–157.

Angrist, J. (2002): "How do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *Quarterly Journal of Economics*, 117, 997–1038.

Anukriti, S. (2013): "The Fertility-Sex Ratio Trade-off: Unintended Consequences of Financial Incentives," .

Bhalotra, S. and T. Cochrane (2010): "Where Have All the Young Girls Gone? Identification of Sex Selection in India," *IZA Discussion Paper No. 5381.*

Edlund, L. (1999): "Son Preference, Sex Ratios and Marriage Patterns," *The Journal of Political Economy*, 107, 1275–1304.

Hu, L. and A. Schlosser (2011): "Prenatal Sex Selection and Girls' Well-Being: Evidence from India," *Unpublished manuscript.*

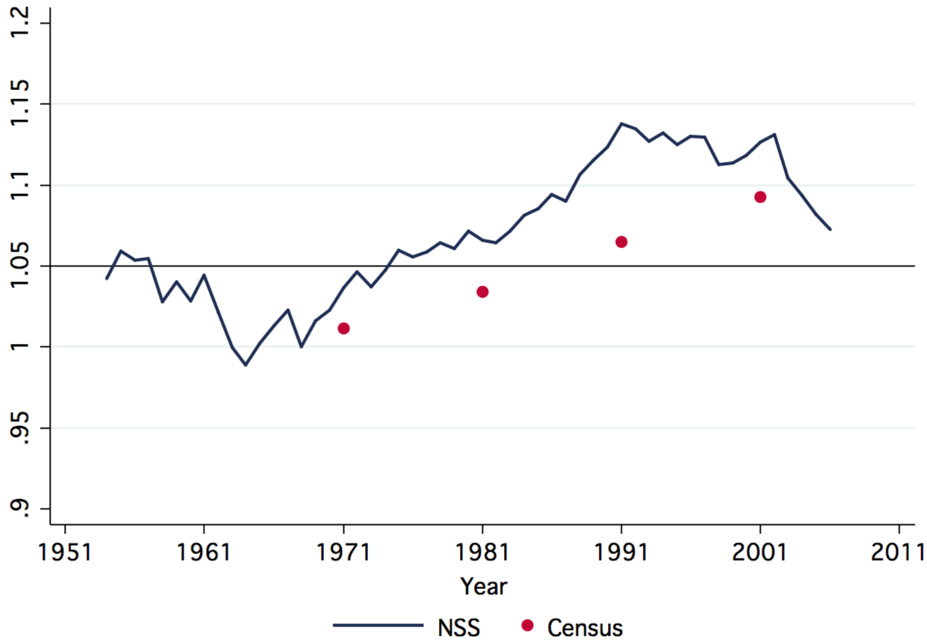
Jensen, R. (2005): "Equal Treatment, Unequal Outcomes? Generating Sex Inequality through Fertility Behavior," *unpublished manuscript, Harvard University.*

Lafortune, J. (2011): "Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market," *American Economic Journal: Applied Economics (forthcoming).*

Lin, M.-J., J.-T. Liu, and N. Qian (2010): "More Missing Women, Fewer Dying Girls: The Impact of Abortion on Sex Ratios at Birth and Excess Female Mortality in Taiwan," *NBER Working Paper 14541.*

