ISER Working Paper Series

Population sex ratios and violence against women: The long-run effects of sex selection in India

Sofia Amaral
Centre for Crime, Justice and Policing
University of Birmingham

Sonia Bhalotra
ISER & Department of Economics
University of Essex

No. 2017-12
October 2017

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

University of Essex

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

www.iser.essex.ac.uk

Non-Technical Summary

Violence against women is major public health and social problem that is estimated to affect one in three women globally. This type of violence is arguably hindering social and economic development and its repercussions for long-term development are large.

In this paper, we investigate the role played of sex imbalance in India's population for violence against women. In India, the demographic imbalance is large and has risen since the introduction of ultrasound technology in the 1980's. A technology that allowed parents to sex-select births of boys in detriment to that of girls. The consequences are visible. For instance, the 2011 Census of India reveals that there were 37 million more men than women in the population, up from about 20 million in 1971. In this time, the youth sex ratio rose from 99.22 to 108, and this led to disruption in the relative availability of grooms to brides in the population. For example, in 2005 is was estimated that there is a surplus of 5% of potential grooms in the population. In these same forty years, although total recorded crime in India fell, violence against women increased and is now the fastest growing crime category. In this paper, we provide the first causal evidence of the relationship between sex ratio and violence against women. To do so, we have assembled a novel database of district-level administrative crime data and census data across over the period between 1971-2011.

Our findings show that a surplus of males in the crime-prone and marriageable ages increase violence crime against women without any effects on other crimes. We also attempt to explain why we find this relationship. We first show that increases in sex ratio increase single rates for men. Second, we show that attitudes to violence against women are less equal for those exposed to a higher sex ratio at birth. Finally, we show that marriage quality measures, including self-reported domestic violence, are negatively related to sex ratios.

Population Sex Ratios and Violence against Women: The Long Run Consequences of Sex Selection in India *

Sofia Amaral
University of Birmingham†

Sonia Bhalotra
University of Essex*

December, 2017

Abstract

This paper investigates the consequences of sex imbalance in India's population for violence against women. We match administrative crime data by category to age-specific sex ratios in census data by district across four decades and, to analyse mechanisms, we also use census data on marriage rates and, household survey data on attitudes to violence against women and marriage quality measures. We find that a surplus of men at age 20-24 increases crimes committed against women, and that this explains about 21% of the rise in gender-based violence. Although less robust, there is some evidence that the youth sex ratio also raises non-gendered forms of violence, but we find no discernible impact upon property and economic crime. In probing mechanisms, we show that marriage rates, marriage quality and attitudes to violence against women are all modified by population sex ratios.

Keywords: sex ratio, violence against women, marriage market, crime, gender attitudes

JEL Classification: J12, J16, N34, K42

*Centre for Crime, Justice and Policing, University of Birmingham. Corresponding author: Sofia Amaral (amaral.sofiafernando@gmail.com)

†ISER & Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, Contact: srbhal@essex.ac.uk

*We thank seminar participants at CESifo, Bath, Royal Holloway Workshop on Domestic Violence, Reading, Leicester, Kent, European Economic Association (Lisbon), European Society for Population Economics (Glasgow), Population Association of America (Chicago), Indian Statistical Institute (Delhi) and ISER for their comments. We also appreciate comments from Helmut Rainer, Espen Villanger, Andreas Kotsadam, Sheetal Sekhri, Abhiroop Mukhopadhyay and Siddhartha Bandyopadhyay. This work was partially supported by the MiSOC Economic and Social Research Council grant ES/L009153/1. Sofia also acknowledges funding from the Royal Economic Society, Portuguese Association for Researchers in the UK and the Fundação para a Ciência e Tecnologia grant SFRH/BD/78034/2011.

1 Introduction

Violence against women (henceforth VAW) is a major public health and social problem that is estimated to affect one in three women globally (Ellsberg et al., 2008). The annual costs of violence against women regarding medical care and productivity shortfalls are significant, for instance, they are estimated to exceed an annual $5.8 billion in the U.S. (Aizer, 2010). In addition, VAW has adverse intergenerational consequences, impacting upon the survival, health and education of children of victimized women (Papageorge et al., 2016; Menon, 2015; Aizer and Currie, 2014; Carrell and Hoekstra, 2010; Rawlings and Siddique, 2014). Civil society and scholarly interest in this phenomenon are on the rise, one reflection of which is that Goal 5.2 of the Sustainable Development Goals released in September 2015 is to "eliminate all forms of violence against all women and girls in public and private spheres". This calls for a clearer understanding of the causes of violence against women.

In this paper, we investigate violence against women in India as a function of the population sex ratio, i.e. the ratio of males to females. Several studies have documented the deficit of girls in India's population (Sen, 1992; Coale, 1991; Klasen, 1994; Anderson and Ray, 2010; Hvistendahl, 2012). Importantly, in India and also in China, the fraction of missing girls at birth has been rising since the widespread availability of prenatal sex detection technology and has come to dominate the rise in the all-age population sex ratio (Sudha and Rajan, 1999; Bhalotra and Cochrane, 2010; Chen et al., 2013; Bongaarts and Guilmoto, 2015).

The average imbalance is significant, for instance, the 2011 Census of India reveals that there were 37 million more men than women in the population, up from about 20 million in 1971. In this time the youth sex ratio rose from 99.22 to 108. Although total recorded crime in India fell during this period, violence against women increased and has been the fastest growing crime category. During 1995 to 2010, a period over which comparable data are available, VAW increased by 34%. In this period, VAW has constituted 15% of all crimes in India and 27% of violent crimes. [1]

We matched district-level census data on age-specific sex ratios and marriage rates to district-level time series administrative data on crime rates by type of crime to create data from 1972 to 2010. The independent variable of interest is the youth sex ratio, in particular, the ratio of men

[1] Rape is the one category of VAW that has been recorded since 1972, and this shows a rise from 0.5 to 1.2 crimes per 100,000 between 1972 and 2010.

to women age 20-24, which is the age at which crime peaks (Theobald and Farrington, 2011).[2] The outcome variable of interest is violence against women, but we estimate identical models for other forms of non-gender based violent crime, and for non-violent crime.

Previous research suggests that economic factors such as income and unemployment shocks often influence property or economic crime but not violent crime (Ciccarelli, 2017; Freedman and Owens, 2016; Machin and Meghir, 2004; Gould et al., 2002). In India, Iyer and Topalova (2014); Blakeslee and Fishman (2014); Fetzer (2014) have similarly found evidence of a positive relationship between income (as measured by rainfall) and property and violent crime, but not VAW. Some recent work highlights the relevance of non-economic factors - such as emotional cues, loss of self-control- in determining violent crime, including violence against women (Blattman et al., 2015; Card and Dahl, 2011; Dahl and DellaVigna, 2009). However, as discussed below, there is a relative scarcity of research on the determinants of violence against women and on the population sex ratio, and marriage market distortions as a causal factor.

Exploiting within-district variation over time, we find that a one standard deviation increase in the sex ratio of the 20-24-year-old population (which is 0.11) is associated with an increase in violence against women that ranges between 12%-16% an elasticity of 0.74. Our estimates indicate that the surplus of young men explains more than a fifth of the rise in gender-based violence since 1995[3]. This result is robust to a number of specification checks. We find a similar elasticity when we study the impact of the youth sex ratio on non-gendered violence, although this is more sensitive to specification. We find no evidence that the youth sex ratio influences acquisitive crime, which includes property and economic crime. Our specifications control for economic and demographic determinants of crime, and for cultures of gendered preferences that may drive a spurious association between the sex ratio and violence against women.

Among mechanisms that might explain our findings are the following. First, men and especially younger men are more likely than women to commit crime (Gavrilova and Campaniello, 2015), so an increasing share of men among youth will mechanically lead to more crime, but potentially to more crime of all sorts. A second reason is that a surplus of men among youth will tend to reduce marriage rates, and there is evidence that unmarried men are more likely

[2]We verify this using the distribution of crime by the age of perpetrator in the Indian crime data too. The precise age range reflects merely the age groupings available in the Indian census. We investigate robustness by including in the equation the sex ratio in neighbouring age groups. We show that the age 20-24 sex ratio reasonably closely tracks its twenty-year lag, the age 0-4 sex ratio, even if there is some modification on account of sex-differentiated mortality after the age of four, and sex-differentiated migration.

[3]Since the sex ratio of 20-24 rose by 8.7% over the period and VAW rose by 34% since 1995.

to commit crime (Visher et al., 2009; Korenman and Neumark, 1991; Akerlof, 1998; Sampson et al., 2006; Siennick et al., 2014). It seems plausible that a rising share of unmarried men will have larger impacts on VAW (which includes rape) than on economic crime, but we are unaware of previous work investigating the association of marriage or fatherhood with violence against women in particular.

Since 44% of violence against women in our sample is domestic violence (committed by men within marriage), a third potential channel is that the quality of marriage is decreasing for women (Hussam, 2016). One may imagine that the growing relative scarcity of women would improve their marriage market outcomes. However, population growth combined with an age gap at marriage (of about five years) raises the supply of potential brides and this, potentially offsets the increased maleness of the sex ratio; see Bhaskar (2011); d'Albis and De La Croix (2012); Rao (2000).[4] A fourth possible channel, particular to crimes against women, is that the attitudes of men towards violence against women are shaped by the sex ratio of their birth cohort.[5] We present some supporting evidence for each of these four possible channels.

Our paper is most closely related to Edlund et al. (2013). Edlund et al. (2013) use data from thirty Chinese provinces through 1988-2004, and document that the increasing share of males among youth is associated with increasing rates of (total) crime, defined as arrests for violent and property crime. The main difference between their study and this paper is that we focus upon crimes against women and, of course, on India.[6] This is a substantively different question and, as discussed, we find no impacts of the youth sex ratio on property crime.[7] We also exploit mechanisms that have not been previously considered in the literature (e.g. attitudes formation).

Our first contribution is to a literature on violence against women. A small literature in economics has analysed how violence against women varies with income shocks (Abiona and Koppensteiner, 2016; Sekhri and Storeygard, 2014), the relative income or employment of women (Aizer, 2010; Anderberg et al., 2016; Bhalotra et al., 2016c), transfers programmes (Angelucci, 2008; Hidrobo et al., 2013; Bobonis et al., 2013; Ramos, 2016) and historical attitudes to women (Tur-Prats, 2015; Alesina et al., 2016). We present the first evidence of the importance of the

[4] Bhaskar (2011) contrasts this with China.

[5] Reasons for this association that we discuss are the intergenerational transmission of gender preferences and exposure to sisters.

[6] Data on crimes against women were not available at the province level for China.

[7] In contrast to India, violent and property crime has been rising in China while crime against women (as measured by rape) has been declining (Edlund et al., 2007).

male to female ratio at age 20-24 in generating violence against women. Our findings have their most direct bearing on India, China and other South and South East Asian countries where the population sex ratio is skewed in favour of men on account of gendered social preferences. However, they have wider relevance on account of a different phenomenon inflicting OECD countries, which is the rising share of unmarried men, which raises the potential for increased crime against women (Akerlof, 1998; Dorn and Hanson, 2015).

Our second contribution is to the literature on the causes and consequences of increasing male-bias in the sex ratio in India and China. Most research in this area investigates causes (Qian, 2008; Ebenstein and Sharygin, 2009; Ebenstein, 2011; Almond et al., 2013; Pino et al., 2016; Anukriti et al., 2016a, 2017; Bhalotra et al., 2016a; Pino et al., 2016). Although there is some evidence on the long run consequences of sex selection in China (Edlund et al., 2013; Ebenstein and Sharygin, 2009; Rao, 2000) ours is one of the few studies to provide evidence for India, where many features of the landscape are different including the absence of the One Child Policy and weaker enforcement of women's rights.[8]

The rest of this paper is organized as follows. Section 2 discusses the empirical strategy. Section 3 describes the data. Section 4 presents the results and, section 5 discusses potential mechanisms. Section 6 concludes.

2 Empirical Specification

The hypothesis of interest is that violence against women (VAW) is increasing in the male-biasedness of the youth sex ratio. Since previous research (mentioned above) suggests that men and, especially, young and unmarried men, are more prone to crime in general, we estimate a series of equations, for all violent and acquisitive crime, and for sub-categories of crime (defined in Table 11). We use the following difference-in-differences specification, where we exploit within district variation in these variables over time, across five decades and 258 districts. The specification is similar to that in Edlund et al. (2013).

$$Crime_{dst} = \gamma_1 SR_{dst}^{20-24} + \pi_1 SR_{dsat}^{0-4} + \beta' X_{dst} + \lambda_d + \omega_t + g_d t + \epsilon_{dst} \quad (1)$$

[8]There is an interesting body of theoretical work discussing consequences. See Becker (2009); Chiappori et al. (2002); Angrist (2002); Abramitzky et al. (2011); La Mattina (2016). When parents have intrinsic son preference, marriage market outcomes may be more nuanced. See Edlund (1999); Bhaskar (2011).

where $Crime$ is the rate of crimes per capita (in logarithm) in a district-state-year (dst) and SR denotes the age-specific ratio of men to women. The coefficient of interest is γ_1, which captures the effect of the youth sex ratio (age 20-24) on crime. This is an age range in which crime rates tend to peak and the period at which the share of unmarried men is high. To the extent that most of the variation in the youth sex ratio is determined in early childhood, in any year in our sample, the youth sex ratio is effectively lagged 20 to 24 years concerning the year in which crime is measured. We control for the early childhood sex ratio to allow this to proxy active son preference - as parents manipulate the sex ratio in this age group when they have more girls than they desire (Anukriti et al., 2016a). Since a male surplus in the 0-4 age group clearly has no direct impact on crime, this control allows us to distinguish the mechanisms linking a youth male surplus to crime against women from underlying son preference. The equation includes district fixed-effects (λ_d) that absorb unobserved persistent determinants of violence against women. Regional variation in predictors of crime have been highlighted by, for instance, Dyson and Moore (1983); Agarwal (1994); Alesina et al. (2016); Tur-Prats (2015). Year fixed-effects (ω_t) control flexibly for shocks common to all districts in India such as world movements in gold prices that could affect VAW (Menon, 2015).[9]

We also include district and state level covariates that control for region-year variation in demographic and economic factors that are potentially predictive of crime, and that may also influence migration by sex and hence the youth sex ratio. The district level controls include the literacy rate, the gender gap in the literacy rate (the log ratio), the share of the population in rural areas, rainfall, and the share of the population identified as belonging to scheduled castes and tribes. In a robustness check, we control for district-level population growth to account for the fact that differential population growth paths affect the ratio of potential brides to grooms (Edlund, 1999; Bhaskar, 2011). The state-level covariates (not available at the district level over time) are per capita income, police personnel per capita and state election dummies that allow for a potential electoral cycle in crime (Levitt, 2002; Ghosh, 2006). Importantly, district rainfall and state income will together control for economic changes; rainfall being an important predictor of income in India (Virmani, 2006).[10] We nevertheless account for economic conditions by controlling for district-level rainfall and state income. To allow that we nonetheless do not

[9] Menon (2015) shows that gold prices at the time of marriage influence domestic violence. Bhalotra et al. (2016a) show that gold prices in pregnancy and the birth month influence the sex ratio age 0-4, but this is effectively a 20-year lag on the youth sex ratio variable that we use. This is discussed further below.

[10] Unemployment rates are less useful given that most people cannot afford to be unemployed, leading to an array of casual and/or part time employment.

capture all potential district-time level variation that may confound the relationship of interest, we include district-specific linear trends ($g_d t$) and in a separate model, state-year fixed effects (θ_{st}). The last term in (1) is the stochastic error term, ϵ_{dst}. In our specifications, we clustered standard-errors at the district-decade level to account for possible correlation in the errors within the district over time (Burgess et al., 2013; Sekhri and Storeygard, 2014; Blakeslee and Fishman, 2014).

The skewness of the youth ratio (age 20-24) tracks the skewness of the child ratio (age 0-4) twenty years earlier. The main determinant of cross-district and within-district time variation in the 20-24 ratio is thus son-preferring behaviour enacted twenty years before, making it unlikely that the 20-24 ratio is correlated with contemporary determinants of crime.[11] What we wish to capture is the excess of young men in the crime and marriage markets conditional upon cultures of son preference.

We include the sex ratio among children age 0-4 so as to condition upon son preference cultures, known to vary across India's districts but to be evolving over time. There is a vast literature documenting that the sex ratio age 0-4 proxies for son preference, involving female-biased abortion and under-investment in girls in early childhood, see for instance, (Miller, 1981; Sen, 2003; Bhalotra and Cochrane, 2010).[12] It provides a falsification test as children of this age do not commit crime.

We further investigate inclusion of the sex ratio of ages bordering age 20-24 (i.e. age 15-19 and 25-34) to examine sensitivity to our choice of age. We also include sex ratios for ages 35 and older. Since the share of unmarried men falls by this age and crime proneness also falls, we interpret these as further checks on the mechanisms driving the relationship of crime with the surplus of 20-24 year old men. In other words, we distinguish the effects of a general population surplus of males from a surplus of those with a proclivity to engage in crime. Finally,

[11] The youth sex ratio will evolve from the child sex ratio on account of differences in migration and mortality after the age of 4 by gender. Mortality falls sharply after the age of five, and any sex differences in mortality after age 4 will tend to follow the same patterns as male-female mortality age 0-4, son preference being the systematic determinant of both. The 2011 Indian census records gross migration rates across districts for men and women of 4.5% and 10%, respectively (Appendix D-3 of the Census 2011 publications, Census of India (2011)). Sex differences in migration arise as men migrate towards job opportunities whereas women migrate primarily for marriage(Mumford et al., 2000). This will only matter for identification if sex-differences in migration are correlated with crime. We expect the controls to adjust for this.

[12] (Bhalotra and Cochrane, 2010) demonstrate a sharp increase in the share of males at birth following the first imports of ultrasound scanners in the mid-1980s, and a further increase after delicensing led to local production in the mid-1990s. These authors cite documentary evidence of the widespread availability of prenatal sex detection at low cost.

in the Appendix we discuss instrumenting the youth sex ratio with the with the availability of ultrasound, which Bhalotra and Cochrane (2010) argue was driven by exogenous changes in import and licensing regulations. While our results are not significant due to a weak instruments problem, the coefficient of interest is positive. We shall conduct a robustness check in which we include older and younger age-groups that are likely to be similarly affected by crime but unequally affected by the marriage market.[13]

After presenting the main results we discuss alternative specifications and robustness checks on the baseline strategy.

3 Data

3.1 Crime Data

We use the yearly publications of the National Crime Records Bureau (NCRB) of the Government of India over the period of 1971-2011, and process the available crime area information to create a district-level data panel for 258 districts across the 16 major states of India. Districts are the lowest administrative units in India, similar to prefectures in China or counties in the United States.[14] The NCRB provides data for 18 categories of crime and using these we construct four different crime rates. These are violence against women, non-gender based violence, property crime and economic crime; see Table 11 for definitions.[15] To avoid outlier effects, all crime variables are winsorized at 1%.

Prior to 1988, the only measure of crime against women in Indian administrative data was rape. Kidnappings disaggregated by gender became available from 1988, and the other categories of VAW were included from 1995. Similarly, the information on other categories of violent crime (such as arson, hurt) was added after 1995. In the analysis, these jumps in the series will be accounted for by year fixed effects. Since the changes were federal they will not exhibit a correlation with district-year variation in the youth sex ratio.

[13] The crime data do not provide victimisation rates by age-group.

[14] Appendix A provides a description of how India gathers crime statistics of India and how we aggregated to the district-level across years.

[15] VAW includes domestic violence, rape, molestation, sexual harassment, kidnapping of women and girls. Non-gender based violence includes murder, riots, kidnapping of males, dacoity, arson and hurt. Property crime includes theft, robbery and burglary. Economic crime includes cheating, counterfeiting and criminal breach of contract. A detailed description of these categories can be found in the Indian Penal Code (Gandhi, 2010).

Figure 2 shows that gender-based crime has tended to increase over the period at least until quite recently (even after visually adjusting for the first and second spikes highlighting the incorporation of the new crime categories in 1988 and 1995 respectively). Over the period the rape rate (which is the one category consistently measured through the period) grew from 0.6 to 1.8 crimes per 100,000 population and gender-based violence is the fastest growing crime rate in the country, having grown from 11 to 14 in 1995-2010. On average there are 13 reports of crimes against women per 100,000 population between 1995-2010. In contrast, both non-gendered violence and acquisitive crime have been decreasing throughout the period. Acquisitive crime accounts for the largest share of crime, with a rate of 53 for property and 5.2 for economic crime, and 26 for non-gender based violence, all per 100,000 population (Table 14).

In addition to administrative changes in coding crime, there are trends in the reporting-behaviour of crime, especially in women reporting crimes against them. It is inherently challenging to disentangle reporting from incidence. However, using a variety of strategies, Iyer et al. (2012) present compelling evidence that the mandated increase in the share of women in local government in India has led to increased reporting by women of crimes against them. There is no evidence one way or the other of trends in reporting for other crimes. We control for crime-specific trends in reporting by estimating crime-specific models and including district trends and state-year fixed effects, in addition to district and year fixed effects and police coverage per capita. Let us nevertheless consider the likely direction of any bias. The possible direction is that women are less likely to report VAW in district-years in which the youth sex ratio is more male-biased. In this case, we would under-estimate the relationship of interest. Thus for under-reporting of VAW to drive our conclusions, it would need to be greater in district-years with less male-biased youth sex ratios, and it would need to evolve in a non-linear manner that exhibits district variation within states conditional upon police coverage. In a robustness exercise, we difference across VAW and property crime, and this specification allows us to control for crime-district fixed-effects and crime-year fixed effects which we expect mitigates the problem.

The NCRB does not publish district or state level crime data disaggregated by age and gender of the offender and/or victim. Consequently, we cannot study the age-gender profile of offenders and victims across districts and over time. However, we were able to identify some relevant descriptive information. First, there is evidence that in India, as elsewhere (Gavrilova and Campaniello, 2015; Cameron et al., 2016), men are more crime prone. According to the 2011 NCRB annual report, males commit 94% of total crime and 85% of VAW (see Table 12).

Second, there is evidence that in India, in line with other countries, crime rates peak among men under the age of 25 (see Table 12). The sociology, criminology and economics literature points to the crime-prone age being between 16-25 (Sampson and Laub, 2005; Nagin and Land, 1993). We do not have age 16-25 identified in the data but we can observe a few critical aspects related to the age-gender profile of offenders in India. These are presented in Table 12. First, total crime is mostly committed by males as the share of crimes committed by men is 94%. Second, men commit the vast majority of VAW with the percentage of VAW crimes committed by males being 85%. When it comes to age, those aged between 18-30 are the primary crime-prone group. Men in this age-group commit 43% of all crime and 42% of VAW. With this, it is plausible to assume that men are more crime-prone and, that crime peaks at 18-30 years of age.

The average age at first marriage for males is 22.6 (s.d. 4.7) (Table 21), and the share of unmarried men is 63% for the age group 20-24 dropping to 19% in the age group 25+, and to close to zero by age 35-44 (see Table 13).[16] Overall, these descriptive statistics motivate our decision to use the sex ratio age 20-24 in the analysis. However, as discussed above, we include other age groups in specification checks.

3.2 Sex Ratio Data

We use decadal census data on age-specific population sex ratios at the district-level from five rounds over the period 1971 to 2011. The Census provides information disaggregated by age groups of 5-year bins starting from age 0 to 4. Since we have annual crime statistics but decadal sex ratios to make use of the yearly-variation in crime we use several interpolation techniques to estimate the yearly sex ratio. Our main specification uses annual sex-ratios by age-group estimated using cubic spline interpolation between census years (Smith et al., 2004; Karlsson and Ziebarth, 2016). We test robustness to replacing this with a linear, linear spline and cubic interpolation. The annual estimates of the sex ratio are similar with these different approaches.[17] In a further specification check, we drop the intercensal years altogether and define the dependent variable as the crime rate averaged over a decade as a function of sex ratios at the start of the decade. The district level census panel was obtained from the Indian District Database of the University of Maryland for the Census rounds of 1971 and 1981 and the Census of India publications for the rounds 1991, 2001 and 2011. Table 14 presents summary statistics. For

[16] Note that the NCRB creates age groups such that 18-30 is the closest to the census age grouping of 20-24.
[17] In Appendix A we show the trend in the sex ratio estimated using these various methods.

1971 and 1981 we only have age-specific population data up to the age of 35 years. As a result, the specification check that includes older age-groups is restricted to data from 1991 onwards.

The population sex ratio in India has been male-biased since the first Indian census was conducted in the mid 19th century Sen (2003). It exhibits considerable regional variation, and the regional pattern has persisted for more than a century, reflecting cultures of son preference that have been argued to emerge from differences in India in kinship structures, property rights and women's participation in agriculture (driven by soil type) (Anttila-Hughes et al., 2016; Carranza, 2014; Dyson and Moore, 1983; Sen, 2003). The sex ratio at birth has shown a particular rise since the mid-1980s, reflecting female foeticide, while the sex ratio in the population older than age five has stabilized (Bhalotra and Cochrane, 2010). The sex ratio in the latest Census of 2011 was 106 males per 100 females which is considerably higher than the biologically normal total ratio (Census of India, 2001; Haub and Gribble, 2011).[18]

The sex ratio of adults age 20-24 has risen over time in line with the birth ratio or the 0-4 ratio (Figure 1);(see Table 14). The child sex ratio (age 0-4) is stable between 1971 and 1981 but, after prenatal sex detection becomes feasible, it jumps from 102 in 1981 to 104 in 1991 and to just under 108 in 2001. This cohort trend in excess of males at birth persists to later ages. The Pearson correlation coefficient is 0.6. In Figure 1 we plot the sex ratio by age-group of those aged 0-4 in the pooled sample. We also present a scatter plot in which district-year sex ratios for the age group 20-24 are plotted against the sex ratio of those same cohorts by district 20 years before. This relationship is clearly positive and its coefficient is about 1.02 (see Figure 1). See Figures 1 and 3. Notice that the spatial distribution of the sex ratios exhibits variation across districts within states.

There was massive variation in levels and trends in crime and sex ratios across the Indian districts (Figure 3). Besides, as we show, there is a correlation between high youth sex ratios, rates of single men and VAW.

3.3 Data on Covariates

Literacy rates and shares of the rural, scheduled caste and scheduled tribe population are measured at the district-level for all five rounds of Census. We also construct the literacy gap

[18] China performs even worse than India. While India and China are outliers, together India and China account for 50.4% of the global population (Hesketh and Xing, 2006).

between males and females as a proxy for district-year gender equality. We account for income shocks by including total annual rainfall. This is gathered from the Global Precipitation Climatology Centre (GPCC)[19]. We use information on total monthly precipitation measured at the 0.5 degree resolution over 1970-2010 and map it to district centroids to calculate total annual rainfall. This measure has been extensively used in the literature and has been shown to accurately measure the spatial heterogeneity in rainfall across Indian districts and, to capture income variability (Iyer and Topalova, 2014).

There are no measures of law and order at the district-level but we have police strength per capita (in logarithms) at the state level from NCRB publications. We define a dummy for state election years to account for spending cycles in law and order using data from the Election Commission of India. We also do not have a consistent time series of income by district, so instead, we use state per capita income (in logarithms) collected from the Reserve Bank of India.[20] Table 14 presents summary statistics.

3.4 Data on Attitudes and Marriage Quality

We obtain data on the share of married men and women by age group and across districts using the two publicly available data from Census 1991 and 2011. We also obtain marriage rates from the third round of the National Family Health Survey (NFHS-3) which surveyed all men and all women of reproductive age. This allows to document the share of men married at age 20-24 by birth cohort. The marriage rates are comparable across the census and survey data, which validates the representativeness of the survey data (see Table 13. The NFHS-3 sampled 94,324 men and 124,385 women, and for consistency with the census analysis, we use the information from the 16 major states only. Also, we can match a randomly selected sample of men and women for which couples identifiers are available. We use the Census data to establish the link between sex ratios and marriage rates. To look at attitudes towards domestic violence and marriage quality, we make use of the NFHS data for men and women and married women and couples, respectively. A potential drawback of using the NFHS-3 is the fact that we cannot identify district (the local marriage market), only state of residence. Summary statistics are presented in Tables 21 and 22.

[19]The data are available at http://www.trfetzer.com/data/, accessed 09.11.2016.

[20]There is no district-year level annual data on unemployment for India and although the NSS does gather unemployment and employment, there is an array of part-time and informal work in response to no unemployment benefits. This makes unemployment a difficult measure of economic opportunities.

4 Results

We present estimates of equation (1) for total crime, violence against women (VAW), non-gendered violence, property and economic crime in Table 1. Moving from columns (1) to (2) we include state-level control variables, and from (2) to (3) we include instead state-year dummies (our preferred specification). The main finding is that a surplus of men age 20-24 increases violent crime but not acquisitive (economically motivated) crime. For violence against women, this result is robust to all controls, including state-year fixed effects and population growth -columns (1) of Table 4)-and the inclusion of a dummy for the introduction of political gender quotas in India (columns (2) of Table 4). Regarding effect sizes, an increase of one standard deviation in sex ratio for ages 20-24 (0.11, around a sample, mean of 1.033) raises gender-based violence by 12.69%. In columns (3) of Table 4 we present estimates of (1) using a log-log model instead (see Table 4), and the elasticity is close to one for violence against women.

Although we see some tendency for the youth sex ratio to raise non-gendered violence, this result is sensitive to functional form (log-log vs semi-log) and to control for population growth. For property and economic crime, the coefficients are smaller and not statistically significant. Total crime, which is the sum of the four categories displayed has an elasticity of close to 0.5 with respect to the youth sex ratio, which is also not significant. A useful perspective on this is provided by comparing estimates of a broadly similar specification for China in Edlund et al. (2013). The Chinese data were not available disaggregated by crime type by province over time, but total crime (violent and property) exhibits an elasticity of 3.7 which is significant at the 5% level.

The child sex ratio is not statistically significant (and tends to attract a negative coefficient) in the equations for VAW which is where it is pertinent as a control for trends in son preference.[21] The primary results are also robust to including in the equation the sex ratios of the age groups neighbouring 20-24, that is the ratios at ages 15-19 and 25-34 (Table 2).

The covariates are displayed in Table 3 in the Online Appendix. In line with Iyer and Topalova (2014), we see that non-gendered crime is higher when rainfall is lower, consistent with a literature suggesting economic motives for crime. Similar to Ghosh (2006), we find

[21] It is statistically significant in the estimations for non-gendered violence. We do not have a clear explanation of this, but since children 0-4 clearly do not commit crime, it suggests that district-years in which there is a higher rate of female foeticide (or excess child mortality amongst girls) are also district-years in which non-gendered violence is high.

evidence of an electoral cycle in crime but only for a non-gendered crime which decreases in the run-up to an election.

4.1 Robustness

In addition to adjusting for covariates, conditioning on district linear trends, state-year effects and sex ratios for ages other than 20-24, we conducted some other specification checks. First, we re-estimated the main equations in Table 1 using just the five census data points, rather than interpolating the sex ratio between decades. We defined the dependent variables as the decadal change in the logarithm of the crime rate and regressed this upon the sex ratio (and other census covariates) measured at the beginning of the decade. So for instance, we allow the surplus of males in 1971 to determine the change in crime between 1971 and 1981.[22] See Table 3.

As discussed briefly in the Introduction, the deficit of women in the age-group 20-24 may not translate directly into a shortage of brides for men age 20-24 if men marry younger women (the age gap is about five years) and if the population is growing. To account for differences in population growth that could impact upon the marriage market and marital match quality, we add population growth as an additional control (see columns 1 of Table 4). This creates no significant change in the main results.

The introduction of gender quotas in villages since the 1990's has been shown to increase reporting rates. We check whether our result is sensitive to including in our specification an indicator for the year in which the state implemented the quota, following (Iyer et al., 2012). We present these results in Table 4 column 2. The results are not sensitive to this.

We also explore the sensitivity of our results to functional form by including sex ratios in logarithms (columns 3 of Table 4). Since crime is already in logarithms, the coefficients in this log-log model are elasticities. We also present results using an alternative measure of sex ratios, the share of males by age-group to total population (columns 4 of Table 4). Finally, we also present results estimates without the inclusion of controls and separately, with the inclusion of gender literacy gap that would account for changing differences in reporting or tolerance towards crime (see Table 19 in Appendix A). The results are robust to these additional tests.

[22] In an alternative specification we followed Dorn and Hanson (2015) and used a first-differenced model. Results remain qualitatively similar and are available upon request.

We estimated a triple-difference type specification, exploiting variation across district, year and crime category. Estimates of the previous double-difference equation show that a surplus of males age 20-24 increases violent crime against women but has no statistically discernible impact on property crime (these results are presented later). On this premise, we specify VAW as the treated crime and property crime as the control crime. This specification does not *require* that the youth sex ratio has a zero impact on property crime (although, statistically, we show it does), but it allows us to estimate the differential impact of the youth sex ratio on VAW relative to property crime. We drop other forms of crime from this sample, namely non-gendered violence and acquisitive crime. This sort of specification has been used before in the health and crime literature (Jayachandran and Lleras-Muney, 2009; Blakeslee and Fishman, 2014). The specification is as follows:

$$Crime_{cdt} = \alpha_0 + \beta_1 SR_{dt}^{20-24} + \gamma_2 (SR_{dt}^{20-24} \times T_c) + \beta_2 T_c + \beta' X_{dt} + \lambda_{cd} + \omega_{ct} + g_d t + \theta_{st} + \epsilon_{cdt} \tag{2}$$

where, as before, $Crime_{cdst}$ is the logarithm of per capita crime of category c in district d of state s measured in year t. We interact the coefficient of interest (SR^{20-24}) with a dummy (T_c) that takes the value one if crime type c is gender-based violence (*treated* crime) and zero if it is property crime (*control* crime). The coefficient on this term, γ_2, captures the influence of the surplus of men age 20-24 on VAW relative to property crime and we expect $\gamma_2 > 0$. Since we pool two sorts of crime here, the district fixed-effects are now crime-specific (λ_{cd}), allowing that regional cultures and other time-invariant unobserved factors are specific to crime type. Similarly the year fixed-effects are crime-specific (ω_{ct}) to allow for instance that national shocks such as the 1991 economic reforms may have affected property crime but not gender-based violence, and to allow for crime-specific variation in national trends in reporting crime (Iyengar, 2009; Aizer and Dal Bó, 2009; Iyer et al., 2012; Miller and Segal, 2014). We report robust standard errors twoway-clustered at the district-decade and crime category category.

For results of this specification to be biased by omitted trends, these would have to vary not only by district and year but also by crime-type. Moerover, as in equation (1), they would have to be correlated with the ratio of men to women age 20-24 while being uncorrelated with the ratio of boys to girls age 0-4. Put this way, we cannot think of a likely confounder.[23] Equation

[23]For instance, any district-year omitted variables that lead to gender differences in migration or mortality that influence the youth sex ratio are likely to be common across crime categories and differencing the impact of the youth sex ratio across the two types of crimes addresses this.

(2) also provides a test of the significance of the difference in the impact of a surplus of young men on violence against women relative to property crime.

See Table 5 which displays a sequence of successively richer controls. The relationship between VAW and the youth sex ratio remains positive, smaller than before but not significantly so (see Table 1).

Next, we present a robustness test to ensure our results are not due to potential measurement error concerns related to the yearly transformation of the decadal data. We make the decadal data (so T=5) in levels and repeat an identical model to that in (1). Also, we also present a first-differences model again, similar to that in (1). We present these results in Table 20 in Appendix A. In both approaches, we find identical estimates to those in Table 1 and the coefficients on 20-24 are not statistically different from those in Table 1 .

Finally, we attempted to use an instrumental variables approach. In Appendix C and in Tables 28, 29 and 30, we present these results and detailed explanation as to why we think this is not the correct approach in this setting.

5 Mechanisms

In the Introduction, we discussed alternative mechanisms that might explain why a male surplus at age 20-24 results in increased VAW. Since the maleness effect- a compositional effect- can occur at most account for an elasticity of 0.5 (see Edlund et al. (2013)), an elasticity greater than this suggests a behavioural effect. In this section, we discuss the evidence on behavioural effects operating through a rising share of unmarried men, deteriorating marriage quality for women, and changing attitudes toward VAW.

5.1 Marriage rates

Previous research, cited earlier, has documented the civilizing effects of marriage on crime. We know of no prior research linking unmarried men to VAW, although this association seems like a natural one to make. Since 1995, when crimes against women other than rape are being recorded, the share of rape, sexual harassment, molestation and kidnapping in all VAW is 56%.

It is plausible that unmarried men are more likely to commit these crimes. To assess the

association between the 20-24 surplus of men and rates of marriage of men in that age-group, we used both census and survey data. First, using district-year level census averages, we estimate:

$$Unmarried_{dc}^{g,k} = \alpha_0 + \gamma SR_{dc}^{k} + X'_{dc} + \alpha_c + \lambda_d + \epsilon_{dc} \tag{3}$$

where $Unmarried_{d}^{g,k}$ is the share of unmarried individuals of gender g, of cohort k in district d at census c. The sex ratio, SR_{dc}^{k}, is the ratio of males to females of cohort k in the district d in census c and the coefficient γ is the coefficient of interest, determining how the age-specific sex ratio translates into the probability of marriage at a particular age. We condition upon a vector of district characteristics (X_{dc}), namely, the share of rural population, the literacy rate and the share of SC/ST population. We allow for cross-age effects, i.e., for the sex ratio outside the k age-group to determine marriage rates in this group. Our specifications always include district fixed-effects (λ_d) and a census round dummy (α_c); recall that data availability restricts us to two census rounds, 1991 and 2011, for this piece of the analysis. We expect increases in the sex ratio of cohort k to be positively correlated with the fraction of never-married men ($\gamma > 0$) in that group and negatively correlated with that of women ($\gamma < 0$).

We use data on age-specific marriage rates for adults age 15-44. Across all age groups, the rates of non-marriage are higher for men than for women. As expected, the probability of being unmarried declines sharply with age, and is small by age 35-44, when it is 3.5% for men and 1.2% for women. In the 20-24 age range, on average 62% of men are unmarried, compared with 23% of women, consistent with an age gap in marriage that averages at about 5.016 in our study period (see Table 13). Using the National Family Heath Survey (NFHS-3) data, that surveyed individuals of reproductive age in 2004-2005, we see broadly consistent rates, with 74.5% of men and 32% of women age 20-24 being single.

The regression estimates using census data are in Table 6. A one standard deviation increase in the maleness of the sex ratio of 20-24-year-olds (0.11) is associated with a 2 percentage point increase in the share of unmarried men age 20-24. This result is most relevant for our analysis. In fact, 20-24 year olds observed in census 1991 were born in 1967-1971, before the introduction of ultrasound prenatal sex detection technology, and 20-24-year-olds enumerated in 2011 were born 1987-1991 after. Since previous research shows that the introduction of ultrasound technology, driven by exogenous changes in trade and industrial policy (Bhalotra and Cochrane, 2010) led to a sharp rise in the maleness of the sex ratio at birth, these coefficients may be deemed to

have a causal interpretation.

We also observe that the impact of a 1 s.d. increase in the maleness of the youth sex ratio on the share of all unmarried men age 15-44 is 1.42 percentage points. The same 0.11 increase in the male surplus at 20-24 translates into a reduction in the share of unmarried women of this age of 1.5 percentage points. In India, eventually, almost everyone marries (see Table 13 and the social stigma associated with being unmarried is high, at least for women. So our attention is focused upon age-specific rates and in particular, the data demonstrate that a large fraction of men 20-24 is unmarried and that this is increasing in the sex ratio in their age group.

5.2 Attitudes towards domestic violence

The third round of the NFHS data contains the first attempt to gather data from men, and it includes all men rather than only the partners of respondent women (which was the case in the earlier rounds). We hypothesize that the (self-reported) attitudes of men and women towards intimate partner violence are shaped by the sex ratio in their birth year. There are two potential channels at play. First, men growing up without sisters or with fewer sisters may have different preferences (and, similarly, for women growing up with more males around them) (Teso, 2016; Oswald and Powdthavee, 2010). Second, previous work cited above shows that the population sex ratio at birth (and in the early years, 0-4) reflects conscious attempts by parents to influence the sex composition of their children through either foeticide or purposive under-investment in survival, for instance, in breastfeeding or immunization. So if a more male-biased sex ratio at the birth of an individual reflects more gender bias among their parents, then it is plausible that their attitudes, growing up with those parents, are more gender biased. We cannot distinguish between the first and second channel as the NFHS does not link adult siblings, or record the number or gender of siblings of respondent adults.

Since we now use the sex ratio at birth rather than the sex ratio of young adults, we have sufficient data in the post-ultrasound period to exploit the sharp pre vs post ultrasound increase in the birth sex ratio to aid identification (this is illustrated in Bhalotra and Cochrane (2010) who argue it as a causal increase). The identifying assumption is that attitudes towards VAW would have been the same in the absence of the exogenous change caused by the introduction of ultrasound that affected sex ratios at birth. We use a conservative separation of the pre and post ultrasound birth cohorts, as those born between 1957-1979 vs those born between 1980-

1990, although our results are not sensitive to cutting the data in 1985. The youngest cohorts, born 1990, are aged 14-15 in NFHS-3 (conducted in 2004-05), and so they just enter the sample of reproductive age adults. To measure the intensity of exposure to ultrasound we merge in state-level data on the child sex ratio (age 0-4) from the Census years of 1961 to 1991 measured at the year of birth of each. We employ the following difference-in-difference specification:

$$Attitudes_{ist} = \alpha_0 + \delta_1(CSR_{st} \times Post_t) + \beta_1 CSR_{st} + \beta_2 Post_t + X'_{ist} + \gamma_s + \lambda_t + \epsilon_{ist} \qquad (4)$$

where CSR_{st} is the ratio of males to females of ages 0-4 at the birth-year t of individual i in state s; $Post_t$ is a dummy variable that takes values 1 for individuals born in or after 1980 and 0 otherwise. The coefficient δ measures the differential impact on attitudes of the maleness of the child sex ratio across those born after versus before the introduction of ultrasound. The coefficients β_1 capture the pre-trend in the child sex ratios by state (which may reflect trends in gender gaps in mortality) and β_2 allows that, for some omitted reason, there is a post-ultrasound break in the underlying trend in attitudes towards gender violence. We include a set of individual-specific covariates, X_{ist}, which are the wealth-index of the household, age, educational attainment and a rural/urban dummy, but the estimates are not sensitive to these controls. We estimate equation (4) separately for men and women and present the summary statistics in the Online Appendix (Table 8).

The dependent variables are created from individual responses to a module of questions on attitudes towards domestic violence. Male and female respondents are asked, for different domains, whether they tolerate domestic violence.[24] Using these responses we create an acceptance index by summing the binary answers, so the index varies between 0 and 5. We also use a binary measure which takes the value 0 if the respondent never justifies physical abuse and 1 if they justify it in at least one domain. On average 40% of men and 42% of women answer that at least some form of physical abuse is justified (Table 8). The results of estimating equation (4) are in Table 7.

We find that among men born after the introduction of ultrasound technology, a one standard deviation increase in the child sex ratio (0.028) created an additional 11% percentage point increase in the binary indicator of tolerance of violence against women. These results are not

[24]The the precise phrasing used in DHS is *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...she burns the food?*; iii) *... she refuses to have sex?*; iv) *...neglects the children?* and, v) *argues?*.

driven by tolerance in one particular domain, suggesting a fairly pervasive change in attitudes (see 10 in the Online Appendix). Estimating the same equation for women, using the child sex ratio in their birth years, we find no evidence that women's attitudes have worsened. In fact, the coefficients are negative suggesting that women's tolerance for violence has decreased. These results are important to our main results since domestic violence represents 44% of total VAW.

These results line up with those found in Grosjean and Khattar (2015) who show that in Australia in areas where the sex ratio of English colonizers was high, current acceptance of more equal gender roles is lower. Hussam (2016) shows, using the same data, that women born after the introduction of ultrasound technology had worse marriage outcomes, but she does not study the attitudes data.

5.3 Marriage quality

One might expect that as the population sex ratio becomes more male-biased so that at any given age there are fewer women than men available on the marriage market, women's marriage prospects would improve (Abramitzky et al., 2011). However, as long as there is positive population growth, this will be offset by women tending to marry older men- the hypothesis of marriage squeeze. d'Albis and De La Croix (2012) estimates that, in the year 2000, India had 1.4 men for every potential bride. We estimate that the relative availability of grooms to brides in the population evolved from a deficit of 98.5 potential grooms to 100 potential brides in 1992 to a surplus of 0.6%, 4.3% and 5.1% by 1998 and 2005, respectively (see Table 10).[25]

To investigate the effects of the youth sex ratio on marriage quality we follow Anukriti et al. (2017) in using two features of the Indian marriage market to identify the effect of sex ratios on marriage quality. First, we use the fact that the marriage market is endogamous along caste (and religious) lines. This feature has been shown to be very persistent within caste groups (Banerjee et al., 2013).[26] The second feature, is the fact that the sex ratio varies by caste lines with the higher castes having more male-biased sex ratios than, the lower castes (scheduled caste/tribe groups (SC/ST)) (Chakraborty and Kim, 2010; Jayachandran, 2017a). Figure 6 plots the sex ratio (left panel) and marriage rates (right panel) by age and caste groups. It shows that the sex ratio for the higher caste groups is higher across the age distribution. The all-India average

[25]To obtain these figures we calculate the Missing Brides Index following d'Albis and De La Croix (2012) and using micro-level data from all available rounds of the National Family Health Survey of India.

[26]Since we use Census data we cannot identify religion as this information is not available by age-group.

sex ratio is 104.37 males per 100 females among the higher castes, 103.59 among SC and 99.77 for ST. This is consistent with marriage patterns which show higher marriage rates in the lower caste groups with more equal sex ratios. The upshot for our purposes is that we can exploit (a) caste endogamy and (b) variation in the male surplus across caste groups. In fact we exploit variation in the youth sex ratio across not only caste but also state and (cohort) exposure to ultrasound technology.

We model the quality of a woman's marital match as a function of the sex ratio of the mating pool available within each state and for each age-caste group, conditioning on the wife's pre-marriage characteristics. We estimate the following triple difference specification:

$$Y_{isca} = \alpha_0 + \delta_2(SR_{sca} \times Post_a) + +\beta_3 SR_{sca} + \beta_4 Post_a + X_i' + \gamma_s + \zeta_c + \lambda_a + \epsilon_{isca} \quad (5)$$

where Y_{isca} is an indicator of the quality of the husband of woman i. The interaction term $(SR_{sca} \times Post_a)$ measures the differential effect of the sex ratio on the husband's quality for a given woman i in-state s of caste group c born before vs after the introduction of ultrasound. The coefficient of interest is δ_2. Pre-determined characteristics of the wife (vector X_i) include rural residence, religion, state, caste and age-group (γ_s, ζ_c and λ_t). The outcomes are years of education, age and height of the husband and the educational, age and height gap between spouses. Higher gaps between spouses have been shown to be relevant risk factors affecting female empowerment (Ackerson et al., 2008; Eswaran and Malhotra, 2011; Erten and Keskin, 2016). We also include measures of controlling behaviour, the incidence of physical, emotional and sexual intimate-partner violence (IPV) and willingness to share decisions about money [27]. We use both the woman's and the couples modules of NFHS-3 to retrieve all available measures of quality.[28] Female education is also a crucial determinant of marriage quality and marriage payments (Lafortune, 2013; Roy, 2015). However, female education is endogenous to sex ratios and while we do not have a formal way of taking into account the fact that endogenous

[27] Our approach to measuring IPV follows the Conflict Tactics Scale (CTS) method employed in DHS surveys Straus et al. (1996). These are self-reported measures where respondents are asked to indicate whether they have been subject to very specific type of behaviours over a specific period (Table 11 presents the description of these variables). We use these measures to construct a single measure of IPV and also use these separately as individual variables.

[28] The NFHS-3 surveyed only a sample of husbands, and this explains the differences in sample sizes. Moreover, while some variables (e.g. education of the spouse) are available in both modules, others (e.g. husband's height) are only available in the couples files. Our results are not sensitive to whether we use the couples or the women's files but to improve precision we report results using only the married women's sample. Results from the couples files are available upon request.

educational investments may affect matching and bargaining outcomes we provide estimations of (5) with and without controlling for female education (see Panel B of Table 8). Results are presented in Table 8 and 25.