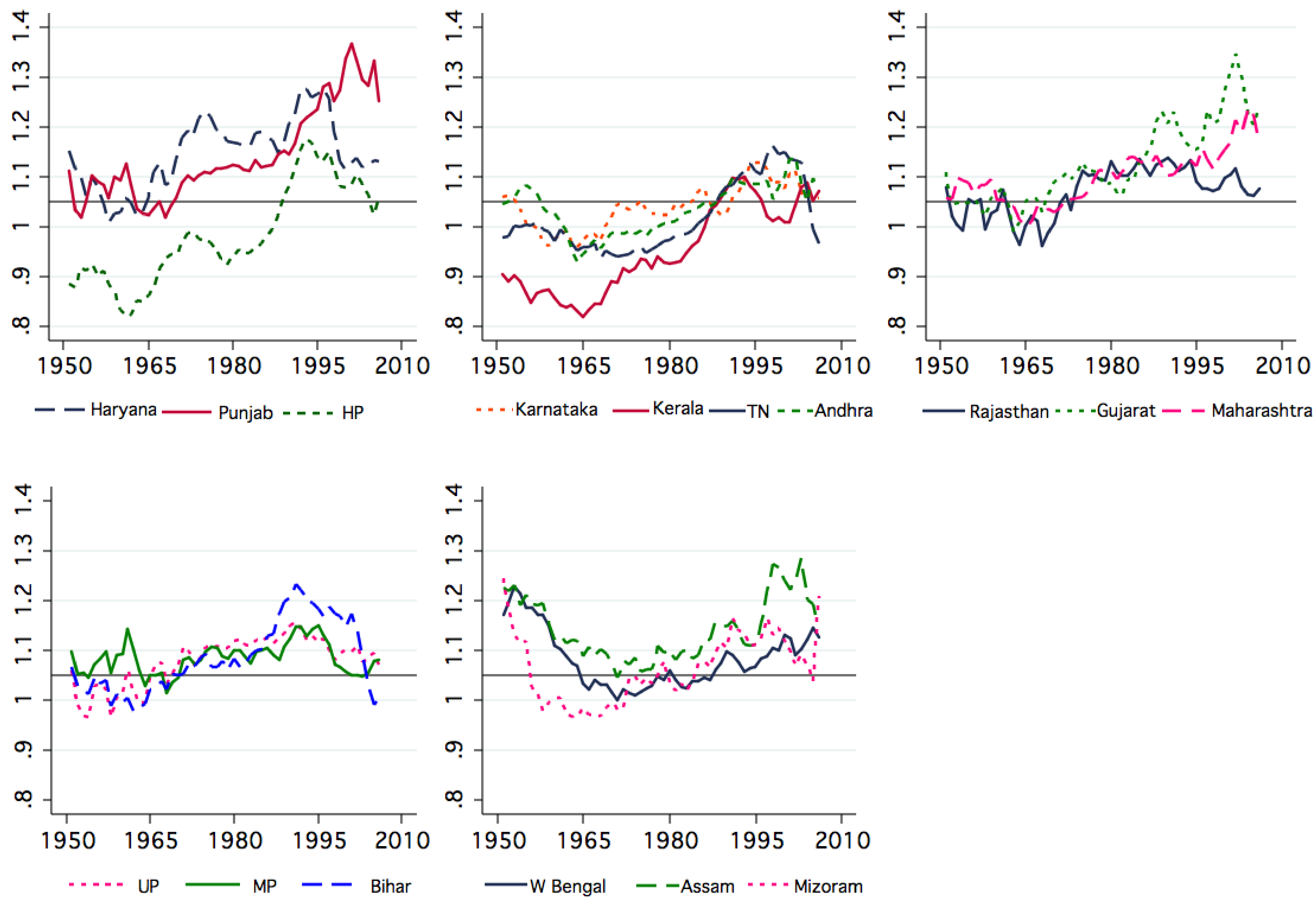
5 Figures

Figure 1: Trends in the Sex Ratio at Birth, India



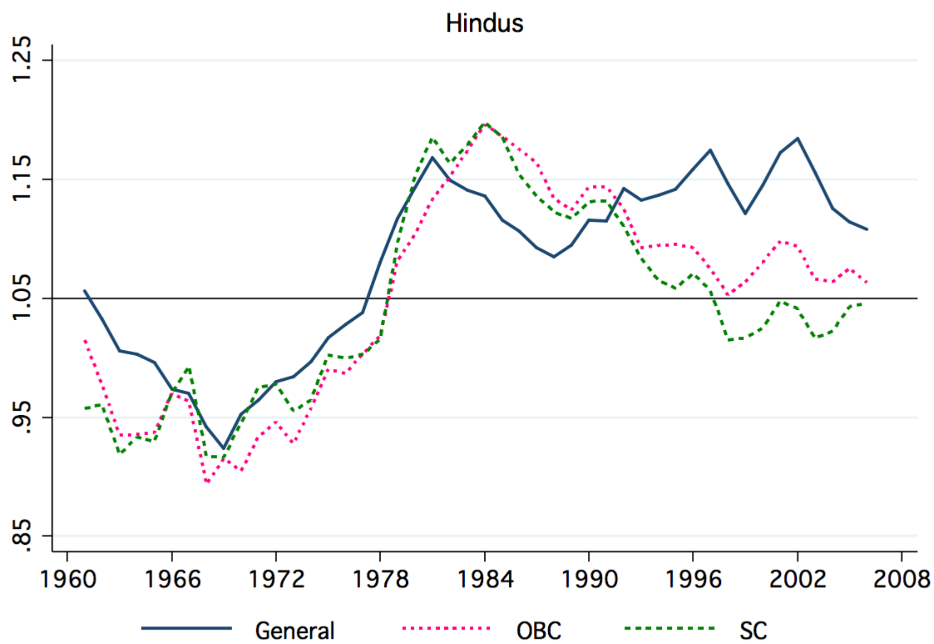
NOTES: NSS plots the 7-year moving average of the sex ratio at birth constructed from pooled NSS data. Census refers to the sex ratio at age 0 using single year age returns from the Census of India. The horizontal line reflects the natural male-female sex ratio at birth, equal to 1.05.

Figure 2: Trends in the Sex Ratio at Birth, by State



NOTES: This graph plots the 7-year moving average of the sex ratio at birth constructed from pooled NSS data. HP is Himachal Pradesh, TN is Tamil Nadu, Andhra is Andhra Pradesh, UP is Uttar Pradesh and includes Uttaranchal, MP is Madhya Pradesh and includes Chhattisgarh, W Bengal is West Bengal, and Bihar includes Jharkhand. The horizontal line reflects the natural male-female sex ratio at birth, equal to 1.05.

Figure 3: Trends in the Sex Ratio at Birth for Hindus, by Caste



NOTES: This graph plots the 7-year moving average of the sex ratio at birth constructed from pooled NSS data. SC stands for Scheduled Castes and OBC represents Other Backward Classes. Only rounds 55, 61, and 66 have been used since earlier rounds do not distinguish between OBC and General castes. The horizontal line reflects the natural male-female sex ratio at birth, equal to 1.05.

6 Tables

Table 1: Summary Statistics, NSS

Round	**38th**	**43rd**	**50th**	**55th**	**61st**	**66th**	**All**
Year of survey	1983	1987-88	1993-94	1999-00	2004-05	2009-10	
N	574,259	607,327	514,281	536,013	543,176	411,737	3,186,793
Hindu	0.791	0.788	0.803	0.784	0.766	0.766	0.784
Muslim	0.138	0.137	0.116	0.139	0.137	0.141	0.135
Sikh	0.024	0.026	0.024	0.024	0.027	0.022	0.025
Christian	0.043	0.045	0.052	0.048	0.067	0.069	0.053
Scheduled Caste	0.148	0.139	0.149	0.156	0.160	0.159	0.151
Scheduled Tribe	0.094	0.098	0.100	0.101	0.120	0.122	0.105
Rural	0.336	0.331	0.369	0.374	0.338	0.615	0.383
Sex Ratio at birth	1.048	1.052	1.062	1.074	1.081	1.084	1.066
Age at survey	21.910	22.409	23.279	23.581	24.203	25.435	23.353
Female	0.485	0.483	0.481	0.484	0.487	0.485	0.484
Children: 5-14 years							
N	76,294	77,076	61,528	66,943	66,466	47,561	395,868
Female	0.470	0.465	0.463	0.470	0.471	0.464	0.467
Attends school (for girls)	0.479	0.573	0.705	0.749	0.842	0.888	0.688
Attends school (for boys)	0.686	0.743	0.831	0.841	0.899	0.908	0.808
Adults: 15-60 years							
N	337,650	368,112	326,137	342,474	355,508	283,067	2,012,948
Female	0.490	0.489	0.487	0.489	0.494	0.493	0.490
Females							
Married	0.850	0.836	0.820	0.808	0.798	0.782	0.816
Literate	0.361	0.422	0.502	0.558	0.615	0.702	0.521
Primary or above	0.279	0.333	0.409	0.468	0.518	0.696	0.444
Secondary or above	0.085	0.121	0.175	0.216	0.226	0.321	0.187
In labor force	0.302	0.289	0.337	0.320	0.367	0.247	0.312
Employed	0.948	0.947	0.956	0.956	0.939	0.945	0.948
Males							
Married	0.693	0.682	0.670	0.662	0.659	0.651	0.670
Literate	0.636	0.692	0.744	0.773	0.818	0.864	0.751
Primary or above	0.514	0.572	0.629	0.670	0.717	0.860	0.654
Secondary or above	0.186	0.238	0.310	0.346	0.358	0.459	0.312
In labor force	0.862	0.849	0.846	0.837	0.842	0.813	0.842
Employed	0.951	0.949	0.962	0.955	0.954	0.971	0.957

NOTES: The sex ratio at birth refers to the the 7-year moving average of the sex ratio at birth constructed from pooled NSS data. The variable employed is defined only for individuals in the labor force.