We find that higher sex ratios among women born after the introduction of ultrasound technology decreased women's marriage quality in several dimensions. In areas with higher sex ratios, women were more likely to marry a husband with fewer education years, be victims of domestic violence and have less decision-making power (Table 8. The coefficients on the spousal height gap (columns 4 and 10 of Table 8) suggest that higher sex ratios widen the height gap between spouses, but this coefficient is not statistically significant. We disaggregated the indicator for intimate-partner violence by types of violence, namely emotional, physical and sexual. We find that the increases in intimate-partner violence are largely due to changes in sexual violence (by 9.2 percentage points). These results are presented in Tables 25 and 24.

Concerning magnitudes, a one standard deviation increase in the sex ratio (0.064) leads to an increase in the self-reported incidence of domestic violence (i.e., any IPV- column 8) after the shock by 4.5 percentage points i.e. an 18% increase in relation to the mean. This is consistent with our findings of section 4 since domestic violence alone represents 44% of total VAW. Women in district-years with more male-biased youth sex ratios are also 5.9 percentages points more likely to face a controlling partner and 8.5 percentages points less likely to be involved in financial decision-making. Overall, these findings show that a marriage-age surplus of men worsens marital outcomes for women. When controlling for female education the magnitudes of the effect sizes decrease (in particular for the education of the spouse) but they remain statistically significant for measures of domestic violence and decision-making. Our findings are broadly in line with Hussam (2016) who uses cluster-level variation in the sex ratio at birth for pre vs post ultrasound cohorts and documents a decline in the quality of marriages for women.[29]

6 Discussion

Our findings are consistent with a mechanical effect associated with increased maleness in the population on crime, given that men commit more crime. Indeed, we find no significant impact on VAW of the sex ratio at ages 25-34, consistent with age 20-24 being the age at which crime

[29] Hussam (2016) defines the local marriage market to be defined at the cluster level. However, Indian women often marry outside their natal village (or cluster) (see for example Rosenzweig and Stark (1989)). In the Indian Human Development Survey of 2005, 87.05% of women married men outside the natal village Amaral (2017).

rates peak. However, the stylized fact, for India and elsewhere, is that men commit more crime of all sorts our effect sizes are consistent with a more nuanced effect that is consistent with sex ratios distorting marriage markets. Also, the largest elasticity consistent with a compositional effect alone is 0.5, and the elasticity we estimate is 0.7.

We find support for three non-competing mechanisms. First, differencing between the pre and post ultrasound birth cohorts observed as 20-24-year-olds in the 1991 and 2011 census respectively, we estimate that a one standard deviation increase in the sex ratio leads to a 2 percentage point increase in the probability that men age 20-24 are unmarried, and we refer to a literature on the civilizing effects of marriage to argue that this contributes to violence. Since consensual sex outside marriage remains uncommon and stigmatized in India, it is plausible that an increasing share of unmarried men to raise VAW, 56% of which is on account of rape, molestation, harassment and kidnapping of women.

Second, we find that the self-reported attitudes of men towards violence against women are a function of the sex ratio at birth (which largely persists to adulthood) and, again, we use the plausibly exogenous introduction of ultrasound technology to identify this relationship. A one standard deviation increase in the sex ratio at birth (0.025) is associated with a 0.05 percentage point change in the probability that men agree that wife beating is justified in at least one of five domains. While our results for women are less robust, we find that women's attitudes have improved and this difference is consistent with a *male backlash* hypothesis (Bobonis et al., 2013). A third mechanism is that the quality of marriages for women is deteriorating, across several measures including, domestic violence (which would explain 44 percent of the total crime rate). Our findings are consistent with (Grosjean and Khattar, 2015).

It is possible that two additional mechanisms explain our findings: i) dowry payments and ii) female labour participation. As to the first, a rise in sex ratios could lead to an increase in dowry prices and this, in turn, would lead to a deterioration of conditions for women- for instance, if dowry payments are lower, women may be subject to increased domestic violence. However, dowry prices appear to be stable since the 1960's Anukriti et al. (2016b) and we are unaware of evidence linking changes in dowry prices driven by changes in sex ratios (due to marriage squeeze as suggested by Rao (2000)). Second, it is possible that a rise in sex ratios that leads to a more regressive gender norms (as we show) can reduce female labour participation and that this, in turn, can lead to increased intimate-partner violence on account of male backlash ((Bloch and

Rao, 2002; Eswaran and Malhotra, 2011; Amaral et al., 2015; Bhalotra et al., 2016b)). However, in India, female labour participation is very low and has decreased over time Afridi et al. (2016); Field et al. (2016).

7 Conclusion

Our main finding is that the increasing male bias in the sex ratio of 20-24-year-olds has led to increasing violence against women in India. The elasticity is unity, and we estimate that a one standard deviation increase in the youth sex ratio results in an 13 percent increase in VAW. Our results suggest that changes in the youth sex ratio account for as much 21% of the change in VAW. This result is robust to a variety of specification checks.

We present the first evidence that male-biased population sex ratios generate violence against women, although Edlund et al. (2013) provides evidence for China that the same phenomenon generates more generalized violence (their measure is arrests for property and violent crime). We also present novel proof on some of the driving mechanisms, and the first evidence that (self-reported) attitudes towards domestic violence are evolving adversely in response to the rising sex ratio at birth.

Our results contribute to research on violence against women; and also to research on the long-term consequences of rising birth sex ratios. In India, prenatal sex detection and female foeticide started to be widespread enough to have population-level impacts in the mid-1980s, and the trend accelerated in the mid-1990s. These cohorts have now reached adulthood, and our estimates suggest that men face a more competitive marriage market and they have less progressive attitudes towards violence against women. This is important as recent evidence indicates that in spite of improvements in health investment, lower fertility and policies aimed at correcting for son preference, the trend in the sex ratio is not expected to revert (Bongaarts and Guilmoto, 2015; Jayachandran, 2017b; Anukriti, 2014).

From a policy perspective, our results indicate that policies aimed at increasing prevention of violence against women and addressing gender attitudes are a potential avenue for interventions as suggested in Bandiera et al. (2012); Anwar et al. (2016); Amaral et al. (2018).

References

Abiona, O., Koppensteiner, M. F., 2016. The impact of household shocks on domestic violence: Evidence from Tanzania. University of Leicester Working Paper No.16/14.

Abramitzky, R., Delavande, A., Vasconcelos, L., 2011. Marrying up: the role of sex ratio in assortative matching. American Economic Journal: Applied Economics, 124–157.

Ackerson, L. K., Kawachi, I., Barbeau, E. M., Subramanian, S., 2008. Effects of individual and proximate educational context on intimate partner violence: a population-based study of women in india. American Journal of Public Health 98 (3), 507–514.

Afridi, F., Dinkelman, T., Mahajan, K., 2016. Why are fewer married women joining the work force in india? a decomposition analysis over two decades.

Agarwal, B., 1994. A field of one's own: Gender and land rights in South Asia. Vol. 58. Cambridge University Press.

Aizer, A., 2010. The gender wage gap and domestic violence. The American Economic Review 100 (4), 1847.

Aizer, A., Currie, J., 2014. The intergenerational transmission of inequality: Maternal disadvantage and health at birth. Science 344 (6186), 856–861.

Aizer, A., Dal Bó, P., 2009. Love, hate and murder: Commitment devices in violent relationships. Journal of Public Economics 93 (3), 412–428.

Akerlof, G. A., 1998. Men without children. The Economic Journal 108 (447), 287–309.

Alesina, A., Brioschi, B., Ferrara, E. L., 2016. Violence against women: A cross-cultural analysis for africa. NBER Working Paper w21901.

Almond, D., Edlund, L., Milligan, K., 2013. Son preference and the persistence of culture: evidence from South and East Asian immigrants to Canada. Population and Development Review 39 (1), 75–95.

Amaral, S., 2017. Do improved property rights decrease violence against women in india? ISER Working Paper Series.

Amaral, S., Bandyopadhyay, S., Sensarma, R., 2015. Employment programmes for the poor and female empowerment: the effect of nregs on gender-based violence in india. Journal of interdisciplinary economics 27 (2), 199–218.

Amaral, S., Bhalotra, S., Prakash, N., 2018. Gender, crime and punishment. mimeo.

Anderberg, D., Rainer, H., Wadsworth, J., Wilson, T., 2016. Unemployment and domestic violence: Theory and evidence. The Economic Journal 126 (597), 1947–1979.

Anderson, S., Ray, D., 2010. Missing women: age and disease. The Review of Economic Studies 77 (4), 1262–1300.

Angelucci, M., 2008. Love on the rocks: Domestic violence and alcohol abuse in rural Mexico. The BE Journal of Economic Analysis & Policy 8 (1).

Angrist, J., 2002. How do sex ratios affect marriage and labor markets? Evidence from America's second generation. The Quarterly Journal of Economics 117 (3), 997–1038.

Anttila-Hughes, J. K., Krause, P., Stopnitzky, Y., 2016. The roots of modern sex ratios.

Anukriti, S., 2014. The fertility-sex ratio trade-off: Unintended consequences of financial incentives. IZA Discussion Paper 8044.

Anukriti, S., Bhalotra, S. R., Tam, H. F., 2016a. On the quantity and quality of girls: New evidence on abortion, fertility, and parental investments.

Anukriti, S., Kwon, S., Prakash, N., 2016b. Dowry: Household responses to expected marriage payments. Unpublished Manuscript.

Anukriti, S., Pop-eleches, C., Salanié, B., Turner, L., 2017. Sex ratios, schooling, and the marriage market in India.

Anwar, S., Bayer, P., Hjalmarsson, R., 2016. A jury of her peers: The impact of the first female jurors on criminal convictions. The Economic Journal.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M., 2012. Empowering adolescent girls: evidence from a randomized control trial in Uganda. World Bank Washington, DC.

Banerjee, A., Duflo, E., Ghatak, M., Lafortune, J., 2013. Marry for what? caste and mate selection in modern india. American Economic Journal: Microeconomics 5 (2), 33–72.

Bardhan, P. K., 1974. On life and death questions. Economic and Political Weekly, 1293–1304.

Basu, A. M., 1992. Culture the status of women and demographic behaviour: illustrated with the case of India.

Becker, G. S., 2009. A Treatise on the Family. Harvard University Press.

Bhalotra, S., Chakravarty, A., Gulesci, S., 2016a. The price of gold: Dowry and death in India. IZA Discussion Paper 9679.

Bhalotra, S., Cochrane, T., 2010. Where have all the young girls gone? Identification of sex selection in India. IZA Discussion Paper 5381.

Bhalotra, S., Kambhampati, U., Rawlings, S., Siddique, Z., 2016b. Crimes against women and economic conditions: Evidence from india. Mimeo.

Bhalotra, S., Rawlings, S., Siddique, S., Kumbathi, U., 2016c. Domestic violence and economic conditions. Working Paper.

Bhaskar, V., 2011. Sex selection and gender balance. American Economic Journal: Microeconomics 3 (1), 214–244.

Blakeslee, D., Fishman, R., 2014. Weather shocks, crime, and agriculture: Evidence from India. Working Paper.

Blattman, C., Jamison, J. C., Sheridan, M., 2015. Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia. NBER Working Paper w21204.

Bloch, F., Rao, V., 2002. Terror as a bargaining instrument: A case study of dowry violence in rural india. The American Economic Review 92 (4), 1029–1043.

Bobonis, G. J., González-Brenes, M., Castro, R., 2013. Public transfers and domestic violence: The roles of private information and spousal control. American Economic Journal: Economic Policy, 179–205.

Bongaarts, J., Guilmoto, C. Z., 2015. How many more missing women? Excess female mortality and prenatal sex selection, 1970–2050. Population and Development Review 41 (2), 241–269.

Burgess, R., Deschenes, O., Donaldson, D., Greenstone, M., 2013. The unequal effects of weather and climate change: Evidence from mortality in India. Working Paper.

Cameron, L. A., Meng, X., Zhang, D., 2016. China's sex ratio and crime: Behavioral change or financial necessity? Working Paper.

Card, D., Dahl, G. B., 2011. Family violence and football: The effect of unexpected emotional cues on violent behavior. The Quarterly Journal of Economics 126 (1), 103–143.

Carranza, E., October 2014. Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India. American Economic Journal: Applied Economics 6 (4), 197–225.

Carrell, S. E., Hoekstra, M. L., 2010. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. American Economic Journal: Applied Economics 2 (1), 211–228.

Census of India, 2001. Government of India.

Census of India, 2011. Government of India.

Chakraborty, T., Kim, S., 2010. Kinship institutions and sex ratios in india. Demography 47 (4), 989–1012.

Chen, Y., Li, H., Meng, L., 2013. Prenatal sex selection and missing girls in china: Evidence from the diffusion of diagnostic ultrasound. Journal of Human Resources 48 (1), 36–70.

Chiappori, P.-A., Fortin, B., Lacroix, G., 2002. Marriage market, divorce legislation, and household labor supply. Journal of Political Economy 110 (1), 37–72.

Ciccarelli, N., 2017. Health insurance premiums, medicaid eligibility, labor market conditions and crime. Working Paper.

Coale, A. J., 1991. Excess female mortality and the balance of the sexes in the population: an estimate of the number of "missing females". The Population and Development Review, 517–523.

Dahl, G., DellaVigna, S., 2009. Does movie violence increase violent crime? The Quarterly Journal of Economics 124 (2), 677–734.

d'Albis, H., De La Croix, D., 2012. Missing daughters, missing brides? Economics Letters 116 (3), 358–360.

Dorn, D., Hanson, G. H., 2015. The labor market and the marriage market: How adverse employment shocks affect marriage, fertility, and children's living circumstances. Working Paper.

Dyson, T., Moore, M., 1983. On kinship structure, female autonomy, and demographic behavior in India. Population and Development Review, 35–60.

Ebenstein, A., 2011. Estimating a dynamic model of sex selection in China. Demography 48 (2), 783.

Ebenstein, A. Y., Sharygin, E. J., 2009. The consequences of the "missing girls" of China. The World Bank Economic Review 23 (3), 399–425.

Edlund, L., 1999. Son preference, sex ratios, and marriage patterns. Journal of Political Economy 107 (6), 1275–1304.

Edlund, L., Li, H., Yi, J., Zhang, J., 2007. Sex ratios and crime: evidence from china's one-child policy.

Edlund, L., Li, H., Yi, J., Zhang, J., 2013. Sex ratios and crime: Evidence from China. Review of Economics and Statistics 95 (5), 1520–1534.

Ellsberg, M., Jansen, H. A., Heise, L., Watts, C. H., Garcia-Moreno, C., et al., 2008. Intimate partner violence and women's physical and mental health in the who multi-country study on women's health and domestic violence: an observational study. The Lancet 371 (9619), 1165–1172.

Erten, B., Keskin, P., 2016. For better or for worse?: Education and the prevalence of domestic violence in turkey. American Economic Journal: Applied Economics.

Eswaran, M., Malhotra, N., 2011. Domestic violence and women's autonomy in developing countries: theory and evidence. Canadian Journal of Economics/Revue canadienne d'économique 44 (4), 1222–1263.

Fetzer, T., 2014. Can workfare programs moderate violence? evidence from India. Working Paper.

Field, E., Pande, R., Rigol, N., Schaner, S., Moore, C. T., 2016. On her account: Can strengthening women's financial control boost female labor supply? Tech. rep., Technical report.

Freedman, M., Owens, E. G., 2016. Your friends and neighbors: localized economic development and criminal activity. Review of Economics and Statistics 98 (2), 233–253.

Gandhi, B. M., 2010. Indian penal code - third edition. Eastern Book Company.

Gavrilova, E., Campaniello, N., 2015. Uncovering the gender participation gap in the crime market. IZA Discussion Paper 8982.

Ghosh, A., 2006. Electoral cycles in crime in a developing country: Evidence from the Indian states.

Gould, E. D., Weinberg, B. A., Mustard, D. B., 2002. Crime rates and local labor market opportunities in the United States: 1979–1997. Review of Economics and Statistics 84 (1), 45–61.

Grosjean, P. A., Khattar, R., 2015. It's raining men! Hallelujah? UNSW Business School Research Paper No. 2014-29C.

Haub, C., Gribble, J., 2011. The world at 7 billion. Population Bulletin 66 (2), 1–12.

Hesketh, T., Xing, Z. W., 2006. Abnormal sex ratios in human populations: Causes and consequences. Proceedings of the National Academy of Sciences 103 (36), 13271–13275.

Hidrobo, M., Peterman, A., Heise, L., 2013. The effect of cash, vouchers and food transfers on intimate partner violence: Evidence from a randomized experiment in Northern Ecuador. International Food Policy Research Institute 84, 1555–1566.

Hussam, R., 2016. Marry rich, poor girl: The implications of sex selection for girl's later life outcomes in India. Working Paper.

Hvistendahl, M., 2012. Unnatural selection: Choosing boys over girls, and the consequences of a world full of men. Public Affairs.

Iyengar, R., 2009. Does the certainty of arrest reduce domestic violence? evidence from mandatory and recommended arrest laws. Journal of Public Economics 93 (1), 85–98.

Iyer, L., Mani, A., Mishra, P., Topalova, P., 2012. The power of political voice: Women's political representation and crime in India. American Economic Review, Applied Economics (11-092).

Iyer, L., Topalova, P. B., 2014. Poverty and crime: evidence from rainfall and trade shocks in India. Harvard Business School BGIE Unit Working Paper (14-067).

Jayachandran, S., 2017a. Fertility decline and missing women. American Economic Journal: Applied Economics 9 (1), 118–39.

Jayachandran, S., 2017b. Fertility decline and missing women. American Economic Journal: Applied Economics 9 (1), 118–39.

Jayachandran, S., Lleras-Muney, A., 2009. Life expectancy and human capital investments: Evidence from maternal mortality declines. The Quarterly Journal of Economics 124 (1), 349–397.

Karlsson, M., Ziebarth, N. R., 2016. Population health effects and health-related costs of heat events: Comprehensive evidence from germany. under review.

Klasen, S., 1994. "missing women" reconsidered. World Development 22 (7), 1061–1071.

Korenman, S., Neumark, D., 1991. Does marriage really make men more productive? Journal of Human Resources, 282–307.

Kumar, H., Somanathan, R., 2009. Mapping Indian districts across census years, 1971-2001.

La Mattina, G., 2016. Civil conflict, sex ratio and intimate partner violence in Rwanda. Journal of Development Economics, forthcoming.

Lafortune, J., 2013. Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. American Economic Journal: Applied Economics 5 (2), 151–178.

Levitt, S. D., 2002. Using electoral cycles in police hiring to estimate the effects of police on crime: Reply. The American Economic Review 92 (4), 1244–1250.

Machin, S., Meghir, C., 2004. Crime and economic incentives. Journal of Human Resources 39 (4), 958–979.

Menon, S., 2015. Unfinished lives: the effect of domestic violence on neonatal & infant mortality. ISER Working Paper No. 2014-27.

Miller, A. R., Segal, C., 2014. Do female officers improve law enforcement quality? effects on crime reporting and domestic violence escalation.

Miller, B. D., 1981. The endangered sex: neglect of female children in rural North India.

Mumford, D. B., Minhas, F. A., Akhtar, I., Akhter, S., Mubbashar, M. H., 2000. Stress and psychiatric disorder in urban Rawalpindi. The British Journal of Psychiatry 177 (6), 557–562.

Nagin, D. S., Land, K. C., 1993. Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. Criminology 31 (3), 327–362.

Oswald, A. J., Powdthavee, N., 2010. Daughters and left-wing voting. The Review of Economics and Statistics 92 (2), 213–227.

Papageorge, N., Pauley, G. C., Cohen, M., Wilson, T. E., Hamilton, B. H., Pollak, R. A., 2016. Health, human capital, and domestic violence. NBER Working Paper w22887.

Pino, F. J., Mookherjee, D., Chakravarty, A., Bhalotra, S., et al., 2016. Property rights and gender bias: Evidence from land reform in West Bengal. Working Paper.

Qian, N., 2008. Missing women and the price of tea in china: The effect of sex-specific earnings on sex imbalance. The Quarterly Journal of Economics 123 (3), 1251–1285.

Ramos, A., 2016. Household decision making with violence: Implications for conditional cash transfer programs. Universitat Autonoma de Barcelona.

Rao, V., 2000. The marriage squeeze interpretation of dowry inflation: response. Journal of Political Economy 108 (6), 1334–1335.

Rawlings, S., Siddique, Z., 2014. Domestic abuse and child health.

Rose, E., 1999. Consumption smoothing and excess female mortality in rural India. Review of Economics and Statistics 81 (1), 41–49.

Rosenzweig, M. R., Stark, O., 1989. Consumption smoothing, migration, and marriage: Evidence from rural India. Journal of Political Economy 97 (4), 905–926.

Roy, S., 2015. Empowering women? inheritance rights, female education and dowry payments in india. Journal of Development Economics 114, 233–251.

Sampson, R. J., Laub, J. H., 2005. A life-course view of the development of crime. The Annals of the American Academy of Political and Social Science 602 (1), 12–45.

Sampson, R. J., Laub, J. H., Wimer, C., 2006. Does marriage reduce crime? A counterfactual approach to within-individual causal effects. Criminology 44 (3), 465–508.

Sekhri, S., Storeygard, A., 2014. Dowry deaths: Response to weather variability in India. Journal of Development Economics 111, 212–223.

Sen, A., 1992. Missing women. BMJ: British Medical Journal 304 (6827), 587.

Sen, A., 2003. Missing women–revisited: reduction in female mortality has been counterbalanced by sex selective abortions. British Medical Journal 327 (7427), 1297–1299.

Siennick, S. E., Staff, J., Osgood, D. W., Schulenberg, J. E., Bachman, J. G., VanEseltine, M., 2014. Partnership transitions and antisocial behavior in young adulthood a within-person, multi-cohort analysis. Journal of Research in Crime and Delinquency.

Smith, L., Hyndman, R. J., Wood, S. N., 2004. Spline interpolation for demographic variables: The monotonicity problem. Journal of Population Research 21 (1), 95–98.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., Sugarman, D. B., 1996. The revised conflict tactics scales (cts2) development and preliminary psychometric data. Journal of family issues 17 (3), 283–316.

Sudha, S., Rajan, S. I., 1999. Female demographic disadvantage in India 1981–1991: Sex selective abortions and female infanticide. Development and Change 30 (3), 585–618.

Teso, E., 2016. The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade.

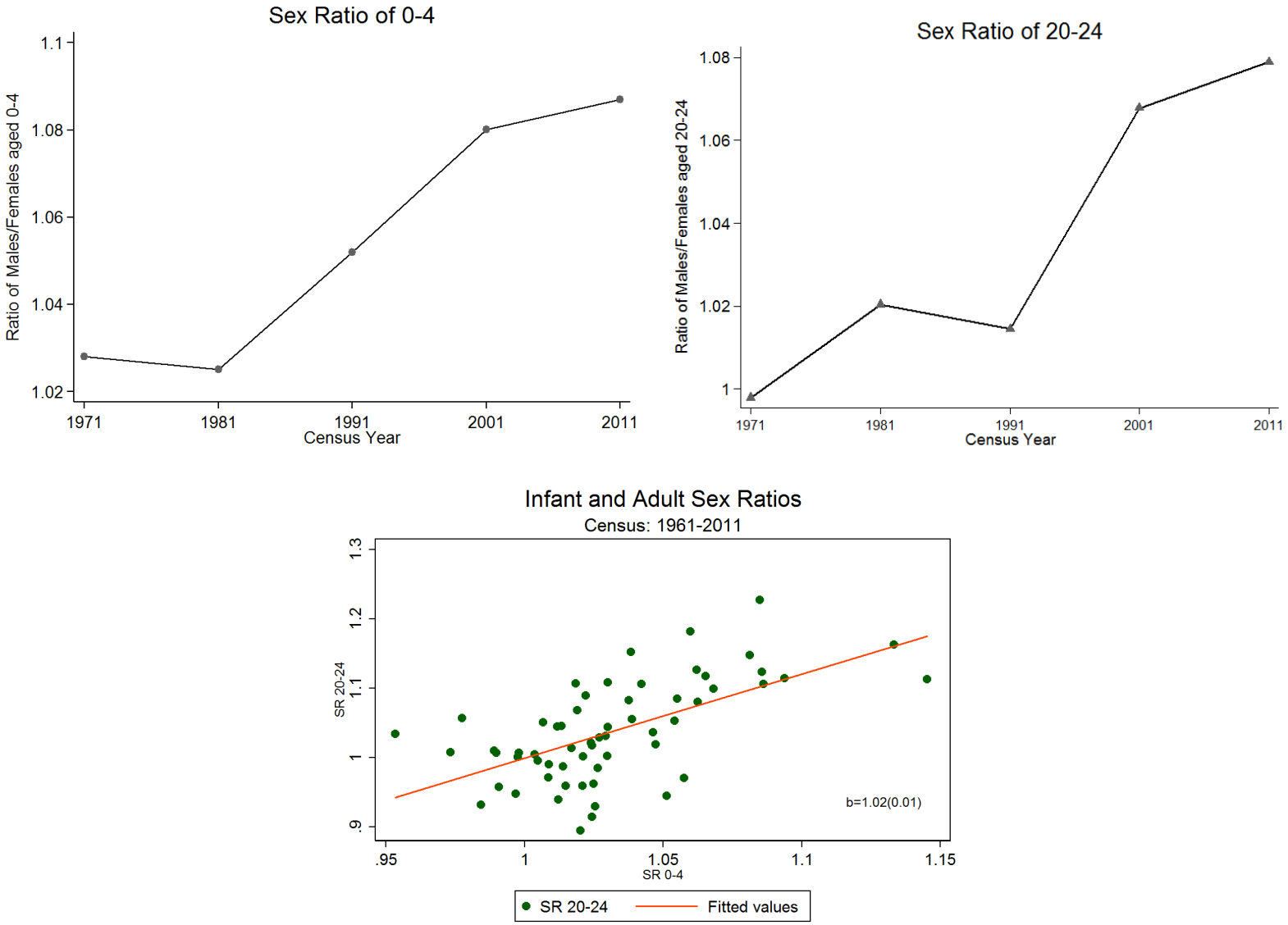
Theobald, D., Farrington, D. P., 2011. Why do the crime-reducing effects of marriage vary with age? British Journal of Criminology 51 (1), 136–158.

Tur-Prats, A., 2015. Family types and intimate-partner violence: A historical perspective. Working Paper.

Virmani, A., 2006. India's economic growth history: Fluctuations, trends, break points and phases. Indian Economic Review, 81–103.

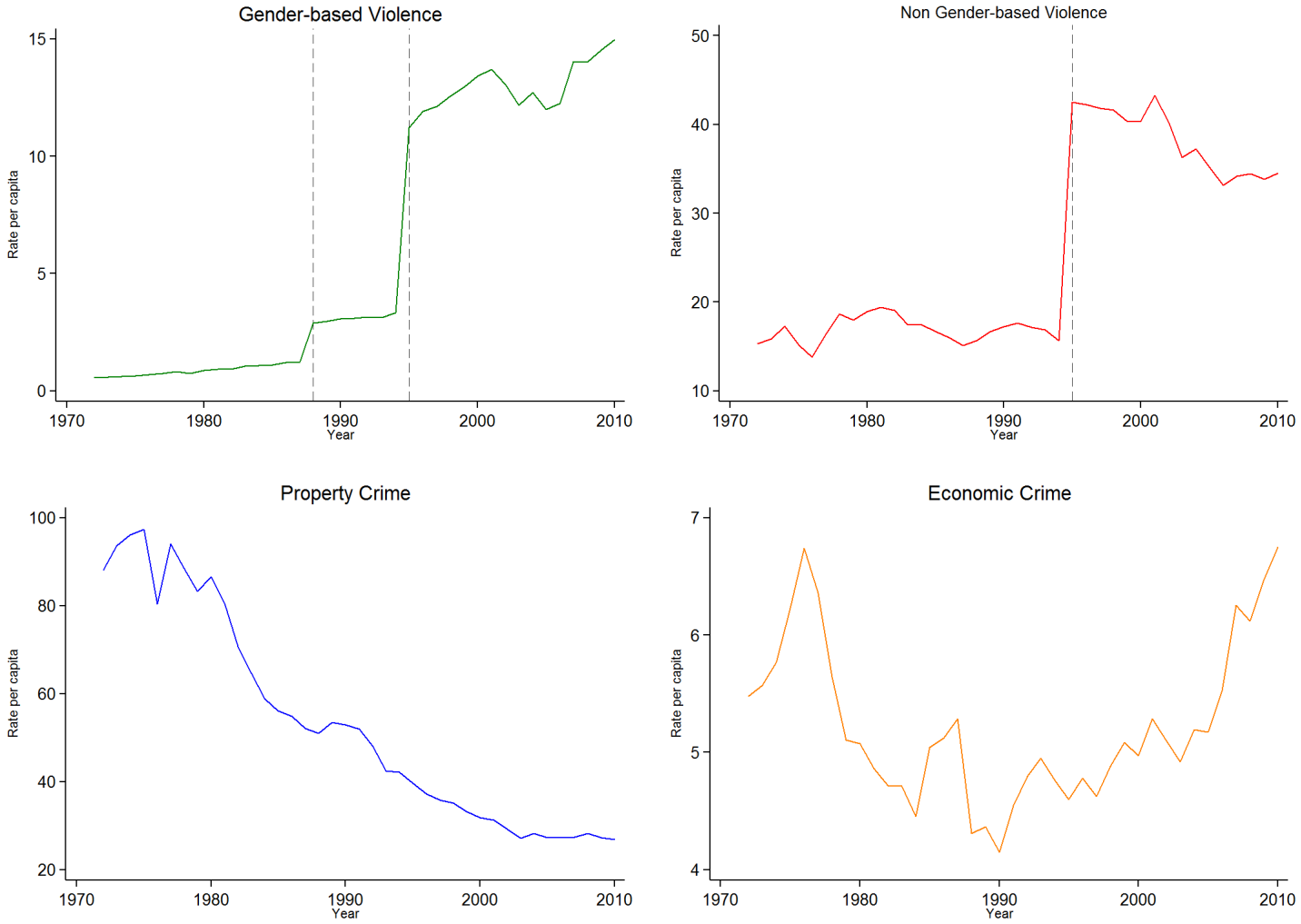
Visher, C. A., Knight, C. R., Chalfin, A., Roman, J. K., 2009. The impact of marital and relationship status on social outcomes for returning prisoners. Washington, DC: Urban Institute.

Figure 1: Trends in Age-specific Sex Ratios



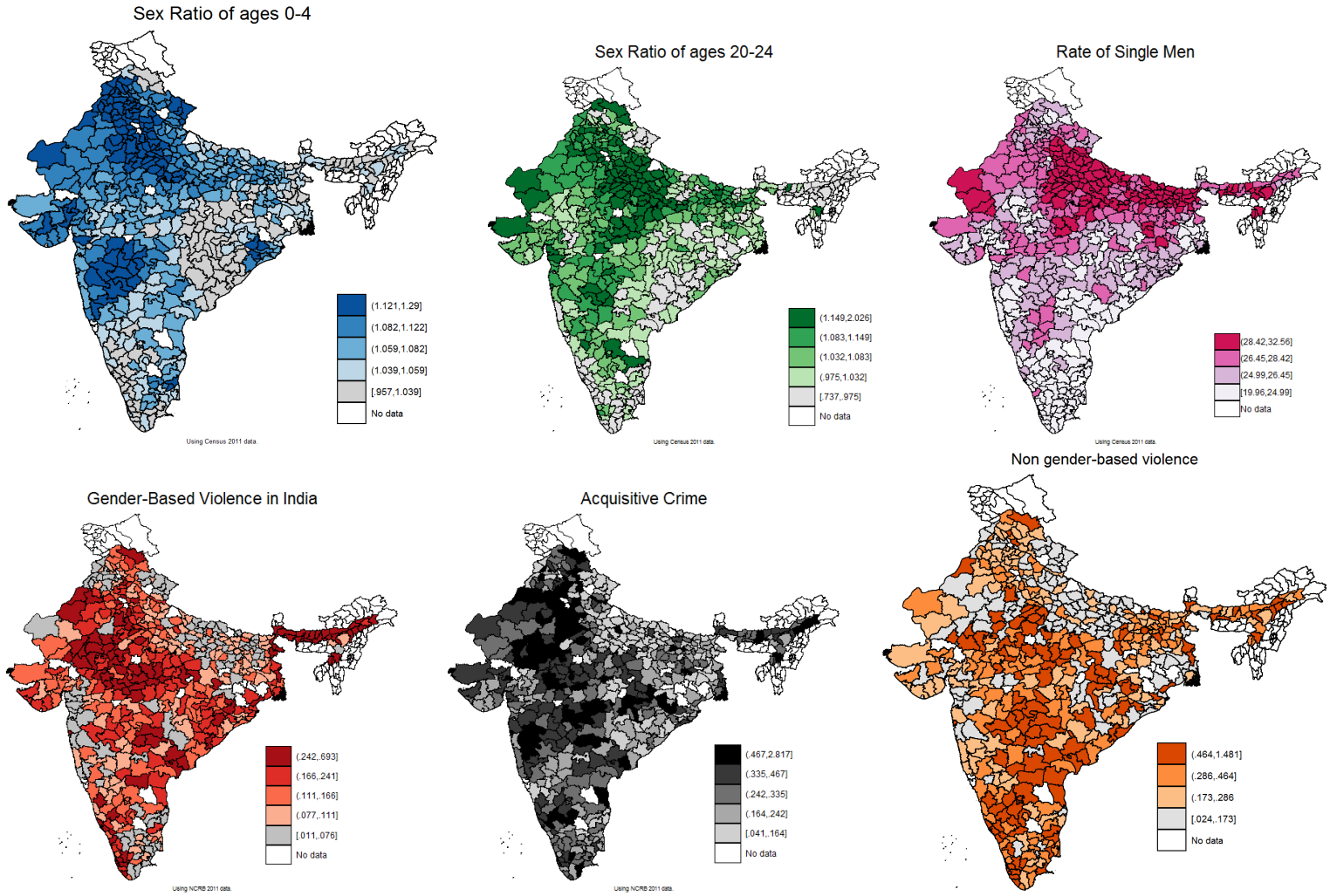
Notes: The top panels show the trend in the average ratio of male to female population of age 0-4 on the left and 20-24 on the right. These trends are obtained using pooled district-level data from Census 1971-2011. The bottom panel shows the linear fit and scatter plot of the ratio of males to females of the same cohort over a 20 year period. This plot uses state-level data from 1961-2011.

Figure 2: Trends in Crime



Notes: Yearly trend in the average crime rate per 100,000 population. Violence against women is the sum of all crimes under the categories of rape (released since 1971), kidnapping and abduction of females (released since 1988), molestation, sexual harassment, cruelty by husband and relatives (all released since 1995). Non-gender based violence is the sum of crimes under riots, murder, dacoity (available since 1971), male kidnappings (released since 1988), arson and hurt (released since 1995). Property crime is the sum of crimes under the categories of burglary, robbery, theft (all available since 1971). Economic crime is the sum of all crime under the categories of criminal breach of trust, cheating (all available since 1971) and counterfeiting (available since 1995). The vertical lines at 1988 and 1995 indicate these points at which the National Crime Records Bureau released additional categories under total violence against women and non-gender based violence. All definitions are also in Table 11 in Appendix.

Figure 3: Distribution across districts in birth and youth sex ratio, crime rates and single men



Notes: Spatial distribution of the main independent and dependent variables. All maps use 2011 data. The sex ratio of 0-4 and 20-24 is the ratio of male to female population of ages 0-4 and 20-24, respectively. The ratio of single men is the ratio of unmarried men to total males and crime rates are per 100,000 population. Acquisitive crime is the sum of property and economic crime. Definitions of all crimes are in Table 11 in Appendix.

Table 1: Crime as a function of the youth sex ratio: Yearly changes estimation

	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	Total			Against Women			Non-Gender Violence			Property			Economic		
SR 20-24	0.749***	0.785***	0.450**	0.863***	0.877***	0.767***	1.770***	1.871***	0.557*	0.248	0.249	0.341	0.264	0.298	0.342
	(0.221)	(0.220)	(0.228)	(0.241)	(0.241)	(0.231)	(0.278)	(0.279)	(0.301)	(0.229)	(0.228)	(0.241)	(0.213)	(0.213)	(0.226)
SR 0-4	-0.216	-0.100	-0.940**	-0.373	-0.333	0.091	-0.460	-0.062	-1.097*	-0.002	-0.057	-0.647	0.829**	0.952**	-0.504
	(0.406)	(0.404)	(0.460)	(0.360)	(0.362)	(0.351)	(0.512)	(0.514)	(0.566)	(0.445)	(0.445)	(0.465)	(0.396)	(0.389)	(0.430)
N	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062
Adj. R-sq.	0.817	0.819	0.855	0.920	0.920	0.945	0.791	0.797	0.854	0.858	0.858	0.894	0.725	0.727	0.783
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No
State*Year	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Coefficients from estimating specification (1). The dependent variable is the log of the crime rate per capita. SR is the (age-specific) sex ratio, males to females. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are double clustered at the district-level and decade level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 2: Crime as a function of the youth sex ratio: Robustness exercise with neighbouring age-groups

	Total	Against Women	Non-Gender Violence	Property	Economics
SR 20-24	0.324	0.866***	0.417	0.195	0.309
	(0.228)	(0.239)	(0.305)	(0.246)	(0.222)
SR 0-4	-0.970**	-0.016	-1.118*	-0.683	-0.536
	(0.475)	(0.366)	(0.586)	(0.476)	(0.437)
SR 15-19	-0.394**	-0.283*	-0.380	-0.464**	-0.216
	(0.196)	(0.149)	(0.274)	(0.198)	(0.170)
SR 25-34	0.492*	-0.149	0.521	0.572**	0.175
	(0.266)	(0.205)	(0.333)	(0.269)	(0.277)
N	10,062	10,062	10,062	10,062	10,062
Adj. R-sq.	0.855	0.945	0.854	0.894	0.783
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes
State*Year	Yes	Yes	Yes	Yes	Yes

Notes: Coefficients from estimating specification (1) by also including the sex ratio of ages 15-19 and 25-34. The dependent variable is the log of the crime rate per capita. SR is the (age-specific) sex ratio, males to females. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are double clustered at the district-level and decade level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 3: Crime as a function of sex ratio - Decadal changes in crime against sex ratios at start of the decade

	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	Δ Total		Δ Against Women		Δ Non-Gender Violence		Δ Property		Δ Economic	
SR 20-24	0.049	-0.572	1.997**	1.859**	1.213	-0.548	-0.246	-0.256	-0.784	-0.821
	(0.841)	(0.911)	(0.825)	(0.888)	(1.132)	(1.251)	(0.926)	(0.992)	(0.969)	(1.013)
SR 0-4	1.704	0.617	1.400	0.073	0.650	0.431	1.875	1.448	3.736	1.121
	(1.726)	(2.111)	(1.775)	(2.141)	(2.369)	(2.644)	(1.963)	(2.289)	(2.560)	(2.460)
N	1,032	1,032	1,032	1,032	1,032	1,032	1,032	1,032	1,032	1,032
Adj. R-sq.d	0.186	0.318	0.578	0.662	0.403	0.563	0.161	0.261	0.040	0.312
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: This tables uses data for the years 1971,1981, 1991, 2001 and 2011. Coefficients from estimating equation (1) but with the dependent variable as the change in crime rates (in logs) across decades. SR is the (age-specific) sex ratio, males to females. Independent variables take the values of the beginning of the decade. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 4: Crime as a function of sex ratio - Robustness Exercises

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pop Growth	Post Gender Quota	Elasticity	Share	Linear Interpolation	Spline	Cubic
Panel B: Crime Against Women							
SR 20-24	0.802***	0.843***	0.741*	3.038***	0.838***	0.767***	0.901***
	(0.233)	(0.242)	(0.249)	(0.997)	(0.268)	(0.231)	(0.236)
SR 0-4	-0.103	-0.438	0.041	-0.896	-0.097	0.091	0.096
	(0.378)	(0.365)	(0.509)	(1.589)	(0.400)	(0.351)	(0.347)
Adj. R-sq.	0.945	0.921	0.945	0.940	0.945	0.945	0.945
Panel D: Property Crime							
SR 20-24	0.270	0.247	0.430	1.439	0.338	0.341	0.400*
	(0.256)	(0.228)	(0.503)	(1.174)	(0.269)	(0.241)	(0.240)
SR 0-4	-0.721	-0.061	-0.680	-4.273**	-0.968*	-0.647	-0.645
	(0.503)	(0.443)	(0.652)	(2.037)	(0.501)	(0.465)	(0.466)
Adj. R-sq.	0.894	0.858	0.894	0.886	0.894	0.894	0.894

Notes: Coefficients from estimating specification (1). In column (1) we add population growth as an additional control; column (2) we control for a post gender-political quota reform dummy; in column (3) the sex ratio is logs; in column (4) the sex ratio measure is the share of the males to total population by age-group (in logs); in column (5) the sex ratio is the linear interpolation of the Census sex ratio; in column (6) the sex ratio is obtained as the sex ratio interpolated through linear spline and in column (7) the sex ratio is obtained through cubic interpolation. The dependent variable is the log of the crime rate per capita. All regressions include district and year fixed-effects, district linear trends and state-year dummies. SR is the (age-specific) sex ratio, males to females. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. Robust standard errors are in parenthesis and are clustered at the district-decade level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 5: Crime as a function of the sex ratio: Differencing across crime types

	(1)	(2)	(3)	(4)	(5)
Dep var: Crime Rate					
SR 20-24*VAW	0.524***	0.546**	0.542*	0.543*	0.545**
	(0.142)	(0.042)	(0.043)	(0.044)	(0.029)
SR 20-24	-0.102	-0.159	-0.378	-0.346	0.088
	(0.227)	(0.138)	(0.125)	(0.119)	(0.153)
Observations	32,255	32,255	32,255	32,255	32,255
Adjusted R-squared	0.923	0.924	0.924	0.924	0.927
District-Crime FE	Yes	Yes	Yes	Yes	Yes
Crime-Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
District Trend	Yes	Yes	Yes	Yes	Yes
Controls*VAW	No	Yes	No	No	Yes
SR 0-4	No	No	Yes	Yes	Yes
Gender Literacy Gap	No	No	No	Yes	Yes
State-Year Dummies	No	No	No	No	Yes

Notes: Table reports coefficients from estimating specification (2). The dependent variable is the log of crime rates per capita by crime category, district and year. SR is the (age-specific) sex ratio, males to females. The interaction term VAW is a dummy that takes values 1 if the crime category is a gender-based violence crime or a 0 if it is a property crime category. The crime categories are as defined in Table 11 in Appendix. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. Standard-errors are clustered at the district-decade level. Significance values at 1,5 and 10% are shown by ***,** or *.

Table 6: Sex Ratio and Share of Never-Married Population

	Share of Unmarried Males			Share of Unmarried Females		
	(1)	(2)	(3)	(1)	(2)	(3)
	15-44	15-44	20-24	15-44	15-44	20-24
SR 15-44	0.0786			-0.0413		
	(0.0604)			(0.0621)		
SR 20-24		**0.129*** **	**0.192*** **		**0.119*** **	**-0.137*** **
		(0.0328)	**(0.0601)**		**(0.0274)**	**(0.0484)**
SR 25-34		0.158***	0.0165		0.124***	0.0878*
		(0.0406)	(0.0698)		(0.0318)	(0.0506)
SR 35-44		-0.108**	0.116		0.0333	0.281***
		(0.0427)	(0.0765)		(0.0374)	(0.0618)
Mean of the Dep. Var.	0.465		0.627	0.228		0.228
N	744	744		744	744	744
Ad. R- sq	0.511	0.610	0.758	0.655	0.801	0.862
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Census 2011 Dummy	Yes	Yes	Yes	Yes	Yes	Yes

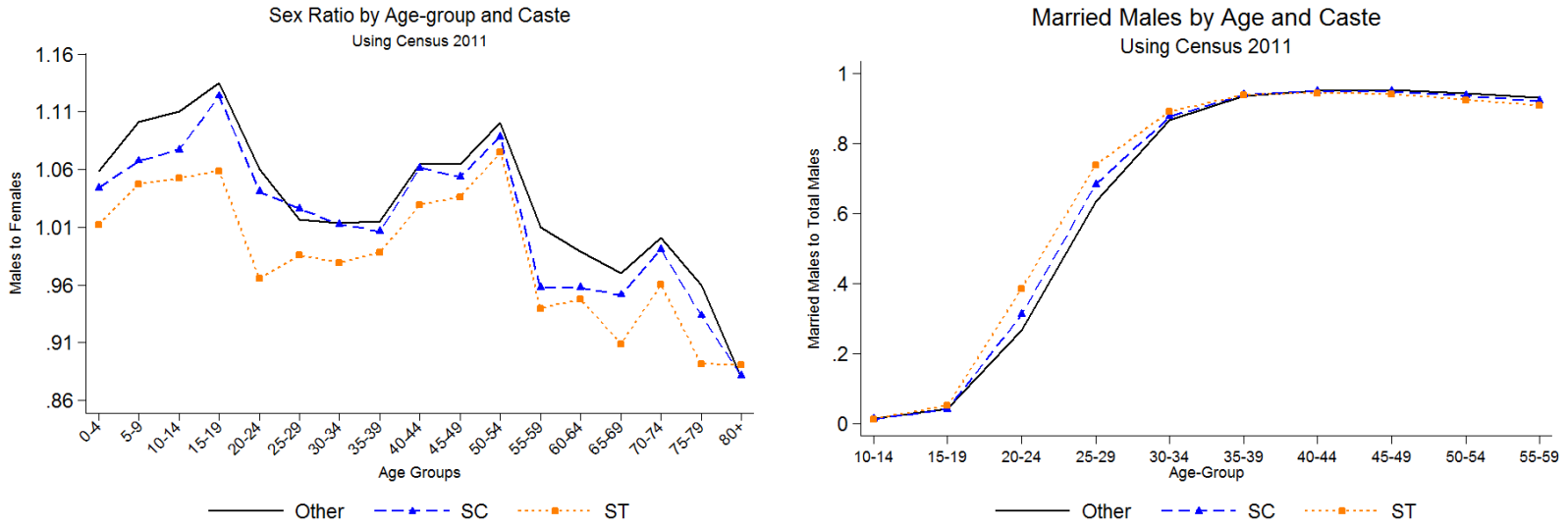
Notes: The dependent variable is the ratio of never married individuals by gender and age over total population of the same age-gender. SR is the (age-specific) sex ratio, males to females. Columns (1) and (2) use the 15-44 age-group sample and columns (3) use the 20-24 age-group sample. We control for district-level rural population, share of SC and ST population, share of literates, literacy gap. In columns (2) and (3) we also control for the sex ratio of those aged 15-19. Robust standard errors in parentheses are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 7: Attitudes Towards Intimate Partner Violence

	Men				Women			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Index	Index	Binary	Binary	Index	Index	Binary	Binary
Post * Child Sex Ratio	1.883**	2.047**	0.533**	0.562**	-0.891	-0.907*	-0.416***	-0.431***
	(0.739)	(0.713)	(0.216)	(0.210)	(0.521)	(0.477)	(0.118)	(0.119)
Post	-1.641*	-1.792**	-0.465*	-0.551**	0.901	1.001**	0.429***	0.488***
	(0.816)	(0.813)	(0.249)	(0.226)	(0.529)	(0.459)	(0.130)	(0.137)
Child Sex Ratio	-1.972	-1.908	-0.590	-0.599	-0.746	-1.041	-0.165	-0.259
	(1.726)	(1.671)	(0.485)	(0.476)	(1.210)	(1.086)	(0.255)	(0.238)
N	53,199	53,199	53,199	53,199	90,184	90,184	90,184	90,184
Adj. R-sq.	0.123	0.124	0.135	0.136	0.138	0.142	0.128	0.131
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE* Hindu	No	Yes	No	Yes	No	Yes	No	Yes
Birth-Year*Hindu	No	Yes	No	Yes	No	Yes	No	Yes
$t_{\delta_1=\beta_2}$	5.42 (0.03)	6.60 (0.02)	7.01 (0.00)	10.71 (0.00)	2.89 (0.11)	4.23 (0.08)	11.65 (0.065)	13.29 (0.00)

Notes: The dependent variable is an index or binary measure of tolerance towards physical abuse. The dependent variables are obtained from the DHS-2005 questions: *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...she burns the food?*; iii) *... she refuses to have sex?*; iv) *...neglects the children?* and, v) *argues?*. Using these responses we create an acceptance index that varies from 0-5 according to each individual's total domains of acceptance of domestic violence. Alternatively, we also use a binary measure which takes the value 0 if an individual never justifies physical abuse and 1 if he/she justifies it in at least 1 circumstance. The child sex ratio is the ratio of males to females of ages 0-4 measured in the state-birth year of each individual. This is gathered from state-level Census information from 1961-2011. Post is a dummy that takes values 1 if an individual was born after 1980 and 0 otherwise. Men and women in the sample are born between 1957 and 1991. All specifications control for individual level controls that include religion dummies, an indicator for rural/urban household, education attainment dummies (no education is the omitted category) and household wealth index. In columns (2) we also include state-Hindu and birth-year*Hindu effects. We also present for each regression the F statistic and p-value for the equality of coefficients $Post * CSR$ and Post. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Figure 4: Sex Ratio and Marriage Rates by Age-Group and Caste



Notes: Using Census 2011 data on sex ratio by age-group and caste. Sex ratio in the ratio of males to females and males marriage rates is the ratio of married males to total male population in the same age-caste group. Caste is defined as male and female population under scheduled caste (SC), scheduled tribe (ST) or others.