Table 2: Effects on Educational Outcomes, using state-year SRB

	Children: 5-14 years		**Adults: 15-60 years**					
	Attends School		*Literate*		*Primary and above*		*Secondary and above*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SRB * Female	-0.2991***	-0.2892***	-0.1428**	-0.1321**	-0.1259*	-0.1145*	0.0076	0.0077
	[0.0934]	[0.0946]	[0.0650]	[0.0627]	[0.0672]	[0.0645]	[0.0392]	[0.0383]
SRB	0.1232	0.1090	0.0063	-0.0067	-0.0196	-0.0312	0.0156	0.0128
	[0.0754]	[0.0677]	[0.0397]	[0.0366]	[0.0427]	[0.0393]	[0.0440]	[0.0447]
N	799,273		2,012,946		2,012,946		2,012,946	
State FE	x	x	x	x	x	x	x	x
YOB FE	x	x	x	x	x	x	x	x
State-Sex FE	x	x	x	x	x	x	x	x
Sex-YOB FE	x	x	x	x	x	x	x	x
Covariates		x		x		x		X
State-specific trends	x	x	x	x	x	x	x	x
Age-Sex FE		x		x		x		x

NOTES: This table reports the coefficients β and γ from specification (1) using data from the NSS. *SRB* stands for the male-female sex ratio at birth. *Female* is equal to one if the person is a female, and zero otherwise. YOB indicates year of birth. Age refers to age at the time of the survey. Each column corresponds to a different regression. Robust standard errors are in brackets are have been clustered at the state level. Covariates are indicators for residence in a rural area, household's religion (Hindu, Muslim, Sikh, Christian), and caste (SC and ST). *** 1%, ** 5%, * 10%.

Table 3: Effects on Educational Outcomes, using Endogamous SRB

	Children: 5-14 years		**Adults: 15-60 years**					
	Attends School		*Literate*		*Primary and above*		*Secondary and above*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SRB * Female	-0.0033*	-0.0030*	-0.0007	-0.0007	-0.0010	-0.0011	-0.0017	-0.0018
	[0.0017]	[0.0016]	[0.0017]	[0.0016]	[0.0025]	[0.0026]	[0.0026]	[0.0025]
SRB	0.0012	0.0003	0.0058*	0.0047**	0.0062	0.0057	0.0037	0.0029
	[0.0027]	[0.0025]	[0.0033]	[0.0022]	[0.0042]	[0.0034]	[0.0029]	[0.0025]
N	353,580		954,574		954,574		2,012,946	
State - YOB FE	x		x		x		x	
Caste - YOB FE	x		x		x		x	
State-specific trends		x		x		x		x
Caste-specific trends		x		x		x		x

NOTES: This table reports the coefficients β and γ from specification (2) using data from the NSS. *SRB* stands for the male-female sex ratio at birth. *Female* is equal to one if the person is a female, and zero otherwise. YOB indicates year of birth. Age refers to age at the time of the survey. Each column corresponds to a different regression and controls for state, YOB, caste, age fixed effects, indicator for residence in a rural area, and interaction fixed effects for sex with state, caste, YOB, and age. Robust standard errors are in brackets are have been clustered at the state level. Covariates are indicators for residence in a rural area, household's religion (Hindu, Muslim, Sikh, Christian), and caste (SC and ST). *** 1%, ** 5%, * 10%.

Table 4: Effects on Educational Outcomes, by Religion

	Hindu (1)	**Muslim** (2)	**Sikh** (3)	**Christian** (4)
	A. Children: 5-14 years			
Attends school				
SRB * Female	-0.3683***	0.0820	-0.1991	0.0116
	[0.0917]	[0.1538]	[0.2000]	[0.0467]
SRB	0.1304*	-0.1170	0.0585	0.0064
	[0.0744]	[0.1130]	[0.0978]	[0.0438]
N	617,735	120,624	18,363	40,041
	B. Adults: 15-60 years			
Literate				
SRB * Female	-0.1749**	-0.1465	-0.2460*	-0.0386
	[0.0721]	[0.1629]	[0.1424]	[0.0864]
SRB	0.0189	0.0043	-0.0291	0.0120
	[0.0439]	[0.0642]	[0.0736]	[0.0432]
Primary and above				
SRB * Female	-0.1394*	-0.1122	-0.3685*	-0.0216
	[0.0744]	[0.1315]	[0.2071]	[0.0939]
SRB	-0.0081	-0.0639	0.0123	-0.0203
	[0.0450]	[0.0819]	[0.0937]	[0.0718]
Secondary and above				
SRB * Female	-0.0112	0.1116**	-0.2940	0.0265
	[0.0392]	[0.0499]	[0.2437]	[0.0712]
SRB	0.0116	0.0603	-0.0898	0.0419
	[0.0507]	[0.0600]	[0.0930]	[0.0598]
N	1,589,572	251,255	52,064	111,482