Table 8: Marriage quality as a function of sex ratio

	(1) Education Years	(2) Education Gap	(3) Age	(4) Age Gap	(5) Joint Decision	(6) Controlling	(7) Any IPV Index	(8) Any IPV	(9) Height	(10) Height Gap
Panel A: Baseline										
Post*SR	-4.129***	-0.275	-0.461	0.004	-1.332***	0.916***	0.924***	0.710**	-0.571	0.404
	(1.390)	(1.038)	(1.188)	(1.212)	(0.313)	(0.258)	(0.271)	(0.311)	(1.169)	(1.369)
Post	4.386***	0.316	-1.898	0.089	1.287***	-0.907***	-0.957***	-0.692**	0.343	-0.827
	(1.401)	(1.089)	(1.207)	(1.224)	(0.315)	(0.268)	(0.268)	(0.313)	(1.217)	(1.383)
SR	3.049*	0.389	-0.086	-0.436	-0.231	-0.846***	-0.834***	-0.840***	0.154	1.040
	(1.651)	(0.872)	(1.316)	(1.123)	(0.309)	(0.276)	(0.229)	(0.270)	(1.870)	(2.144)
$t_{\delta_2=\beta_4}$	9.32 (0.00)	0.08 (0.78)	0.36 (0.55)	0.60 (0.57)	17.40 (0.00)	12.63 (0.00)	12.23 (0.00)	5.67 (0.02)	0.10 (0.75)	0.03 (0.61)
Panel B: Controlling for Female Education										
Post*SR	-1.597**	-1.597**	-0.693	-0.248	-1.257***	0.796***	0.737***	0.527*	0.302	0.544
	(0.691)	(0.691)	(1.196)	(1.198)	(0.293)	(0.250)	(0.227)	(0.300)	(1.083)	(1.420)
Post	1.713**	1.713**	-1.652	0.356	1.208***	-0.780***	-0.762***	-0.499	-0.524	-0.967
	(0.733)	(0.733)	(1.215)	(1.210)	(0.294)	(0.260)	(0.223)	(0.308)	(1.070)	(1.434)
SR	1.301	1.301	0.075	-0.261	-0.284	-0.796***	-0.724***	-0.754***	-0.337	0.983
	(1.014)	(1.014)	(1.323)	(1.134)	(0.325)	(0.288)	(0.199)	(0.272)	(1.908)	(2.163)
	71,443	71,443	67,614	67,614	71,154	52,779	53,077	53,077	29,854	29,854
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$t_{\delta_2=\beta_4}$	5.42 (0.03)	0.08 (0.78)	0.16 (0.69)	0.06 (0.804)	17.65 (0.00)	9.59 (0.00)	11.17 (0.00)	2.86 (0.09)	0.15 (0.70)	0.20 (0.66)

Notes: The dependent variables are measures of husband quality or spousal gap (e.g. the difference between husband's and wife's age). Using the sample of married women from DHS-2005. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. In Panel A we control for women's religion dummies, whether the woman resides in a rural area and in Panel B we also control for women's educational attainment. All regressions include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Joint-decision making is a dummy that takes values 1 if the woman has any say in the husband decision over his income and 0 otherwise; controlling is a measure of controlling behaviour that takes values 1 if the husband exhibits any of the following behaviours:the husband is jealous if respondent talks with other men; husband accuses respondent of unfaithfulness; husband does not permit respondent to meet her girlfriends; husband tries to limit respondent's contact with family; husband insists on knowing where respondent is; husband doesn't trust respondent with money. Any IPV and IPV are an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. We also present for each regression the F statistic and p-value for the equality of coefficients $Post * SR$ and Post. (Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 9: Intimate-partner violence as a function of sex ratio

	(1) Emotional Index	(2) Physical Index	(3) Sexual Index	(4) Emotional Binary	(5) Physical Binary	(6) Sexual Binary
Panel A: Baseline						
Post*SR	0.175	0.629**	0.242***	0.561*	0.497	1.438***
	(0.106)	(0.233)	(0.071)	(0.318)	(0.371)	(0.355)
Post	-0.179*	-0.665**	-0.242***	-0.575*	-0.484	-1.438***
	(0.102)	(0.234)	(0.069)	(0.323)	(0.372)	(0.346)
SR	-0.126*	-0.594***	-0.258*	-0.562**	-0.565**	-1.210***
	(0.068)	(0.149)	(0.124)	(0.251)	(0.244)	(0.322)
Panel B: Controlling for Female Education						
Post*SR	0.133	0.472**	0.225***	0.412	0.295	1.328***
	(0.101)	(0.205)	(0.064)	(0.306)	(0.368)	(0.332)
Post	-0.135	-0.501**	-0.224***	-0.419	-0.272	-1.325***
	(0.096)	(0.208)	(0.063)	(0.308)	(0.376)	(0.325)
SR	-0.101	-0.502***	-0.248*	-0.503**	-0.464*	-1.167***
	(0.063)	(0.139)	(0.120)	(0.236)	(0.256)	(0.315)
N	53,074	53,075	53,072	53,074	53,075	53,072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of intimate-partner violence by domains: emotional, physical and sexual. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste. For each measure we present an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix-Table 11. Sample of married women of DHS-2005. In Panel A we control for women's religion dummies, whether the woman resides in a rural area and in Panel B we also control for women's educational attainment. All regressions include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Online Appendix to: "Population Sex Ratios and Violence against Women: The Long Run Consequences of Sex Selection in India"

Sofia Amaral
University of Birmingham[2]

Sonia Bhalotra
University of Essex[1]

February, 2018

[*]Centre for Crime, Justice and Policing, University of Birmingham. Corresponding author: Sofia Amaral (amaral.sofiafernando@gmail.com)

[†]ISER & Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, Contact: srbhal@essex.ac.uk

List of Figures

List of Tables

Appendix

Table 10: Missing Brides Index

Survey	a_m	a_w	β	α	σ	Index
1992-1993	22.704	17.180	3.446	5.189	0.899	0.985
1998-1999	23.395	17.402	3.058	4.796	0.899	1.006
d'Albis and de la Croix-2000	24.770	20.200	2.640	6.580	0.921	1.043
2004-2005	23.032	17.935	2.717	4.574	0.899	1.051

Notes: This table presents estimates of the Missing Brides Index developed by d'Albis and De La Croix (2012). We use the measure of fertility rate as in d'Albis and De La Croix (2012) for the years of the survey round (1992-1993; 1998-1999 and 2004-2005). This variable is collected using data from the U.N. population statistics. We created each of the components of the index in d'Albis and De La Croix (2012) using the mean variables of each round of National Family Health Surveys (rounds 1,2,3 for the years 1992-1993, 1998-1999 and 2005-2006, respectively). The index uses information on the average age at marriage for men and women (a_m and a_w); total fertility rate corrected for infant mortality of under five children (β); and age of childbearing (α, which is the mean age at childbearing minus the age at marriage and; the sex ratio (females to males) of pre-marriageable age (i.e. of age 10-14) cohorts (σ).

Table 11: Variable Definitions

Measure	Description
	Panel A: Violence against women and other crime
Crimes against Women	Includes the incidents registered under rape (released since 1971), kidnapping and abduction of females (released since 1988), molestation, sexual harassment, cruelty by husband and relatives (all released since 1995)
Non-gender based Violence	Includes the incidents registered underunder riots, murder, dacoity (available since 1971), male kidnappings (released since 1988), arson and hurt (released since 1995).
Property	Includes the incidents registered under under the categories of burglary, robbery, theft (all available since 1971).
Economic	Includes the incidents registered under under the categories of criminal breach of trust, cheating (all available since 1971) and counterfeiting (available since 1995).
	Panel B: Surveyed intimate partner violence
Humiliation	Spouse ever humiliated her
Threat	Spouse ever threatened her with harm
Insult	Spouse ever insult or make feel bad
Push	Spouse ever pushed, shook or threw something
Slap	Spouse ever slapped
Punch	Spouse ever punched with fist or something harmful
Kick	Spouse ever kicked or dragged
Strangle	Spouse ever tried to strangle or burn
Threat with weapon	Spouse ever threatened with knife/gun or other weapon
Pull Hair	Spouse ever twisted her arm or pulled her hair
Forced sex	Spouse ever physically forced sex when not wanted
Forced sexual Acts	Spouse ever forced other sexual acts when not wanted

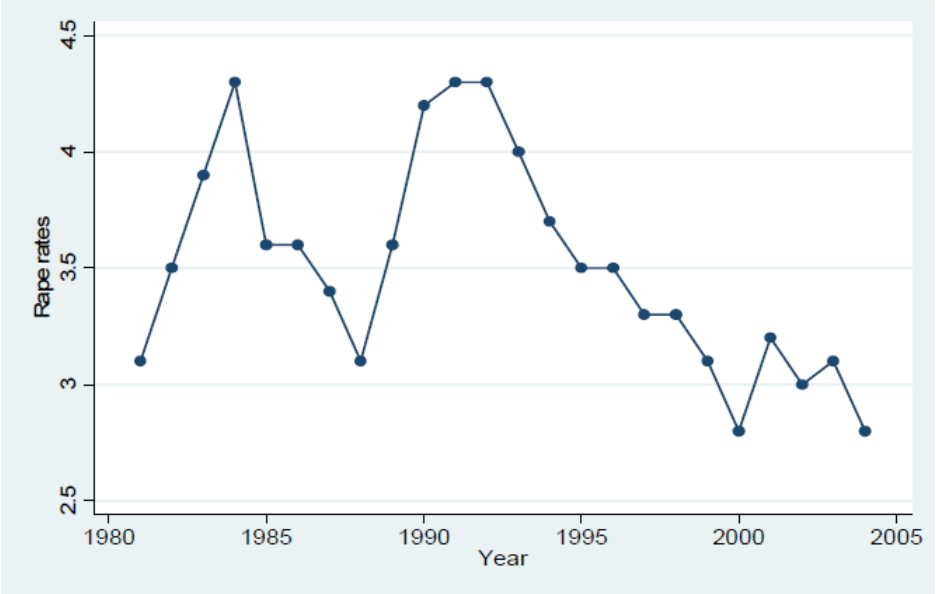
Notes: Panel A presents the definitions of crime rates variables from the National Crime Records Bureau (NCRB). Crime rates is defined as total incidents per 100, 000 population. Crime definitions are from the Indian Penal Code. Panel B present definitions of self-reported measures of intimate-partner violence from the National Family Health Survey, wave III (NFHS). The measures come from DHS questions of the following form: "Did your (last) (husband/partner) ever do any of the following things to you..." for each of the measures outlined above. Using these responses we create variables for emotional, physical and sexual violence as an index (i.e. the sum of responses by domain) and as dummy variables that take values 1 if the respondent was a victim of any of the forms of abuse listed. We also construct an aggregate measure formed of all responses listed in the table.

Table 12: Summary Statistics - By Age-Gender-Crime Type Offender Profile

Age-group Crime Categories	$<$ 18 Years	Between 18-30 Years	Between 30-45 Years	Between 45-60 Years	$>$ 60 Years	Total (by gender)
			Panel A: Male Offenders			
Violence against Women	0.841%	41.826%	28.700%	11.850%	1.968%	85.186%
Non-Gender based Crime	1.088%	43.256%	34.136%	14.836%	1.712%	95.028%
Total IPC Crime	1.063%	43.110%	33.579%	14.530%	1.738%	94.019%
			Panel B: Female Offenders			
Violence against Women	0.082%	4.583%	5.309%	4.018%	0.824%	14.814%
Non-Gender based Crime	0.047%	1.879%	2.066%	0.866%	0.114%	4.972%
Total IPC Crime	0.051%	2.156%	2.398%	1.189%	0.187%	5.981%

Notes: Based on the National Crime Records Bureau Report of 2011. Each cell represents the share of **arrested** offenders by age-group and crime type for males (Panel A) and females (Panel B). Violence against women comprises all Indian Penal Code (IPC) crime under the categories of rape, kidnapping and abduction of women and girls, dowry deaths, molestation, sexual harassment, cruelty by husband and relatives, importation of women and girls. Non-gender based crime comprises arrests for all crimes under acquisitive crime and non-gender based violence as defined in 11. The age-group decomposition shown is the decomposition available from the NCRB reports. For example 94.019% of total IPC crime is committed by males whereas the remaining (5.981%) is committed by females.When looking at VAW, 85.186% is committed by males and 14.814% is committed by females against other females. When looking by age-group, 46.409% of VAW is committed by those age between 18-30, with men in this age-group committing 41.826% and females 4.583%.

Figure 5: Trend in Rape in China



Notes: Figure from Edlund et.al, (2013). Rape rates are defined as counts of rapes registered by the police per 10,000 population. This figure shows a recent trend decline in the rape rate in China which contrast to the case of India.

Table 13: Summary Statistics- Share of unmarried individuals

		Census		NFHS-2005	
		Mean	SD	Mean	SD
Men					
	Age-group: 15-44	0.465	0.053	0.507	0.499
	Age-group: 15-19	0.924	0.080	0.980	0.001
	Age-group: 20-24	**0.627**	**0.171**	**0.756**	**0.004**
	Age-group: 25-34	0.188	0.097	0.384	0.005
	Age-group: 35-44	0.035	0.018	0.086	0.002
Women					
	Age-group: 15-44	0.228	0.078	0.246	0.430
	Age-group: 15-19	0.712	0.175	0.776	0.003
	Age-group: 20-24	**0.228**	**0.138**	**0.309**	**0.003**
	Age-group: 25-34	0.043	0.036	0.053	0.001
	Age-group: 35-44	0.012	0.013	0.011	0.000

Notes: The top panel shows rates of unmarried men by different age-groups and the bottom panel the rates for women. The first 2 columns use district-level Census data from 1991 and 2001. Columns 3 and 4 use data from all men and women surveyed in the NFHS-III across the 16 major states of India.

Table 14: Summary Statistics- Crime and Sex Ratios Data

		Mean	SD	Δ 1972-2010	Δ 1995-2010
Panel A: Crime Rates (per 100,000)					
Total Crime		90.430	58.586	-0.241	-0.152
Total against Women		6.233	8.325	26.15	0.336
	Rape	1.361	1.424	2.007	0.0382
	Domestic Violence*	5.149	4.875		0.845
	Molestation*	3.926	3.818		-0.0230
	Sexual Harassment*	0.784	1.465		0.575
	Kidnapping of Females*	1.497	1.616		0.0973
Total non-gender based		25.61	23.83	1.259	-0.189
	Murder	3.319	2.278	-0.251	-0.372
	Dacoity	1.016	1.685	-0.865	-0.557
	Riots	9.689	10.33	-0.525	-0.532
	Kidnapping of Males*	1.908	1.957		-0.228
	Arson*	1.059	1.112		-0.486
	Hurt*	26.79	26.25		-0.0234
Property		53.37	48.10	-0.695	-0.323
	Robbery	2.610	3.064	-0.561	-0.291
	Thefts	33.41	33.27	-0.657	-0.247
	Burglary	17.35	15.48	-0.776	-0.470
Economic		5.225	4.641	0.232	0.469
	Counterfeiting	1.993	2.197		-0.257
	Cheating	3.098	3.629	1.756	0.860
	Criminal Breach of Trust	0.134	0.718	-0.694	-0.255
Panel B: Sex Ratio					
	SR 0-4	1.054	0.0504	0.058	0.021
	SR 10-19	1.136	0.102	-0.024	0.191
	SR 20-24	1.033	0.111	0.087	0.300
	SR 25-34	1.011	0.0985	0.015	0.265
	SR 35-44	1.119	0.629		0.187
	SR45-64	1.097	0.684		0.235

Notes: Panel A shows the crime rates per 100,000 population. Crime data comes from yearly publications of the NCRB. Panel B shows sex ratios using Census Publications of 1972 to 2010 for age-groups 0-4, 10-19, 15-19, 20-24 and 25-34; and using Census publications from 1991 to 2011 for the age-groups 35-44 and 45-64. * denote crime categories that are not available since 1972. Δ stands for total growth rate over the designated period. Crime definitions are provided in Table 11 in Appendix.

Crime Reporting in India, the NCRB database and the construction of compatible districts across years

We use police reported data from the National Crime Records Bureau (NCRB) of India to measure violence against women and other forms of crime. The NCRB publishes information at the crime-area level for cognizable crimes prescribed under the Indian Penal Code. This is the major source of administrative data on law and order in India. The data released is based on information gathered from two processes. First, once an incident occurs and is reported, the police are required to register a First Information Report (FIR)- see also Iyer et al. (2012) for an overview. The aggregated information from these reports is then gathered by each police station and reported to the NCRB. The information at the crime area level is what we collect from on-line publications.

In order to match these with districts at any given year-state we use the following process. In order to match a given crime area to its district we used general information from the web and government websites to construct an auxiliary dataset to aggregate crime-area units to its districts. For most cases crime areas are large metropolitan cities or rural areas. For example, Kochwe is a city in the district of Ernakulam in Kerala (http://ernakulam.nic.in/history.htm). Thus, crime area statistics of Kochwe are aggregated to the district-level statistics of Ernakulam. This process was cross-checked twice in order to remove any potential errors. It is worth noting that it is not possible to conduct the analysis at a lower administrative level than the district as data reporting is not consistent over time and space for lower administrative areas.

The division of the Indian administrative boundaries for districts and states varies considerably over the period of our sample. In order to compare the same geographic unit over time we aggregate the information to each district boundary of 1971. This process was done via the database constructed by Kumar and Somanathan (2009) and on our own data construction based on publicly available histories of each district. In 1971, there were 306 districts at the in our sample of the 16 major states of India. Our tracking process allows us to track about 84% of the districts since 1971 [1].

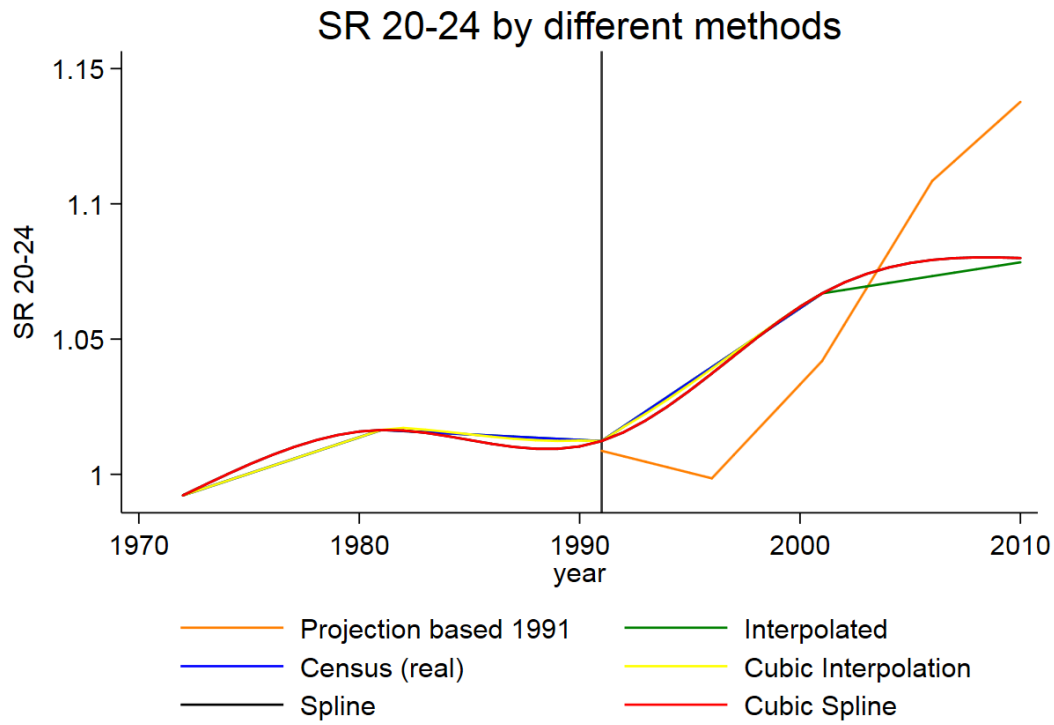
[1] However, our final sample of districts differs from the sample had these splits not occurred. Our sample has higher crime rates, higher sex ratios, less rural population and is less gender equal (see Table 1 of this Appendix.) The states included in the sample are Andhra Pradesh, Bihar, Gujarat, Haryana, Himachal Pradesh, Punjab, Madhya Pradesh, Rajasthan, Uttar Pradesh, Karnataka, Kerala, Tamil Nadu, West Bengal. The newly created states of Telangana, Jharkhand, Chhattisgarh and Uttaranchal are merged with their pre-2001 state boundary definitions.

Table 15: Balance Test

	Not-Traceable		Traceable			
	Mean	SD	Mean	SD	Difference	SD
Violence against Women	10.88	7.369	14.41	8.766	-3.529***	(0.747)
Non-gender based violence	39.61	30.98	43.58	29.26	-3.978	(2.773)
Property	27.65	34.21	34.03	24.34	-6.371**	(2.724)
Economic	4.824	5.560	5.470	5.135	-0.646	(0.492)
Total	82.96	62.10	97.49	53.37	-14.524***	(5.323)
Scheduled Caste (%)	16.14	8.182	17.04	7.033	-0.894	(0.701)
Scheduled Tribe (%)	10.74	18.14	9.901	15.65	0.838	(1.557)
Rural Population (%)	79.90	16.60	76.27	15.21	3.637**	(1.465)
Literacy Rate	52.67	12.98	53.50	10.25	-0.838	(1.074)
Gender gap in Literacy	11.81	0.06	12.27	0.04	-0.455***	(0.070)
SR 0-4	1.074	0.0676	1.079	0.0591	-0.006	(0.006)
SR 10-19	1.126	0.0903	1.143	0.0930	-0.017**	(0.008)
SR 20-24	1.039	0.127	1.073	0.111	-0.035***	(0.011)
SR 25-34	0.985	0.0900	0.998	0.0767	-0.013*	(0.008)
SR 35-44	1.083	0.0921	1.094	0.0843	-0.012	(0.008)
SR 45-64	1.070	0.0885	1.063	0.0733	0.007	(0.007)

Notes: Difference in means between districts traceable and not traceable across 1971, 1981 and 1991 and 2001 . Using data from the 2001 Census. All variable definitions are as defined in Table 1.

Figure 6: Decadal Census sex ratio 20-24: Interpolation Alternatives



Notes: Trend in the sex ratio of 20-24 age group by method: i) projection based on 1991 Census and following the method of Edlund et al, 2013; ii) spline; iii) linear interpolation; iv) cubic interpolation and v) cubic spline.