NOTES: This table reports the coefficients β and γ from specification (1). SRB stands for the male-female sex ratio at birth. $Female$ is equal to one if the child is a female, and zero otherwise. Each column corresponds to a different regression. The control variables included are same as in column (2) of Table 3.3.2, except religion indicators. Robust standard errors are in brackets are have been clustered at the state level. *** 1%, ** 5%, * 10%.

Table 5: Effects on Educational Outcomes among Hindus, by Caste

	SC	**ST**	**OBC**	**General**	**General/OBC**
	(1)	(2)	(3)	(4)	(5)
			A. Children: 5-14 years		
Attends school					
SRB * Female	-0.4717***	-0.2243	-0.2108**	-0.0964	-0.3586***
	[0.1107]	[0.1464]	[0.0988]	[0.0576]	[0.0895]
SRB	0.0383	0.1497	0.0148	0.0744	0.1440*
	[0.1175]	[0.1226]	[0.0961]	[0.0526]	[0.0788]
N	121,373	55,518	110,744	78,387	443,354
			B. Adults: 15-60 years		
Literate					
SRB * Female	-0.4380***	-0.3986**	-0.3021**	0.1989**	-0.1109
	[0.1320]	[0.1639]	[0.1139]	[0.0735]	[0.0668]
SRB	0.0717	0.0751	-0.0789	-0.1101***	0.0168
	[0.0805]	[0.0777]	[0.0710]	[0.0355]	[0.0418]
Primary					
SRB * Female	-0.3091**	-0.2312*	-0.2931**	0.2074***	-0.1136*
	[0.1227]	[0.1346]	[0.1237]	[0.0629]	[0.0629]
SRB	0.0155	0.051	-0.0697	-0.0806**	0.0019
	[0.0734]	[0.0620]	[0.0755]	[0.0362]	[0.0452]
Secondary					
SRB * Female	-0.0121	-0.0251	-0.0464	0.1667***	-0.0232
	[0.0499]	[0.0485]	[0.0774]	[0.0556]	[0.0386]
SRB	0.0037	0.0173	-0.0135	0.0039	0.0277
	[0.0467]	[0.0297]	[0.0956]	[0.0532]	[0.0616]
N	281,648	130,525	303,330	263,666	1,185,972

NOTES: This table reports the coefficients β and γ from specification (1). The sample is restricted to Hindus. SRB stands for the male-female sex ratio at birth. $Female$ is equal to one if the child is a female, and zero otherwise. Each column corresponds to a different regression. The control variables included are same as in column (2) of Table 3.3.2, except religion and caste covariates. Robust standard errors are in brackets are have been clustered at the state level. *** 1%, ** 5%, * 10%.

Table 6: Effects on Age at First Marriage, NFHS-3 data

								Hindus		
	All		**Hindu**	**Muslim**	**Sikh**	**Christian**	**SC**	**ST**	**OBC**	**General**
Age at First Marriage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SRB * Female	-2.7012***	-0.0122	-2.3457**	-3.0455	-8.3238	-3.2593**	-3.9786*	-0.8886	-1.3619	-2.3825
	[0.6113]	[0.0892]	[0.9511]	[2.6053]	[6.2032]	[1.1745]	[2.0293]	[3.3064]	[1.4339]	[1.7407]
SRB	1.9623*	0.1250**	0.8398	1.231	8.4132	5.0999***	2.0441	-0.8508	-0.8074	1.0805
	[1.0061]	[0.0633]	[1.2128]	[1.9733]	[8.0530]	[0.6472]	[2.0937]	[3.7362]	[1.6217]	[1.4899]
N	129,670	116,083	97,492	16,394	2,588	9,408	19,850	7,485	37,323	32,834
Endogamous SRB		X								

NOTES: This table reports the coefficients β and γ from specification (1) using NFHS-3 data, except column (2) which reports results from specification (2). The sample is restricted to individuals older than 14 years and further restricted to only Hindus in columns (7) - (10). *SRB* stands for the male-female sex ratio at birth. *Female* is equal to one if the child is a female, and zero otherwise. Each column corresponds to a different regression. All covariates, interaction and main fixed effects are included in all regressions, except caste covariates in columns (2) - (6) and religion and caste covariates in columns (7) - (10). Robust standard errors are in brackets are have been clustered at the state level. *** 1%, ** 5%, * 10%.