Table 16: Individual Crime Regressions

	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	Total		Against Women		Non-Gender Violence		Property		Economic	
SR 20-24	0.826**	0.653**	0.932**	1.024***	1.920***	0.774*	0.209	0.549	0.267	0.469
	(0.333)	(0.315)	(0.392)	(0.352)	(0.406)	(0.432)	(0.335)	(0.341)	(0.324)	(0.321)
SR 0-4	-0.371	-1.448**	-0.538	-0.477	-0.337	-1.578*	-0.331	-1.118*	0.697	-0.826
	(0.575)	(0.692)	(0.518)	(0.484)	(0.725)	(0.866)	(0.620)	(0.671)	(0.564)	(0.651)
SR 10-19	-0.220	0.006	0.811**	0.343	-0.467	0.321	-0.138	-0.128	0.037	0.209
	(0.370)	(0.369)	(0.363)	(0.369)	(0.484)	(0.526)	(0.410)	(0.392)	(0.357)	(0.339)
SR 15-19	-0.503*	-0.318	0.148	-0.171	-0.476	-0.232	-0.491*	-0.377	-0.518**	-0.231
	(0.267)	(0.332)	(0.241)	(0.262)	(0.387)	(0.469)	(0.279)	(0.330)	(0.231)	(0.270)
SR 25-34	0.913***	0.449	0.699*	1.019***	0.584	0.989**	0.178	0.447	1.458***	0.930**
	(0.337)	(0.274)	(0.396)	(0.387)	(0.405)	(0.388)	(0.398)	(0.385)	(0.450)	(0.461)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Each cell is a coefficient and standard-error from a separate regression.The dependent variable is the log of the crime rate per capita. SR is the (age-specific) sex ratio, males to females. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 17: Showing all controls in Table 1

	Total	Against Women	Non-gender Based	Property	Economic
SR 20-24	0.062	0.913**	0.864*	0.195	0.301
	(0.329)	(0.398)	(0.475)	(0.337)	(0.315)
SR 0-4	1.423**	-0.418	2.166**	-0.305	0.737
	(0.651)	(0.540)	(0.854)	(0.623)	(0.574)
% SC	0.034***	0.005	0.066***	0.007	0.005
	(0.012)	(0.016)	(0.014)	(0.015)	(0.011)
%ST	0.028***	0.006	0.040***	0.012	0.011
	(0.009)	(0.010)	(0.010)	(0.008)	(0.007)
%Rural	-0.005	-0.002	-0.008	0.011**	0.008
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Literacy	0.010	-0.007	0.023***	-0.005	0.003
	(0.007)	(0.006)	(0.008)	(0.006)	(0.006)
Gender Gap in Literacy	0.019	0.043**	-0.024	0.026	0.004
	(0.012)	(0.017)	(0.015)	(0.022)	(0.017)
Annual Rainfall (log of)	-0.042**	-0.035**	-0.018	-0.053***	-0.037**
	(0.019)	(0.014)	(0.026)	(0.016)	(0.016)
Income p.c.	0.173***	0.060*	0.124***	0.005	0.112***
	(0.033)	(0.033)	(0.042)	(0.033)	(0.030)
Police Strength	1.129***	0.128***	1.213***	0.232***	0.112***
	(0.076)	(0.041)	(0.092)	(0.043)	(0.037)
Election Year	-0.013*	-0.000	-0.040***	0.003	-0.022***
	(0.007)	(0.007)	(0.009)	(0.006)	(0.007)
N	10,062	10,062	10,062	10,062	10,062
Adj. R-sq.	0.291	0.908	0.416	0.758	0.430
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes

Notes: Coefficients as in Table 1. The dependent variable is the log of crime rates per capita. SR is the (age-specific) sex ratio, males to females. We always control for district specific effects and year dummies in addition to controlling for rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs) and, a dummy for state election years, state police per capita and state GDP per capita. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

Table 18: Replication of (Edlund et al., 2013)

	(1) Total	(2) Total	(1) Against Women	(2) Against Women	(1) Non-Gender Violence	(2) Non-Gender Violence	(1) Property	(2) Property	(1) Economic	(2) Economic
SR 10-19	-0.083	(0.347)	-0.061	0.226	-0.023	-0.514	-0.159	0.100	-0.757	0.444
	(0.396)	-0.275	(0.574)	(0.491)	(0.614)	(0.497)	(0.437)	(0.439)	(0.652)	(0.544)
SR 20-24	0.637	(0.486)	1.801**	0.802	1.605*	-1.025	-0.028	-0.181	1.486*	0.976
	(0.557)	-0.178	(0.736)	(0.653)	(0.878)	(0.870)	(0.502)	(0.526)	(0.810)	(0.772)
SR 25-34	-1.008	(0.581)	-1.516	-0.714	-0.916	-0.087	-0.959	-0.535	-2.754***	-0.724
	(0.677)	-0.102	(0.993)	(0.816)	(1.047)	(0.896)	(0.667)	(0.652)	(1.043)	(0.819)
SR 35-44	-1.211	(0.773)	-0.619	-0.318	-0.125	1.185	-0.905	-1.121	-1.539	-0.432
	(0.798)	0.488	(1.088)	(1.083)	(1.211)	(1.052)	(0.800)	(0.889)	(1.089)	(1.154)
SR 45-64	1.486	(0.901)	0.317	0.017	-0.510	0.354	1.825*	1.580*	3.071**	-0.209
	(1.030)		(1.293)	(1.244)	(1.399)	(1.181)	(1.027)	(0.952)	(1.285)	(1.119)
N	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760
Adj. R-sq	0.536	0.652	0.842	0.886	0.632	0.760	0.653	0.701	0.426	0.499
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficients from estimating specification (1) with additional age-specific sex ratios (10-19, 25-34, 35-44 and 45-64) following (Edlund et al., 2013). SR is the (age-specific) sex ratio, males to females. Because age-specific population data from ages 35 onwards are only available since the Census of 1991 this Tables uses data from 1991 onwards only. The dependent variable is the log of the crime rate per capita. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 19: Additional Robustness Checks

	Total			Against Women			Non-Gender Violence			Property			Economic		
	(1)	(2)	from Table 1	(1)	(2)	from Table 1	(1)	(2)	from Table 1	(1)	(2)	from Table 1	(1)	(2)	from Table 1
SR 20-24	0.758***	0.749***	0.450**	0.863***	0.863***	0.767***	1.768***	1.770***	0.557*	0.254	0.248	0.341	0.281	0.264	0.342
	(0.220)	(0.221)	(0.228)	(0.244)	(0.241)	(0.231)	(0.279)	(0.278)	(0.301)	(0.228)	(0.229)	(0.241)	(0.213)	(0.213)	(0.226)
SR 0-4	-0.207	-0.216	-0.940**	-0.452	-0.373	0.091	-0.508	-0.460	-1.097*	-0.010	-0.002	-0.647	0.878**	0.829**	-0.504
	(0.403)	(0.406)	(0.460)	(0.370)	(0.360)	(0.351)	(0.508)	(0.512)	(0.566)	(0.444)	(0.445)	(0.465)	(0.390)	(0.396)	(0.430)
N	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062
Adj. R-sq.	0.817	0.817	0.855	0.920	0.920	0.945	0.789	0.791	0.854	0.857	0.858	0.894	0.725	0.725	0.783
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender Lit. Gap	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No
$State \times Year$	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: Coefficients from estimating specification (1) where in columns (1) we do not include any control, in columns (2) we include as set of controls the share of rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs), election year dummy and state police per capita and GDP per capita as well as gender literacy gap. In columns (3) we present the most parsimounious regression as show in Baseline from Table **??**. Robust standard errors are in parenthesis and are clustered at the district-decade-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 20: Addressing measurement error concerns

	(1)	(2)	(3)	(4)	(5)
	Total	Women	Non-Gender	Property	Economics
			Panel A: Levels		
SR 20-24	1.058***	0.916**	1.860***	0.427	0.038
	(0.390)	(0.448)	(0.464)	(0.387)	(0.376)
SR 0-4	1.015*	0.211	0.522	1.020*	0.526
	(0.560)	(0.616)	(0.681)	(0.598)	(0.657)
Observations	1,290	1,290	1,290	1,290	1,290
Adj. R-sq.	0.688	0.936	0.778	0.823	0.596
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
			Panel B: First- Differences		
SR 20-24	0.406	1.019***	1.274***	-0.012	-0.189
	(0.323)	(0.327)	(0.425)	(0.347)	(0.336)
SR 0-4	0.250	0.347	-0.206	0.571	0.658
	(0.602)	(0.658)	(0.723)	(0.677)	(0.758)
N	1,032	1,032	1,032	1,032	1,032
Adjusted R-squared	0.318	0.661	0.563	0.262	0.311
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: This tables uses data for the years 1971,1981, 1991, 2001 and 2011. Coefficients from estimating equation (1) but with the dependent variable in levels in Panel A and in first-differences in Panel B. SR is the (age-specific) sex ratio, males to females and is measured in levels in Panel A and in first-differences in Panel B. Independent variables are also defined in levels or first-differences, respectively in Panel A and B. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 21: NFHS III-Summary Statistics: Attitudes towards Intimate Partner Violence

	Men				Women			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Acceptance Index	0.943	1.390	0	5	1.147	1.629	0	5
Acceptance Binary	0.403	0.490	0	1	0.422	0.494	0	1
Acceptance by Domain:								
Movement	0.226	0.418	0	1	0.257	0.437	0	1
Children	0.289	0.453	0	1	0.314	0.464	0	1
Argues	0.243	0.429	0	1	0.269	0.444	0	1
Sex	0.0722	0.259	0	1	0.128	0.334	0	1
Food	0.113	0.317	0	1	0.179	0.384	0	1
Controls:								
No Education	0.148	0.355	0	1	0.347	0.476	0	1
Primary Education	0.154	0.361	0	1	0.143	0.350	0	1
Secondary Education	0.544	0.498	0	1	0.407	0.491	0	1
Higher Education	0.154	0.361	0	1	0.103	0.304	0	1
Hindu	0.807	0.395	0	1	0.799	0.401	0	1
Muslim	0.133	0.340	0	1	0.133	0.340	0	1
Christian	0.0220	0.147	0	1	0.0223	0.148	0	1
Other	0.0374	0.190	0	1	0.0417	0.200	0	1
Rural	0.480	0.500	0	1	0.549	0.498	0	1
Household Size	5.902	3.082	1	35	6.055	3.059	1	35
Wealth Index	2.305	0.848	1	3	2.250	0.871	1	3
Age Gap at First Marriage*	22.583	4.651	1	48	17.567	3.616	2	44
N	53,240				90,184			

Notes: Using the sample of men and women born between 1957 and 1991 across the major 17 Indian states. Variables acceptance index and binary are obtained from the DHS questions: *"Do you justify wife-beating if..."*: i) *... goes out without telling you?*; ii) *... burns the food?*; iii) *... refuses to have sex?*; iv) *...neglects the children?* and, v) *argues?*. Acceptance index is a measure that varies from 0-5 according to each response of total domains of acceptance of domestic violence. The binary measure takes values 0 if the respondent never justifies physical abuse and 1 if justifies in at least 1 circumstance. The remaining variables are used as controls in specification (3). All data comes from NFHS-III. *: This information uses only the sample of married individuals (N=32.681 for men and N=71.929 for women.

Table 22: NFHS III-Summary Statistics: Married women Sample

	(1) Mean	(2) SD	(3) Min	(4) Max
Education Years	7.122	5.147	0	24
Education Gap	5.443	4.494	-7	21
Height	164.7	6.448	106.4	196.2
Height Gap	12.61	7.591	-50.50	58.90
Age	31.61	8.014	15	49
Age Gap	5.481	4.054	-14	32
Join-decision	0.749	0.433	0	1
Controlling	0.318	0.466	0	1
Any IPV	0.250	0.433	0	1
Any IPV Index	0.643	1.502	0	11
Emotional	0.102	0.303	0	1
Emotional Index	0.168	0.555	0	3
Physical	0.206	0.405	0	1
Physical Index	0.492	1.203	0	7
Sexual	0.0508	0.220	0	1
Sexual Index	0.0697	0.320	0	2
Sex Ratio	1.038	0.0644	0.810	1.326

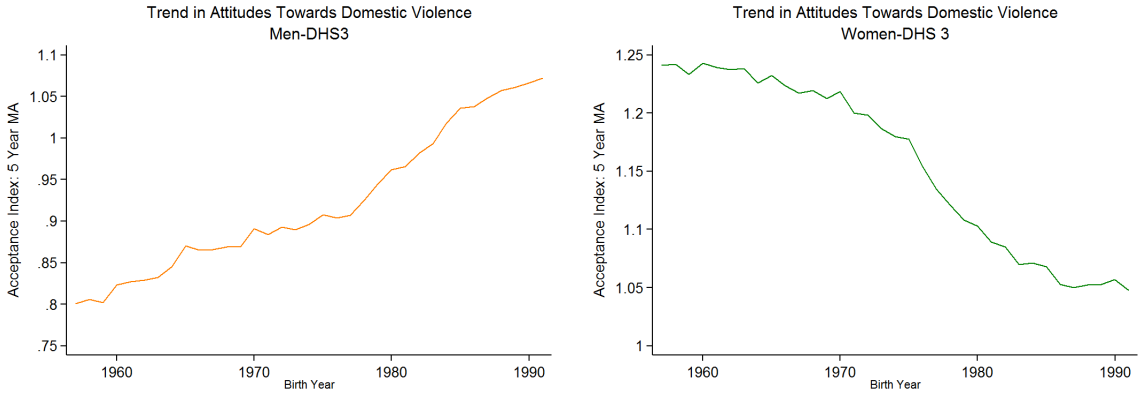
Notes: Using the sample of married women of NFHS III. Variable definitions are in Table 11. Sex Ratio is the ratio of males to females defined by age- group, caste and state.

Table 23: Sex Ratio and Male Attitudes towards domestic violence - Domain Specific Regressions

	(1) Movement	(2) Children	(3) Arguing	(4) Sex	(5) Food
			Panel A: Woman		
Post * Child Sex Ratio	-0.286*	-0.252**	-0.167	-0.109	-0.078
	(0.157)	(0.116)	(0.138)	(0.151)	(0.105)
Post	0.304*	0.272**	0.150	0.106	0.069
	(0.164)	(0.124)	(0.142)	(0.159)	(0.107)
Child Sex Ratio	-0.461	-0.161	-0.050	-0.222	0.147
	(0.265)	(0.323)	(0.279)	(0.287)	(0.261)
N	90,184	90,184	90,184	90,184	90,184
Adj. R-sq.	0.118	0.112	0.087	0.068	0.070
			Panel B: Man		
Post * Child Sex Ratio	0.170	0.425**	0.615***	0.292	0.382***
	(0.223)	(0.197)	(0.169)	(0.171)	(0.108)
Post	-0.066	-0.346	-0.604***	-0.295	-0.330**
	(0.236)	(0.221)	(0.182)	(0.184)	(0.127)
Child Sex Ratio	-0.692	-0.683	-0.136	-0.035	-0.425*
	(0.545)	(0.414)	(0.355)	(0.357)	(0.216)
N	53,199	53,199	53,199	53,199	53,199
Adj. R-sq.	0.091	0.130	0.069	0.030	0.047
Controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Birth-Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a a dummy that takes values 1 if an individual tolerates physical abuse under each of the 5 domains. The dependent variables are obtain from the DHS questions: *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...neglects the children?*; iii) *argues?*; iii) *... she refuses to have sex?* and, v) *...she burns the food?*, respectively for columns (1)-(5). Panel A presents results for the women's sample and panel B for the man's sample. The child sex ratio is the ratio of males to females of ages 0-4 in the state-birth year of each individual. This is gathered from state-Census year information from 1961-2011. Post is a dummy that takes values 1 if an individual was born after 1980 and 0 otherwise. All specifications control religion dummies, an indicator for rural/urban household, education attainment dummies (no education is the omitted category) and household wealth index. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Figure 7: Trends in Attitudes Towards Domestic Violence by Gender



Notes: Trend using the 5 year moving average of the Index measure of acceptance of domestic violence across birth years. Using NFHS-III data for all men and women.

Table 24: Intimate-Partner Violence as a function of sex ratio-separate domains

	Humiliation	Threat	Insult	Push	Slap	Punch	Kick	Strangle	Threat with weapon	Pull Hair	Forced sex	Forced sexual Acts
Post*SR	0.612**	0.367	0.458	0.691*	0.529	0.939***	0.552	1.119	1.604**	0.775**	1.450***	0.971**
	(0.271)	(0.467)	(0.390)	(0.354)	(0.347)	(0.346)	(0.408)	(0.774)	(0.743)	(0.390)	(0.342)	(0.408)
Post	-0.645**	-0.342	-0.464	-0.754**	-0.510	-1.019***	-0.611	-1.170	-1.813**	-0.814**	-1.443***	-0.978**
	(0.269)	(0.461)	(0.384)	(0.358)	(0.345)	(0.361)	(0.421)	(0.826)	(0.777)	(0.389)	(0.335)	(0.397)
SR	-0.597***	-0.036	-0.205	-0.302	-0.645***	-0.638**	-0.613**	-0.302	-0.669	-0.874***	-1.191***	-1.048***
	(0.204)	(0.260)	(0.332)	(0.253)	(0.230)	(0.264)	(0.271)	(0.679)	(0.750)	(0.250)	(0.325)	(0.378)
N	53,074	53,073	53,074	53,062	53,075	53,071	53,071	53,067	53,072	53,061	53,072	53,072
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of intimate-partner violence by separate domains. All dependent variables are dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Using the sample married women questionnaire of DHS-2005. We control for women's religion dummies, whether the woman resides in a rural area. All regressions also include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 25: Marriage Quality as a function of sex ratio- Heterogeneous effect by Non- Scheduled Caste/Scheduled Tribe

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Education Years	Education Gap	Age	Age Gap	Any IPV Index	Any IPV	Joint Decision	Controlling	Height	Heigh Gap
Post*Other Castes*SR	-3.340**	0.920	-0.989	-0.159	0.872**	0.634	-1.332***	0.808***	-1.997	-1.011
	(1.421)	(0.945)	(1.071)	(1.175)	(0.340)	(0.397)	(0.344)	(0.306)	(1.696)	(1.604)
Other Castes*SR	3.495*	0.772	-0.370	-0.844	-0.624**	-0.598**	-0.225	-0.802***	0.201	1.448
	(1.672)	(0.820)	(1.319)	(1.178)	(0.257)	(0.289)	(0.304)	(0.285)	(2.035)	(2.035)
Post*SR	-1.218	-0.848	-0.407	-0.886	1.254**	1.059**	-0.983**	1.052***	5.886**	6.950*
	(1.398)	(1.556)	(2.023)	(1.915)	(0.568)	(0.467)	(0.403)	(0.381)	(2.700)	(3.708)
N	71,449	71,443	67,614	67,614	53,077	53,064	71,154	52,779	29,855	29,855
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of husbands quality or spousal gap (e.g. the difference between husband's and wife's age). Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Other Castes is a dummy that takes values 1 if the woman is from a non SC/ST caste and 0 if it is. Using the sample married women questionnaire of DHS-2005. We control for women's religion dummies, whether the woman resides in a rural area. All regressions also include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. The interaction term is a dummy that takes values 1 if a woman belongs to the category other caste. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Joint-decision making is a dummy that takes values 1 if the woman has any say in the husband decision over his income and 0 otherwise; controlling is a measure of controlling behaviour that takes values 1 if the husband exhibits any of the following behaviours:the husband is jealous if respondent talks with other men; husband accuses respondent of unfaithfulness; husband does not permit respondent to meet her girlfriends; husband tries to limit respondent's contact with family; husband insists on knowing where respondent is; husband doesn't trust respondent with money. Any IPV and IPV are an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 26: Marriage quality as a function of sex ratio- disaggregated measures of Intimate-Partner Violence with effects by Non-Scheduled Caste/Scheduled Tribe

	(1) Emotional Index	(2) Physical Index	(3) Sexual Index	(4) Emotional	(5) Physical	(6) Sexual
Post*Other Castes*SR	0.197*	0.573*	0.466*	0.112*	0.124	0.634
	(0.109)	(0.277)	(0.225)	(0.055)	(0.116)	(0.397)
Other Castes* SR	-0.089	-0.434**	-0.315*	-0.070	-0.130	-0.598**
	(0.073)	(0.193)	(0.155)	(0.047)	(0.081)	(0.289)
Post*SR	0.091	0.967*	0.765*	0.117	0.241	1.059**
	(0.187)	(0.498)	(0.390)	(0.101)	(0.164)	(0.467)
Observations	53,074	53,075	53,075	53,073	53,065	53,064
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of intimate-partner violence by domains: emotional, physical and sexual. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Other Castes is a dummy that takes values 1 if the woman is from a non SC/ST caste and 0 if it is. For each measure we present an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix-Table **??**. Using the sample married women questionnaire of DHS-2005. In Panel A we control for women's religion dummies, whether the woman resides in a rural area and in Panel B we also control for women's educational attainment. All regressions include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 27: Attitudes Towards Intimate-Partner Violence by Marriage Status

	Men		Women	
	Index	Binary	Index	Binary
Married	0.214***	0.071***	0.485***	0.138***
	(0.030)	(0.011)	(0.060)	(0.013)
N	47,879	47,879	90,645	90,645

The dependent variable is a a dummy that takes values 1 if an individual tolerates physical abuse under each of the 5 domains. The dependent variables are obtain from the DHS questions: *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...neglects the children?*; iii) *argues?*; iii) *... she refuses to have sex?* and, v) *...she burns the food?*, respectively for columns (1)-(5). We present results using the index and binary measures as defined in the text. Married is a dummy that takes values one if the individual is married or not. All specifications control religion dummies, an indicator for rural/urban household, education attainment dummies (no education is the omitted category) and household wealth index. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Instrumental variables

We explore instrumenting the youth sex ratio with the availability of ultrasound, which Bhalotra and Cochrane (2010) argue was driven by exogenous changes in import and licensing regulations. The first post-ultrasound cohorts, born in the mid-1980s have matured to age 20-24 by the census of 2011. However, although we see a strong trend in the 0-4 sex ratio in the three decades since the mid-80s, because we use the 20-24 sex ratio, we have only one post-ultrasound cohort in our data. Also district variation in ultrasound availability will be endogenous (a function of demand for sex-selective abortion), so the exogenous variation in availability of ultrasound is at the national level. As a result, the power of this potential instrument is weak. We explored interactions of an indicator for post-ultrasound cohorts with an indicator for the Northwest of the country, known to have stronger son preference (Dyson and Moore, 1983; Sen, 2003). We also investigated alternative instruments, motivated by previous work. These include rainfall shocks (Rose, 1999), gold price shocks (Bhalotra et al., 2016a) and prices of rice and wheat weighted by land area under these crops (Basu, 1992; Bardhan, 1974), all measured in the birth years of the cohorts who are 20-24 years old in each census year. However we continued to face a weak instruments problem. lthough we have multiple instruments, the first-stage results reveal that we have a weak instruments problem with an F-statistic well below 10 (bottom panel of Table 28, 29 and 30). Nonetheless, we observe a positive and significant coefficient on the instrumented youth sex ratio for violence against women in the second-stage estimates. However, these estimates are imprecise and we do not rely upon them. Results are below in Tables 28- 30.

Table 28: 2SLS Estimates - Set of Instruments 1

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	1.613	10.717	-4.366	-0.834	8.526
	(5.528)	(7.252)	(7.898)	(6.292)	(5.806)
First-Stage Dep. Var.:	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.008	-0.008	-0.008	-0.008	-0.008
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Rice	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.002	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D.Rainfall	0.016*	0.016*	0.016*	0.016*	0.016*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
D.Rice	0.001	0.001	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
D. Wheat	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Post*Rainfall	0.005	0.005	0.005	0.005	0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Post*Rice	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post*Wheat	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Post*D.Rainfall	-0.014	-0.014	-0.014	-0.014	-0.014
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Post*D.Rice	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Post*D.Wheat	-0.005	-0.005	-0.005	-0.005	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Kleibergen-Paap F-stat	0.778	0.778	0.778	0.778	0.778
N	4,528	4,528	4,528	4,528	4,528
Adj R-sq	0.896	0.931	0.890	0.913	0.805
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: The top part of the table reports the second-stage estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of the crime rate per capita. The set of instruments measured at birth are log of total rainfall, $price \times cultivatedarea$ of rice and wheat, the 1 year difference of these instruments and the interactions of with a dummy that takes values 1 for the birth years post 1980. We control for controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 29: 2SLS Estimates- Set of Instruments 2

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	-2.532	1.768	-4.267	-1.291	2.218
	(5.069)	(4.699)	(6.019)	(5.356)	(5.107)
First-Stage- Dep Var	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$Fertilizer_{Nitrogen}$	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Fertilizer_{Phosphorus}$	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Fertilizer_{Potassium}$	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rice	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	4,599	4,599	4,599	4,599	4,599
Adj. R-sq.	0.896	0.946	0.890	0.913	0.834
Kleibergen-Paap F-stat	1.137	1.137	1.137	1.137	1.137
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: The top part of the table reports the second-stage estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of crime rates per capita. The set of instruments measured at birth are *price* $\times$ *quantities* sold of fertilizers nitrogen, phosphorus and potassium measured in Rs/ton. Fertilizer prices are controlled by the Government of India and are based on reported maximum sale prices of fertilizer compounds adjusted for the proportion of the nutrient present in each compound. We always control for district-specific linear trends, state-year dummies, district specific effects and year dummies in addition to controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

Table 30: 2SLS Estimates - Set of Instruments 3

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	3.313	8.684	-1.906	-1.012	12.479
	(6.252)	(8.828)	(7.158)	(7.585)	(9.160)
First-Stage- Dep. Var.	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.006	-0.006	-0.006	-0.006	-0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Rice	0.001	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D. Rainfall	0.012	0.012	0.012	0.012	0.012
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
D. Rice	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D. Wheat	-0.005	-0.005	-0.005	-0.005	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
N	4,502	4,502	4,502	4,502	4,502
Adj. R-sq.	0.893	0.936	0.892	0.913	0.768
Kleibergen-Paap F-stat	0.507	0.507	0.507	0.507	0.507
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: Top part of the table report IV estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of crime rates per capita. The set of instruments measured at birth are log of total rainfall, $price \times cultivatedarea$ of rice and wheat and the 1 year difference of these instruments. We always control for district-specific linear trends, state-year dummies, district specific effects and year dummies in addition to controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

Online Appendix to: "Population Sex Ratios and Violence against Women: The Long Run Consequences of Sex Selection in India"

Sofia Amaral
University of Essex[†]

Sonia Bhalotra
University of Essex[*]

February, 2018

[*]Centre for Crime, Justice and Policing, University of Birmingham. Corresponding author: Sofia Amaral (amaral.sofiafernando@gmail.com)

[†]ISER & Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, Contact: srbhal@essex.ac.uk

List of Figures

List of Tables

Appendix A: Additional Aggreate Level Figures and Tables

Table 1: **Missing Brides Index**

Survey	a_m	a_w	β	α	σ	Index
1992-1993	22.704	17.180	3.446	5.189	0.899	0.985
1998-1999	23.395	17.402	3.058	4.796	0.899	1.006
d'Albis and de la Croix-2000	24.770	20.200	2.640	6.580	0.921	1.043
2004-2005	23.032	17.935	2.717	4.574	0.899	1.051

Notes: This table presents estimates of the Missing Brides Index developed by d'Albis and De La Croix (2012). We use the measure of fertility rate as in d'Albis and De La Croix (2012) for the years of the survey round (1992-1993; 1998-1999 and 2004-2005). This variable is collected using data from the U.N. population statistics. We created each of the components of the index in d'Albis and De La Croix (2012) using the mean variables of each round of National Family Health Surveys (rounds 1,2,3 for the years 1992-1993, 1998-1999 and 2005-2006, respectively). The index uses information on the average age at marriage for men and women (a_m and a_w); total fertility rate corrected for infant mortality of under five children (β); and age of childbearing (α, which is the mean age at childbearing minus the age at marriage and; the sex ratio (females to males) of pre-marriageable age (i.e. of age 10-14) cohorts (σ).

Table 2: **Variable Definitions**

Measure	Description
	Panel A: Violence against women and other crime
Crimes against Women	Includes the incidents registered under rape (released since 1971), kidnapping and abduction of females (released since 1988), molestation, sexual harassment, cruelty by husband and relatives (all released since 1995)
Non-gender based Violence	Includes the incidents registered underunder riots, murder, dacoity (available since 1971), male kidnappings (released since 1988), arson and hurt (released since 1995).
Property	Includes the incidents registered under under the categories of burglary, robbery, theft (all available since 1971).
Economic	Includes the incidents registered under under the categories of criminal breach of trust, cheating (all available since 1971) and counterfeiting (available since 1995).
	Panel B: Surveyed intimate partner violence
Humiliation	Spouse ever humiliated her
Threat	Spouse ever threatened her with harm
Insult	Spouse ever insult or make feel bad
Push	Spouse ever pushed, shook or threw something
Slap	Spouse ever slapped
Punch	Spouse ever punched with fist or something harmful
Kick	Spouse ever kicked or dragged
Strangle	Spouse ever tried to strangle or burn
Threat with weapon	Spouse ever threatened with knife/gun or other weapon
Pull Hair	Spouse ever twisted her arm or pulled her hair
Forced sex	Spouse ever physically forced sex when not wanted
Forced sexual Acts	Spouse ever forced other sexual acts when not wanted

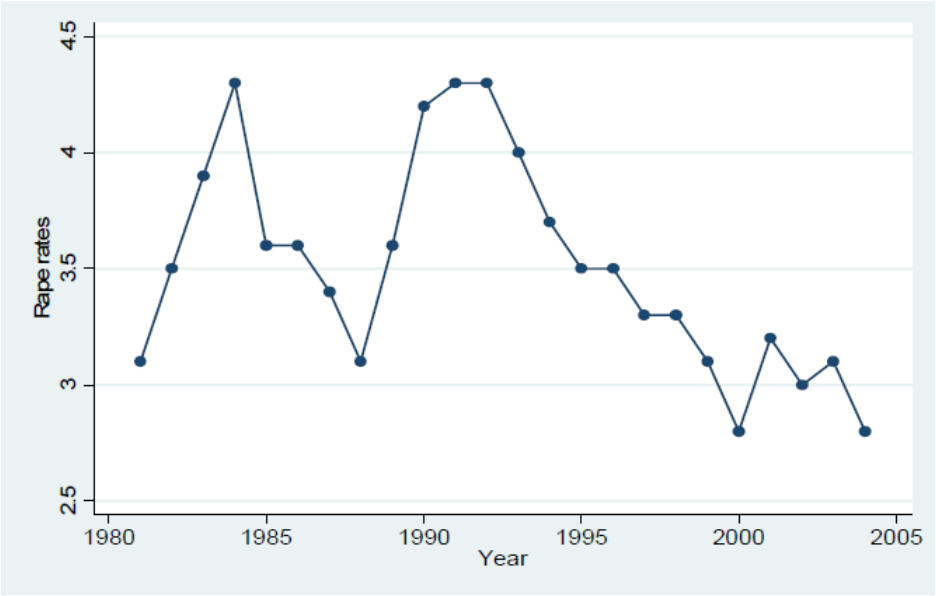
Notes: Panel A presents the definitions of crime rates variables from the National Crime Records Bureau (NCRB). Crime rates is defined as total incidents per 100, 000 population. Crime definitions are from the Indian Penal Code. Panel B present definitions of self-reported measures of intimate-partner violence from the National Family Health Survey, wave III (NFHS). The measures come from DHS questions of the following form: "Did your (last) (husband/partner) ever do any of the following things to you..." for each of the measures outlined above. Using these responses we create variables for emotional, physical and sexual violence as an index (i.e. the sum of responses by domain) and as dummy variables that take values 1 if the respondent was a victim of any of the forms of abuse listed. We also construct an aggregate measure formed of all responses listed in the table.

Table 3: **Summary Statistics - By Age-Gender-Crime Type Offender Profile**

Age-group Crime Categories	< 18 Years	Between 18-30 Years	Between 30-45 Years	Between 45-60 Years	> 60 Years	Total (by gender)
			Panel A: Male Offenders			
Violence against Women	0.841%	41.826%	28.700%	11.850%	1.968%	85.186%
Non-Gender based Crime	1.088%	43.256%	34.136%	14.836%	1.712%	95.028%
Total IPC Crime	1.063%	43.110%	33.579%	14.530%	1.738%	94.019%
			Panel B: Female Offenders			
Violence against Women	0.082%	4.583%	5.309%	4.018%	0.824%	14.814%
Non-Gender based Crime	0.047%	1.879%	2.066%	0.866%	0.114%	4.972%
Total IPC Crime	0.051%	2.156%	2.398%	1.189%	0.187%	5.981%

Notes: Based on the National Crime Records Bureau Report of 2011. Each cell represents the share of **arrested** offenders by age-group and crime type for males (Panel A) and females (Panel B). Violence against women comprises all Indian Penal Code (IPC) crime under the categories of rape, kidnapping and abduction of women and girls, dowry deaths, molestation, sexual harassment, cruelty by husband and relatives, importation of women and girls. Non-gender based crime comprises arrests for all crimes under acquisitive crime and non-gender based violence as defined in 2. The age-group decomposition shown is the decomposition available from the NCRB reports. For example 94.019% of total IPC crime is committed by males whereas the remaining (5.981%) is committed by females.When looking at VAW, 85.186% is committed by males and 14.814% is committed by females against other females. When looking by age-group, 46.409% of VAW is committed by those age between 18-30, with men in this age-group committing 41.826% and females 4.583%.

Figure 1: **Trend in Rape in China**



Notes: Figure from Edlund et.al, (2013). Rape rates are defined as counts of rapes registered by the police per 10,000 population. This figure shows a recent trend decline in the rape rate in China which contrast to the case of India.

Table 4: **Summary Statistics- Share of unmarried individuals**

		Census		NFHS-2005	
		Mean	SD	Mean	SD
Men					
	Age-group: 15-44	0.465	0.053	0.507	0.499
	Age-group: 15-19	0.924	0.080	0.980	0.001
	Age-group: 20-24	**0.627**	**0.171**	**0.756**	**0.004**
	Age-group: 25-34	0.188	0.097	0.384	0.005
	Age-group: 35-44	0.035	0.018	0.086	0.002
Women					
	Age-group: 15-44	0.228	0.078	0.246	0.430
	Age-group: 15-19	0.712	0.175	0.776	0.003
	Age-group: 20-24	**0.228**	**0.138**	**0.309**	**0.003**
	Age-group: 25-34	0.043	0.036	0.053	0.001
	Age-group: 35-44	0.012	0.013	0.011	0.000

Notes: The top panel shows rates of unmarried men by different age-groups and the bottom panel the rates for women. The first 2 columns use district-level Census data from 1991 and 2001. Columns 3 and 4 use data from all men and women surveyed in the NFHS-III across the 16 major states of India.

Table 5: **Summary Statistics- Crime and Sex Ratios Data**

		Mean	SD	Δ 1972-2010	Δ 1995-2010
Panel A: Crime Rates (per 100,000)					
Total Crime		90.430	58.586	-0.241	-0.152
Total against Women		6.233	8.325	26.15	0.336
	Rape	1.361	1.424	2.007	0.0382
	Domestic Violence*	5.149	4.875		0.845
	Molestation*	3.926	3.818		-0.0230
	Sexual Harassment*	0.784	1.465		0.575
	Kidnapping of Females*	1.497	1.616		0.0973
Total non-gender based		25.61	23.83	1.259	-0.189
	Murder	3.319	2.278	-0.251	-0.372
	Dacoity	1.016	1.685	-0.865	-0.557
	Riots	9.689	10.33	-0.525	-0.532
	Kidnapping of Males*	1.908	1.957		-0.228
	Arson*	1.059	1.112		-0.486
	Hurt*	26.79	26.25		-0.0234
Property		53.37	48.10	-0.695	-0.323
	Robbery	2.610	3.064	-0.561	-0.291
	Thefts	33.41	33.27	-0.657	-0.247
	Burglary	17.35	15.48	-0.776	-0.470
Economic		5.225	4.641	0.232	0.469
	Counterfeiting	1.993	2.197		-0.257
	Cheating	3.098	3.629	1.756	0.860
	Criminal Breach of Trust	0.134	0.718	-0.694	-0.255
Panel B: Sex Ratio					
	SR 0-4	1.054	0.0504	0.058	0.021
	SR 10-19	1.136	0.102	-0.024	0.191
	SR 20-24	1.033	0.111	0.087	0.300
	SR 25-34	1.011	0.0985	0.015	0.265
	SR 35-44	1.119	0.629		0.187
	SR45-64	1.097	0.684		0.235

Notes: Panel A shows the crime rates per 100,000 population. Crime data comes from yearly publications of the NCRB. Panel B shows sex ratios using Census Publications of 1972 to 2010 for age-groups 0-4, 10-19, 15-19, 20-24 and 25-34; and using Census publications from 1991 to 2011 for the age-groups 35-44 and 45-64. * denote crime categories that are not available since 1972. Δ stands for total growth rate over the designated period. Crime definitions are provided in Table 2 in Appendix.

Crime Reporting in India, the NCRB database and the construction of compatible districts across years

We use police reported data from the National Crime Records Bureau (NCRB) of India to measure violence against women and other forms of crime. The NCRB publishes information at the crime-area level for cognizable crimes prescribed under the Indian Penal Code. This is the major source of administrative data on law and order in India. The data released is based on information gathered from two processes. First, once an incident occurs and is reported, the police are required to register a First Information Report (FIR)- see also Iyer et al. (2012) for an overview. The aggregated information from these reports is then gathered by each police station and reported to the NCRB. The information at the crime area level is what we collect from on-line publications.

In order to match these with districts at any given year-state we use the following process. In order to match a given crime area to its district we used general information from the web and government websites to construct an auxiliary dataset to aggregate crime-area units to its districts. For most cases crime areas are large metropolitan cities or rural areas. For example, Kochwe is a city in the district of Ernakulam in Kerala (http://ernakulam.nic.in/history.htm). Thus, crime area statistics of Kochwe are aggregated to the district-level statistics of Ernakulam. This process was cross-checked twice in order to remove any potential errors. It is worth noting that it is not possible to conduct the analysis at a lower administrative level than the district as data reporting is not consistent over time and space for lower administrative areas.

The division of the Indian administrative boundaries for districts and states varies considerably over the period of our sample. In order to compare the same geographic unit over time we aggregate the information to each district boundary of 1971. This process was done via the database constructed by Kumar and Somanathan (2009) and on our own data construction based on publicly available histories of each district. In 1971, there were 306 districts at the in our sample of the 16 major states of India. Our tracking process allows us to track about 84% of the districts since 1971 [1].

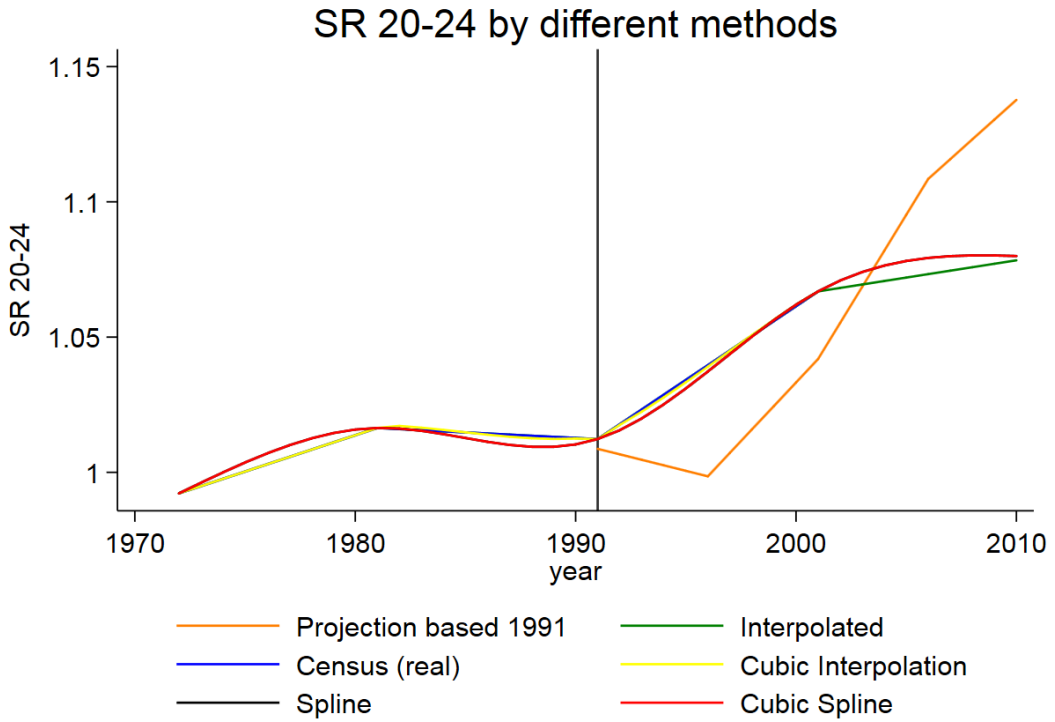
[1]However, our final sample of districts differs from the sample had these splits not occurred. Our sample has higher crime rates, higher sex ratios, less rural population and is less gender equal (see Table 1 of this Appendix.) The states included in the sample are Andhra Pradesh, Bihar, Gujarat, Haryana, Himachal Pradesh, Punjab, Madhya Pradesh, Rajasthan, Uttar Pradesh, Karnataka, Kerala, Tamil Nadu, West Bengal. The newly created states of Telangana, Jharkhand, Chhattisgarh and Uttaranchal are merged with their pre-2001 state boundary definitions.

Table 6: **Balance Test**

	Not-Traceable		Traceable			
	Mean	SD	Mean	SD	Difference	SD
Violence against Women	10.88	7.369	14.41	8.766	-3.529***	(0.747)
Non-gender based violence	39.61	30.98	43.58	29.26	-3.978	(2.773)
Property	27.65	34.21	34.03	24.34	-6.371**	(2.724)
Economic	4.824	5.560	5.470	5.135	-0.646	(0.492)
Total	82.96	62.10	97.49	53.37	-14.524***	(5.323)
Scheduled Caste (%)	16.14	8.182	17.04	7.033	-0.894	(0.701)
Scheduled Tribe (%)	10.74	18.14	9.901	15.65	0.838	(1.557)
Rural Population (%)	79.90	16.60	76.27	15.21	3.637**	(1.465)
Literacy Rate	52.67	12.98	53.50	10.25	-0.838	(1.074)
Gender gap in Literacy	11.81	0.06	12.27	0.04	-0.455***	(0.070)
SR 0-4	1.074	0.0676	1.079	0.0591	-0.006	(0.006)
SR 10-19	1.126	0.0903	1.143	0.0930	-0.017**	(0.008)
SR 20-24	1.039	0.127	1.073	0.111	-0.035***	(0.011)
SR 25-34	0.985	0.0900	0.998	0.0767	-0.013*	(0.008)
SR 35-44	1.083	0.0921	1.094	0.0843	-0.012	(0.008)
SR 45-64	1.070	0.0885	1.063	0.0733	0.007	(0.007)

Notes: Difference in means between districts traceable and not traceable across 1971, 1981 and 1991 and 2001 . Using data from the 2001 Census. All variable definitions are as defined in Table 1.

Figure 2: **Decadal Census sex ratio 20-24: Interpolation Alternatives**



Notes: Trend in the sex ratio of 20-24 age group by method: i) projection based on 1991 Census and following the method of Edlund et al, 2013; ii) spline; iii) linear interpolation; iv) cubic interpolation and v) cubic spline.

Table 7: **Individual Crime Regressions**

	(1) Total	(2) Total	(1) Against Women	(2) Against Women	(1) Non-Gender Violence	(2) Non-Gender Violence	(1) Property	(2) Property	(1) Economic	(2) Economic
SR 20-24	0.826**	0.653**	0.932**	1.024***	1.920***	0.774*	0.209	0.549	0.267	0.469
	(0.333)	(0.315)	(0.392)	(0.352)	(0.406)	(0.432)	(0.335)	(0.341)	(0.324)	(0.321)
SR 0-4	-0.371	-1.448**	-0.538	-0.477	-0.337	-1.578*	-0.331	-1.118*	0.697	-0.826
	(0.575)	(0.692)	(0.518)	(0.484)	(0.725)	(0.866)	(0.620)	(0.671)	(0.564)	(0.651)
SR 10-19	-0.220	0.006	0.811**	0.343	-0.467	0.321	-0.138	-0.128	0.037	0.209
	(0.370)	(0.369)	(0.363)	(0.369)	(0.484)	(0.526)	(0.410)	(0.392)	(0.357)	(0.339)
SR 15-19	-0.503*	-0.318	0.148	-0.171	-0.476	-0.232	-0.491*	-0.377	-0.518**	-0.231
	(0.267)	(0.332)	(0.241)	(0.262)	(0.387)	(0.469)	(0.279)	(0.330)	(0.231)	(0.270)
SR 25-34	0.913***	0.449	0.699*	1.019***	0.584	0.989**	0.178	0.447	1.458***	0.930**
	(0.337)	(0.274)	(0.396)	(0.387)	(0.405)	(0.388)	(0.398)	(0.385)	(0.450)	(0.461)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Each cell is a coefficient and standard-error from a separate regression.The dependent variable is the log of the crime rate per capita. SR is the (age-specific) sex ratio, males to females. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 8: **Showing all controls in Table ??**

	Total	Against Women	Non-gender Based	Property	Economic
SR 20-24	0.062	0.913**	0.864*	0.195	0.301
	(0.329)	(0.398)	(0.475)	(0.337)	(0.315)
SR 0-4	1.423**	-0.418	2.166**	-0.305	0.737
	(0.651)	(0.540)	(0.854)	(0.623)	(0.574)
% SC	0.034***	0.005	0.066***	0.007	0.005
	(0.012)	(0.016)	(0.014)	(0.015)	(0.011)
%ST	0.028***	0.006	0.040***	0.012	0.011
	(0.009)	(0.010)	(0.010)	(0.008)	(0.007)
%Rural	-0.005	-0.002	-0.008	0.011**	0.008
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Literacy	0.010	-0.007	0.023***	-0.005	0.003
	(0.007)	(0.006)	(0.008)	(0.006)	(0.006)
Gender Gap in Literacy	0.019	0.043**	-0.024	0.026	0.004
	(0.012)	(0.017)	(0.015)	(0.022)	(0.017)
Annual Rainfall (log of)	-0.042**	-0.035**	-0.018	-0.053***	-0.037**
	(0.019)	(0.014)	(0.026)	(0.016)	(0.016)
Income p.c.	0.173***	0.060*	0.124***	0.005	0.112***
	(0.033)	(0.033)	(0.042)	(0.033)	(0.030)
Police Strength	1.129***	0.128***	1.213***	0.232***	0.112***
	(0.076)	(0.041)	(0.092)	(0.043)	(0.037)
Election Year	-0.013*	-0.000	-0.040***	0.003	-0.022***
	(0.007)	(0.007)	(0.009)	(0.006)	(0.007)
N	10,062	10,062	10,062	10,062	10,062
Adj. R-sq.	0.291	0.908	0.416	0.758	0.430
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes	Yes

Notes: Coefficients as in Table **??**. The dependent variable is the log of crime rates per capita. SR is the (age-specific) sex ratio, males to females. We always control for district specific effects and year dummies in addition to controlling for rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs) and, a dummy for state election years, state police per capita and state GDP per capita. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

Table 9: **Replication of (Edlund et al., 2013)**

	(1) Total	(2) Total	(1) Against Women	(2) Against Women	(1) Non-Gender Violence	(2) Non-Gender Violence	(1) Property	(2) Property	(1) Economic	(2) Economic
SR 10-19	-0.083	(0.347)	-0.061	0.226	-0.023	-0.514	-0.159	0.100	-0.757	0.444
	(0.396)	-0.275	(0.574)	(0.491)	(0.614)	(0.497)	(0.437)	(0.439)	(0.652)	(0.544)
SR 20-24	0.637	(0.486)	1.801**	0.802	1.605*	-1.025	-0.028	-0.181	1.486*	0.976
	(0.557)	-0.178	(0.736)	(0.653)	(0.878)	(0.870)	(0.502)	(0.526)	(0.810)	(0.772)
SR 25-34	-1.008	(0.581)	-1.516	-0.714	-0.916	-0.087	-0.959	-0.535	-2.754***	-0.724
	(0.677)	-0.102	(0.993)	(0.816)	(1.047)	(0.896)	(0.667)	(0.652)	(1.043)	(0.819)
SR 35-44	-1.211	(0.773)	-0.619	-0.318	-0.125	1.185	-0.905	-1.121	-1.539	-0.432
	(0.798)	0.488	(1.088)	(1.083)	(1.211)	(1.052)	(0.800)	(0.889)	(1.089)	(1.154)
SR 45-64	1.486	(0.901)	0.317	0.017	-0.510	0.354	1.825*	1.580*	3.071**	-0.209
	(1.030)		(1.293)	(1.244)	(1.399)	(1.181)	(1.027)	(0.952)	(1.285)	(1.119)
N	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760	4,760
Adj. R-sq	0.536	0.652	0.842	0.886	0.632	0.760	0.653	0.701	0.426	0.499
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficients from estimating specification (1) with additional age-specific sex ratios (10-19, 25-34, 35-44 and 45-64) following (Edlund et al., 2013). SR is the (age-specific) sex ratio, males to females. Because age-specific population data from ages 35 onwards are only available since the Census of 1991 this Tables uses data from 1991 onwards only. The dependent variable is the log of the crime rate per capita. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). State-level controls include an election year dummy, state police per capita and GDP per capita. There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 10: **Additional Robustness Checks**