Table 7: Effects on Match Quality

	Age of spouse		**Education of spouse**	
	(1)	(2)	(3)	(4)
SRB * Female	2.9873**	0.1451	-0.2182	0.1057
	[1.1419]	[0.1159]	[0.9331]	[0.1176]
SRB	-1.0355	-0.1101	-0.1562	0.0589
	[0.9340]	[0.0845]	[0.5859]	[0.1051]
N	123,365	112,557	126,448	115,309
Endogamous SRB		X		X

NOTES: Columns (1) and (3) report the coefficients β and γ from specification (1) and columns (2) and (4) report the same coefficients for specification (2) using NFHS-3 data. All covariates, interaction and main fixed effects are included in all regresions. The sample is restricted to individuals older than 14 years. *SRB* stands for the male-female sex ratio at birth. *Female* is equal to one if the child is a female, and zero otherwise. Robust standard errors are in brackets are have been clustered at the state level. *** 1%, ** 5%, * 10%.

Table 8: Effects on Labor Force Participation

						Hindus				
	All	**Hindu**	**Muslim**	**Sikh**	**Christian**	**SC**	**ST**	**OBC**	**General**	**General/OBC**
In LF = 1	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SRB * Female	-0.1885***	-0.2273***	-0.2706***	0.0500	0.0998	-0.3305***	-0.2222	-0.1802	-0.1573**	-0.2299***
	[0.0597]	[0.0700]	[0.0968]	[0.1388]	[0.1041]	[0.0737]	[0.1777]	[0.1289]	[0.0571]	[0.0691]
SRB	0.0798*	0.1031**	0.0132	-0.0443	0.0341	0.1381***	0.1952**	0.0715	0.1110*	0.1030**
	[0.0394]	[0.0465]	[0.0766]	[0.0971]	[0.0618]	[0.0475]	[0.0812]	[0.0854]	[0.0587]	[0.0447]
N	2,012,946	1,589,572	251,255	52,064	111,482	281,648	130,525	303,330	263,666	1,185,972

NOTES: This table reports the coefficients β and γ from specification (1). The sample is restricted to Hindus in columns (6) - (10). SRB stands for the male-female sex ratio at birth. $Female$ is equal to one if the child is a female, and zero otherwise. Each column corresponds to a different regression. The control variables included are same as in column (2) of Table 3.3.2, except caste covariate in columns (2) - (5) and religion and caste covariates in columns (6) - (10). Robust standard errors are in brackets are have been clustered at the state level. *** 1%, ** 5%, * 10%.

7 Appendix

Table 9: Effects on Educational Outcomes using NFHS data

	Children: 5-14 years		**Adults: 15-60 years**					
	Attends School		*Literate*		*Primary and above*		*Secondary and above*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SRB * Female	-0.3031**	-0.2853**	-0.1405**	-0.1322**	-0.1380**	-0.1296**	-0.1070	-0.0986
	[0.1162]	[0.1151]	[0.0653]	[0.0634]	[0.0645]	[0.0626]	[0.0805]	[0.0792]
SRB	0.2314	0.2727*	-0.0490	-0.0612	-0.0501	-0.0620	-0.0411	-0.0429
	[0.1560]	[0.1519]	[0.0487]	[0.0393]	[0.0485]	[0.0392]	[0.0532]	[0.0444]
N	315,006		830,991		830,991		830,991	
State FE	x	x	x	x	x	x	x	x
YOB FE	x	x	x	x	x	x	x	x
State-Sex FE	x	x	x	x	x	x	x	x
Sex-YOB FE	x	x	x	x	x	x	x	x
Covariates		x		x		x		X
State-specific trends	x	x	x	x	x	x	x	x
Age-Sex FE		x		x		x		x

NOTES: This table reports the coefficients β and γ from specification (1). SRB stands for the male-female sex ratio at birth. $Female$ is equal to one if the person is a female, and zero otherwise. YOB indicates year of birth. Age refers to age at the time of the survey. Each column corresponds to a different regression. Robust standard errors are in brackets are have been clustered at the state level. Covariates are indicators for residence in a rural area, household's religion (Hindu, Muslim, Sikh, Christian), and caste (SC and ST). *** 1%, ** 5%, * 10%.