	Total			Against Women			Non-Gender Violence			Property			Ec	
	(1)	(2)	from Table ??	(1)	(2)	from Table ??	(1)	(2)	from Table ??	(1)	(2)	from Table ??	(1)	(2)
SR 20-24	0.758***	0.749***	0.450**	0.863***	0.863***	0.767***	1.768***	1.770***	0.557*	0.254	0.248	0.341	0.281	0.26
	(0.220)	(0.221)	(0.228)	(0.244)	(0.241)	(0.231)	(0.279)	(0.278)	(0.301)	(0.228)	(0.229)	(0.241)	(0.213)	(0.21
SR 0-4	-0.207	-0.216	-0.940**	-0.452	-0.373	0.091	-0.508	-0.460	-1.097*	-0.010	-0.002	-0.647	0.878**	0.829
	(0.403)	(0.406)	(0.460)	(0.370)	(0.360)	(0.351)	(0.508)	(0.512)	(0.566)	(0.444)	(0.445)	(0.465)	(0.390)	(0.39
N	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,062	10,0
Adj. R-sq.	0.817	0.817	0.855	0.920	0.920	0.945	0.789	0.791	0.854	0.857	0.858	0.894	0.725	0.72
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
District Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender Lit. Gap	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
$State \times Year$	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

Notes: Coefficients from estimating specification (1) where in columns (1) we do not include any control, in columns (2) we include as set of controls the share of literacy rate, SC/ST share of the population, total annual rainfall (in logs), election year dummy and state police per capita and GDP per capita as well gap. In columns (3) we present the most parsimounious regression as show in Baseline from Table **??**. Robust standard errors are in parenthesis and ar district-decade-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 11: **Addressing measurement error concerns**

	(1) Total	(2) Women	(3) Non-Gender	(4) Property	(5) Economics
			Panel A: Levels		
SR 20-24	1.058***	0.916**	1.860***	0.427	0.038
	(0.390)	(0.448)	(0.464)	(0.387)	(0.376)
SR 0-4	1.015*	0.211	0.522	1.020*	0.526
	(0.560)	(0.616)	(0.681)	(0.598)	(0.657)
Observations	1,290	1,290	1,290	1,290	1,290
Adj. R-sq.	0.688	0.936	0.778	0.823	0.596
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
			Panel B: First- Differences		
SR 20-24	0.406	1.019***	1.274***	-0.012	-0.189
	(0.323)	(0.327)	(0.425)	(0.347)	(0.336)
SR 0-4	0.250	0.347	-0.206	0.571	0.658
	(0.602)	(0.658)	(0.723)	(0.677)	(0.758)
N	1,032	1,032	1,032	1,032	1,032
Adjusted R-squared	0.318	0.661	0.563	0.262	0.311
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: This tables uses data for the years 1971,1981, 1991, 2001 and 2011. Coefficients from estimating equation (1) but with the dependent variable in levels in Panel A and in first-differences in Panel B. SR is the (age-specific) sex ratio, males to females and is measured in levels in Panel A and in first-differences in Panel B. Independent variables are also defined in levels or first-differences, respectively in Panel A and B. District-level controls include rural population, literacy rate, SC/ST share of the population, total annual rainfall (in logs). There are 258 districts. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Appendix B: Additional Individual-Level Results

Table 12: **NFHS III-Summary Statistics: Attitudes towards Intimate Partner Violence**

	Men				Women			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Acceptance Index	0.943	1.390	0	5	1.147	1.629	0	5
Acceptance Binary	0.403	0.490	0	1	0.422	0.494	0	1
Acceptance by Domain:								
Movement	0.226	0.418	0	1	0.257	0.437	0	1
Children	0.289	0.453	0	1	0.314	0.464	0	1
Argues	0.243	0.429	0	1	0.269	0.444	0	1
Sex	0.0722	0.259	0	1	0.128	0.334	0	1
Food	0.113	0.317	0	1	0.179	0.384	0	1
Controls:								
No Education	0.148	0.355	0	1	0.347	0.476	0	1
Primary Education	0.154	0.361	0	1	0.143	0.350	0	1
Secondary Education	0.544	0.498	0	1	0.407	0.491	0	1
Higher Education	0.154	0.361	0	1	0.103	0.304	0	1
Hindu	0.807	0.395	0	1	0.799	0.401	0	1
Muslim	0.133	0.340	0	1	0.133	0.340	0	1
Christian	0.0220	0.147	0	1	0.0223	0.148	0	1
Other	0.0374	0.190	0	1	0.0417	0.200	0	1
Rural	0.480	0.500	0	1	0.549	0.498	0	1
Household Size	5.902	3.082	1	35	6.055	3.059	1	35
Wealth Index	2.305	0.848	1	3	2.250	0.871	1	3
Age Gap at First Marriage*	22.583	4.651	1	48	17.567	3.616	2	44
N	53,240				90,184			

Notes: Using the sample of men and women born between 1957 and 1991 across the major 17 Indian states. Variables acceptance index and binary are obtained from the DHS questions: *"Do you justify wife-beating if..."*: i) *... goes out without telling you?*; ii) *... burns the food?*; iii) *... refuses to have sex?*; iv) *...neglects the children?* and, v) *argues?*. Acceptance index is a measure that varies from 0-5 according to each response of total domains of acceptance of domestic violence. The binary measure takes values 0 if the respondent never justifies physical abuse and 1 if justifies in at least 1 circumstance. The remaining variables are used as controls in specification (3). All data comes from NFHS-III. *: This information uses only the sample of married individuals (N=32.681 for men and N=71.929 for women.

Table 13: **NFHS III-Summary Statistics: Married women Sample**

	(1) Mean	(2) SD	(3) Min	(4) Max
Education Years	7.122	5.147	0	24
Education Gap	5.443	4.494	-7	21
Height	164.7	6.448	106.4	196.2
Height Gap	12.61	7.591	-50.50	58.90
Age	31.61	8.014	15	49
Age Gap	5.481	4.054	-14	32
Join-decision	0.749	0.433	0	1
Controlling	0.318	0.466	0	1
Any IPV	0.250	0.433	0	1
Any IPV Index	0.643	1.502	0	11
Emotional	0.102	0.303	0	1
Emotional Index	0.168	0.555	0	3
Physical	0.206	0.405	0	1
Physical Index	0.492	1.203	0	7
Sexual	0.0508	0.220	0	1
Sexual Index	0.0697	0.320	0	2
Sex Ratio	1.038	0.0644	0.810	1.326

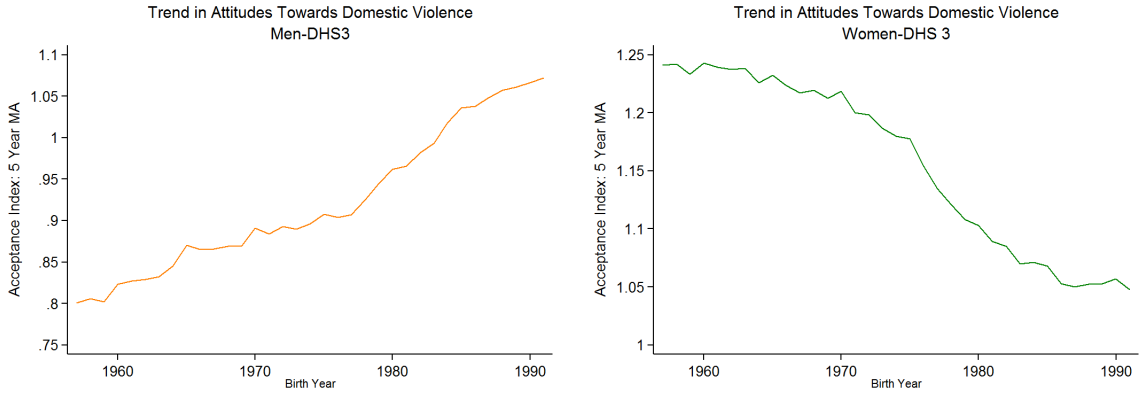
Notes: Using the sample of married women of NFHS III. Variable definitions are in Table 2. Sex Ratio is the ratio of males to females defined by age- group, caste and state.

Table 14: **Sex Ratio and Male Attitudes towards domestic violence - Domain Specific Regressions**

	(1) Movement	(2) Children	(3) Arguing	(4) Sex	(5) Food
			Panel A: Woman		
Post * Child Sex Ratio	-0.286*	-0.252**	-0.167	-0.109	-0.078
	(0.157)	(0.116)	(0.138)	(0.151)	(0.105)
Post	0.304*	0.272**	0.150	0.106	0.069
	(0.164)	(0.124)	(0.142)	(0.159)	(0.107)
Child Sex Ratio	-0.461	-0.161	-0.050	-0.222	0.147
	(0.265)	(0.323)	(0.279)	(0.287)	(0.261)
N	90,184	90,184	90,184	90,184	90,184
Adj. R-sq.	0.118	0.112	0.087	0.068	0.070
			Panel B: Man		
Post * Child Sex Ratio	0.170	0.425**	0.615***	0.292	0.382***
	(0.223)	(0.197)	(0.169)	(0.171)	(0.108)
Post	-0.066	-0.346	-0.604***	-0.295	-0.330**
	(0.236)	(0.221)	(0.182)	(0.184)	(0.127)
Child Sex Ratio	-0.692	-0.683	-0.136	-0.035	-0.425*
	(0.545)	(0.414)	(0.355)	(0.357)	(0.216)
N	53,199	53,199	53,199	53,199	53,199
Adj. R-sq.	0.091	0.130	0.069	0.030	0.047
Controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Birth-Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a a dummy that takes values 1 if an individual tolerates physical abuse under each of the 5 domains. The dependent variables are obtain from the DHS questions: *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...neglects the children?*; iii) *argues?*; iii) *... she refuses to have sex?* and, v) *...she burns the food?*, respectively for columns (1)-(5). Panel A presents results for the women's sample and panel B for the man's sample. The child sex ratio is the ratio of males to females of ages 0-4 in the state-birth year of each individual. This is gathered from state-Census year information from 1961-2011. Post is a dummy that takes values 1 if an individual was born after 1980 and 0 otherwise. All specifications control religion dummies, an indicator for rural/urban household, education attainment dummies (no education is the omitted category) and household wealth index. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Figure 3: **Trends in Attitudes Towards Domestic Violence by Gender**



Notes: Trend using the 5 year moving average of the Index measure of acceptance of domestic violence across birth years. Using NFHS-III data for all men and women.

Table 15: **Intimate-Partner Violence as a function of sex ratio-separate domains**

	Humiliation	Threat	Insult	Push	Slap	Punch	Kick	Strangle	Threat with weapon	Pull Hair	Forced sex	Forced sexual Acts
Post*SR	0.612**	0.367	0.458	0.691*	0.529	0.939***	0.552	1.119	1.604**	0.775**	1.450***	0.971**
	(0.271)	(0.467)	(0.390)	(0.354)	(0.347)	(0.346)	(0.408)	(0.774)	(0.743)	(0.390)	(0.342)	(0.408)
Post	-0.645**	-0.342	-0.464	-0.754**	-0.510	-1.019***	-0.611	-1.170	-1.813**	-0.814**	-1.443***	-0.978**
	(0.269)	(0.461)	(0.384)	(0.358)	(0.345)	(0.361)	(0.421)	(0.826)	(0.777)	(0.389)	(0.335)	(0.397)
SR	-0.597***	-0.036	-0.205	-0.302	-0.645***	-0.638**	-0.613**	-0.302	-0.669	-0.874***	-1.191***	-1.048***
	(0.204)	(0.260)	(0.332)	(0.253)	(0.230)	(0.264)	(0.271)	(0.679)	(0.750)	(0.250)	(0.325)	(0.378)
N	53,074	53,073	53,074	53,062	53,075	53,071	53,071	53,067	53,072	53,061	53,072	53,072
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of intimate-partner violence by separate domains. All dependent variables are dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Using the sample married women questionnaire of DHS-2005. We control for women's religion dummies, whether the woman resides in a rural area. All regressions also include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 16: **Marriage Quality as a function of sex ratio- Heterogeneous effect by Non- Scheduled Caste/Scheduled Tribe**

	(1) Education Years	(2) Education Gap	(3) Age	(4) Age Gap	(5) Any IPV Index	(6) Any IPV	(7) Joint Decision	(8) Controlling	(9) Height	(10) Heigh Gap
Post*Other Castes*SR	-3.340**	0.920	-0.989	-0.159	0.872**	0.634	-1.332***	0.808***	-1.997	-1.011
	(1.421)	(0.945)	(1.071)	(1.175)	(0.340)	(0.397)	(0.344)	(0.306)	(1.696)	(1.604)
Other Castes*SR	3.495*	0.772	-0.370	-0.844	-0.624**	-0.598**	-0.225	-0.802***	0.201	1.448
	(1.672)	(0.820)	(1.319)	(1.178)	(0.257)	(0.289)	(0.304)	(0.285)	(2.035)	(2.035)
Post*SR	-1.218	-0.848	-0.407	-0.886	1.254**	1.059**	-0.983**	1.052***	5.886**	6.950*
	(1.398)	(1.556)	(2.023)	(1.915)	(0.568)	(0.467)	(0.403)	(0.381)	(2.700)	(3.708)
N	71,449	71,443	67,614	67,614	53,077	53,064	71,154	52,779	29,855	29,855
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of husbands quality or spousal gap (e.g. the difference between husband's and wife's age). Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Other Castes is a dummy that takes values 1 if the woman is from a non SC/ST caste and 0 if it is. Using the sample married women questionnaire of DHS-2005. We control for women's religion dummies, whether the woman resides in a rural area. All regressions also include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. The interaction term is a dummy that takes values 1 if a woman belongs to the category other caste. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Joint-decision making is a dummy that takes values 1 if the woman has any say in the husband decision over his income and 0 otherwise; controlling is a measure of controlling behaviour that takes values 1 if the husband exhibits any of the following behaviours:the husband is jealous if respondent talks with other men; husband accuses respondent of unfaithfulness; husband does not permit respondent to meet her girlfriends; husband tries to limit respondent's contact with family; husband insists on knowing where respondent is; husband doesn't trust respondent with money. Any IPV and IPV are an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 17: **Marriage quality as a function of sex ratio- disaggregated measures of Intimate-Partner Violence with effects by Non- Scheduled Caste/Scheduled Tribe**

	(1) Emotional Index	(2) Physical Index	(3) Sexual Index	(4) Emotional	(5) Physical	(6) Sexual
Post*Other Castes*SR	0.197*	0.573*	0.466*	0.112*	0.124	0.634
	(0.109)	(0.277)	(0.225)	(0.055)	(0.116)	(0.397)
Other Castes* SR	-0.089	-0.434**	-0.315*	-0.070	-0.130	-0.598**
	(0.073)	(0.193)	(0.155)	(0.047)	(0.081)	(0.289)
Post*SR	0.091	0.967*	0.765*	0.117	0.241	1.059**
	(0.187)	(0.498)	(0.390)	(0.101)	(0.164)	(0.467)
Observations	53,074	53,075	53,075	53,073	53,065	53,064
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Age-Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are measures of intimate-partner violence by domains: emotional, physical and sexual. Post is a dummy that takes values 1 if the woman was born after 1980 and 0 otherwise. SR is the ratio of males to females in a given state-age-group-caste of the woman. Other Castes is a dummy that takes values 1 if the woman is from a non SC/ST caste and 0 if it is. For each measure we present an index and a dichotomous measure of self-reported incidence of domestic violence as defined in Appendix- Table **??**. Using the sample married women questionnaire of DHS-2005. In Panel A we control for women's religion dummies, whether the woman resides in a rural area and in Panel B we also control for women's educational attainment. All regressions include state, age-group and caste dummies. Caste groups are defined as scheduled castes, scheduled tribes and Other. Age-group dummies are 5-year bins groups starting from 15-19 to 45-49. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 18: **Attitudes Towards Intimate-Partner Violence by Marriage Status**

	Men		Women	
	Index	Binary	Index	Binary
Married	0.214***	0.071***	0.485***	0.138***
	(0.030)	(0.011)	(0.060)	(0.013)
N	47,879	47,879	90,645	90,645

The dependent variable is a a dummy that takes values 1 if an individual tolerates physical abuse under each of the 5 domains. The dependent variables are obtain from the DHS questions: *"Do you justify wife-beating if..."*: i) *...she goes out without telling you?*; ii) *...neglects the children?*; iii) *argues?*; iii) *... she refuses to have sex?* and, v) *...she burns the food?*, respectively for columns (1)-(5). We present results using the index and binary measures as defined in the text. Married is a dummy that takes values one if the individual is married or not. All specifications control religion dummies, an indicator for rural/urban household, education attainment dummies (no education is the omitted category) and household wealth index. Robust standard errors in parentheses are clustered at the state-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Appendix C: Instrumental variables

We explore instrumenting the youth sex ratio with the availability of ultrasound, which Bhalotra and Cochrane (2010) argue was driven by exogenous changes in import and licensing regulations. The first post-ultrasound cohorts, born in the mid-1980s have matured to age 20-24 by the census of 2011. However, although we see a strong trend in the 0-4 sex ratio in the three decades since the mid-80s, because we use the 20-24 sex ratio, we have only one post-ultrasound cohort in our data. Also district variation in ultrasound availability will be endogenous (a function of demand for sex-selective abortion), so the exogenous variation in availability of ultrasound is at the national level. As a result, the power of this potential instrument is weak. We explored interactions of an indicator for post-ultrasound cohorts with an indicator for the Northwest of the country, known to have stronger son preference (Dyson and Moore, 1983; Sen, 2003). We also investigated alternative instruments, motivated by previous work. These include rainfall shocks (Rose, 1999), gold price shocks (Bhalotra et al., 2016) and prices of rice and wheat weighted by land area under these crops (Basu, 1992; Bardhan, 1974), all measured in the birth years of the cohorts who are 20-24 years old in each census year. However we continued to face a weak instruments problem. lthough we have multiple instruments, the first-stage results reveal that we have a weak instruments problem with an F-statistic well below 10 (bottom panel of Table 19, 20 and 21). Nonetheless, we observe a positive and significant coefficient on the instrumented youth sex ratio for violence against women in the second-stage estimates. However, these estimates are imprecise and we do not rely upon them. Results are below in Tables 19- 21.

Table 19: **2SLS Estimates - Set of Instruments 1**

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	1.613	10.717	-4.366	-0.834	8.526
	(5.528)	(7.252)	(7.898)	(6.292)	(5.806)
First-Stage Dep. Var.:	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.008	-0.008	-0.008	-0.008	-0.008
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Rice	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.002	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D.Rainfall	0.016*	0.016*	0.016*	0.016*	0.016*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
D.Rice	0.001	0.001	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
D. Wheat	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Post*Rainfall	0.005	0.005	0.005	0.005	0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Post*Rice	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post*Wheat	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Post*D.Rainfall	-0.014	-0.014	-0.014	-0.014	-0.014
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Post*D.Rice	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Post*D.Wheat	-0.005	-0.005	-0.005	-0.005	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Kleibergen-Paap F-stat	0.778	0.778	0.778	0.778	0.778
N	4,528	4,528	4,528	4,528	4,528
Adj R-sq	0.896	0.931	0.890	0.913	0.805
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: The top part of the table reports the second-stage estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of the crime rate per capita. The set of instruments measured at birth are log of total rainfall, $price \times cultivatedarea$ of rice and wheat, the 1 year difference of these instruments and the interactions of with a dummy that takes values 1 for the birth years post 1980. We control for controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 20: **2SLS Estimates- Set of Instruments 2**

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	-2.532	1.768	-4.267	-1.291	2.218
	(5.069)	(4.699)	(6.019)	(5.356)	(5.107)
First-Stage- Dep Var	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
$Fertilizer_{Nitrogen}$	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Fertilizer_{Phosphorus}$	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Fertilizer_{Potassium}$	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rice	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	4,599	4,599	4,599	4,599	4,599
Adj. R-sq.	0.896	0.946	0.890	0.913	0.834
Kleibergen-Paap F-stat	1.137	1.137	1.137	1.137	1.137
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: The top part of the table reports the second-stage estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of crime rates per capita. The set of instruments measured at birth are *price* $\times$ *quantities* sold of fertilizers nitrogen, phosphorus and potassium measured in Rs/ton. Fertilizer prices are controlled by the Government of India and are based on reported maximum sale prices of fertilizer compounds adjusted for the proportion of the nutrient present in each compound. We always control for district-specific linear trends, state-year dummies, district specific effects and year dummies in addition to controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

Table 21: **2SLS Estimates - Set of Instruments 3**

Second-Stage Dep. Var	(1) Total	(2) Against Women	(3) Non-gender based	(4) Property	(5) Economics
SR 20-24	3.313	8.684	-1.906	-1.012	12.479
	(6.252)	(8.828)	(7.158)	(7.585)	(9.160)
First-Stage- Dep. Var.	SR 20-24	SR 20-24	SR 20-24	SR 20-24	SR 20-24
Rainfall	-0.006	-0.006	-0.006	-0.006	-0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Rice	0.001	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Wheat	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D. Rainfall	0.012	0.012	0.012	0.012	0.012
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
D. Rice	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
D. Wheat	-0.005	-0.005	-0.005	-0.005	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
N	4,502	4,502	4,502	4,502	4,502
Adj. R-sq.	0.893	0.936	0.892	0.913	0.768
Kleibergen-Paap F-stat	0.507	0.507	0.507	0.507	0.507
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
District Linear Trends	Yes	Yes	Yes	Yes	Yes
State-Year	Yes	Yes	Yes	Yes	Yes

Notes: Top part of the table report IV estimates and the bottom panel reports first-stage results. SR is the (age-specific) sex ratio, males to females. All instruments are the 4-year moving average over the years of births of the 4 year bins sex ratio. For example in 1991 the sex ratio of the 20-24 is affected by conditions measured at birth over the period 1967-1971. The dependent variable is the log of crime rates per capita. The set of instruments measured at birth are log of total rainfall, $price \times cultivatedarea$ of rice and wheat and the 1 year difference of these instruments. We always control for district-specific linear trends, state-year dummies, district specific effects and year dummies in addition to controlling for rural population, literacy rate, gender literacy gap, SC/ST share of the population and infant sex ratio. Robust standard errors are in parenthesis and are clustered at the district-level. Significance values at 1,5 and 10% are shown by ***,** or *.

References

Bardhan, P. K., 1974. On life and death questions. Economic and Political Weekly, 1293–1304.

Basu, A. M., 1992. Culture the status of women and demographic behaviour: illustrated with the case of India.

Bhalotra, S., Chakravarty, A., Gulesci, S., 2016. The price of gold: Dowry and death in India. IZA Discussion Paper 9679.

Bhalotra, S., Cochrane, T., 2010. Where have all the young girls gone? Identification of sex selection in India. IZA Discussion Paper 5381.

d'Albis, H., De La Croix, D., 2012. Missing daughters, missing brides? Economics Letters 116 (3), 358–360.

Dyson, T., Moore, M., 1983. On kinship structure, female autonomy, and demographic behavior in India. Population and Development Review, 35–60.

Edlund, L., Li, H., Yi, J., Zhang, J., 2013. Sex ratios and crime: Evidence from China. Review of Economics and Statistics 95 (5), 1520–1534.

Iyer, L., Mani, A., Mishra, P., Topalova, P., 2012. The power of political voice: Women's political representation and crime in India. American Economic Review, Applied Economics (11-092).

Kumar, H., Somanathan, R., 2009. Mapping Indian districts across census years, 1971-2001.

Rose, E., 1999. Consumption smoothing and excess female mortality in rural India. Review of Economics and Statistics 81 (1), 41–49.

Sen, A., 2003. Missing women–revisited: reduction in female mortality has been counterbalanced by sex selective abortions. British Medical Journal 327 (7427), 1297–1299.