

Amaral, Sofia; Bhalotra, Sonia R.

Working Paper

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

ISER Working Paper Series, No. 2017-12

Provided in Cooperation with:

Institute for Social and Economic Research (ISER), University of Essex

Suggested Citation: Amaral, Sofia; Bhalotra, Sonia R. (2017) : Population sex ratios and violence against women: The long-run effects of sex selection in India, ISER Working Paper Series, No. 2017-12, University of Essex, Institute for Social and Economic Research (ISER), Colchester

This Version is available at:

<https://hdl.handle.net/10419/200361>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

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

Sofia Amaral

ISER, University of Essex

Sonia Bhalotra

ISER & Department of Economics, University of Essex

No. 2017-12

October 2017



INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

Non-Technical Summary

Violence against women is a major public health and social problem that is estimated to affect one in three women globally. In this paper, we investigate the role played by sex imbalance in India's population in explaining violence against women. In India, the demographic imbalance is large and has risen since the introduction of ultrasound technology and sex-selective abortion in the 1980s. 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 (the ratio of men to women) rose from 99.22 to 108. 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 linking the sex ratio of youth (age 20-24) to violence against women. Using district level census data matched to administrative crime data over the period 1971-2011, we estimate that increases in the youth sex ratio can account for about a third of the increase in gender-based violence since 1995. In probing mechanisms we present evidence that men are more prone to crime than women, that the share of unmarried men is increasing in the youth sex ratio, that attitudes to violence against women are more regressive for individuals exposed to a higher sex ratio at birth and also that indicators of marriage quality including self-reported domestic violence are negatively related to sex ratios.

Population Sex ratios and Violence against Women: the long-run effects of sex selection in India *

Sofia Amaral
University of Essex[†]

Sonia Bhalotra
University of Essex*

October, 2017

Abstract

This paper investigates the consequences of sex imbalance in India's population for violence against women. We match district level administrative crime data by category to age-specific sex ratios in census data across four decades and, to analyse mechanisms, we also use administrative data on marriage rates and household survey data on attitudes to violence against women and marriage quality. We estimate that the elasticity of violence against women with respect to the surplus of men age 20-24 is unity, and that this explains about 35% of the rise in gender-based violence since 1995. 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 argue that men are more prone to crime than women, that the share of unmarried men is increasing in the youth sex ratio, that attitudes to violence against women are evolving as a function of the sex ratio at birth and marriage quality measures, including self-reported domestic violence, are negatively related to sex ratios.

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

JEL Classification: J12, J16, N34, K42

*ISER & Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, e-mail: srbhal@essex.ac.uk

[†]Institute for Social and Economic Research (ISER), University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. e-mail: sfmont@essex.ac.uk

*We thank seminar participants at Leicester, Kent, ESPE (Glasgow), PPA (Chicago), ISI (Delhi) and ISER for their comments. We also appreciate comments received by Sheetal Sekhri and Abhiroop Mukhopadhyay. This work was 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 in terms of medical care and productivity shortfalls are large, for instance, they are estimated to exceed \$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 is 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 the 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 imbalance is large, 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. At the same time, the relative availability of grooms to brides in the population is estimated to have gone from a deficit of 98.5 potential grooms to 100 potential brides in 1992 to a surplus of 0.6 and 5.1% by 1998 and 2005, respectively (see Table 1).¹ In these same forty years, although total recorded crime in India fell, violence against women increased and is the fastest growing crime category. As explained below, the scope of the VAW data was expanded in 1995, so we cannot provide a growth rate for the entire study period but from 1995 to 2010, VAW increased by 31%.² Since 1995, VAW has constituted 15%

¹To 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.

²Rape is the one category of VAW that has been recorded since 1972 and this shows a rise. See Tables 2 and 3.

of all crimes in India and 27% of violent crimes.

There was massive variation in levels and trends in crime and sex ratios across the Indian districts (Figure 3). We matched district-level census data on age-specific sex ratios and marriage rates across five census years to district-level time series administrative data on crime rates by type of crime to create data from 1972 to 2010 for 258 districts. The independent variable of interest is the youth sex ratio, in particular, the ratio of men to women age 20-24, which is the age at which crime peaks (Theobald and Farrington, 2011).³ 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.

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%. The elasticity is 1.1. Our estimates indicate that the surplus of young men explains more than a third of the rise in gender-based violence over the study period. 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. They also allow for differential trends in the reporting of different types of crime. This is pertinent since violence against women is typically under-reported because of stigma and the fear of retaliation, and because recent evidence suggests that it is changing (Iyer et al., 2012).

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). Some recent work highlights the relevance of non-economic factors such as emotional cues in determining violent crime, including violence against women (Blattman et al., 2015; Card and Dahl, 2011; Dahl and DellaVigna,

³We verify this using the distribution of crime by age of perpetrator in the Indian crime data too. The precise age range simply reflects 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 fairly 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.

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 as a causal factor.

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, potentially 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 to commit crime (Visser 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 crime against women in particular.

However, 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 marriages is decreasing for women (Hussam, 2016). One may imagine that the growing relative scarcity of women would improve their marriage market outcomes. d’Albis and De La Croix (2012) estimate that, in the year 2000, India had 1.4 men for every potential bride. Our own estimates following the authors reveal an upward trend of missing women in the population with a surplus of 5.1% males by 2005 (Table 1).

However it has been argued that population growth combined with an age gap at marriage (of about five years) raises the supply of potential brides, potentially offsetting 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. We present evidence in favour of each of these four potential channels.

Our paper is most closely related to Edlund et al. (2013) and to recent papers by Anukriti et al. (2017) and Hussam (2016). 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

⁴Bhaskar (2011) contrasts this with China.

main difference between their study and this paper is that we focus upon crimes against women.⁵ This is a substantively different question and, as discussed, we find no impacts of the youth sex ratio on property crime.⁶ Anukriti et al. (2017) finds that between 1987 and 2010, imbalanced sex ratios decreased female education attendance by 1 percentage point. In states with higher population imbalance-like Punjab- this effect goes up to 6 percentage points. Hussam (2016) shows that women born after the post-ultrasound increase in the sex ratio at birth in India are in poorer quality marriages, and domestic violence is one of several indicators of the quality of marriage. Our paper differs in focusing upon crime rather than marriage quality. Using administrative data on crime and survey data on domestic violence, we investigate all forms of violence against women, as well as non-gendered crime and we discuss mechanisms among which changes in the marriage market is one.

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 (Angelucci, 2008; Aizer, 2010; Anderberg et al., 2016; Hidrobo et al., 2013; Bhalotra et al., 2016b), and historical attitudes to women (Tur-Prats, 2015; Alesina et al., 2016). We present the first evidence of the importance of the 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, a feature we argue increases crime (Akerlof, 1998; Dorn and Hanson, 2015).

Our second contribution is to a 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., 2016, 2017). Although there is some evidence of long run marriage market and crime effects of sex selection in China (Edlund et al., 2013; Ebenstein and Sharygin, 2009; Rao, 2000) ours is one of the few 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

⁵Data on crimes against women were not available at the province level for China.

⁶In 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).

rights.⁷

The rest of this paper is organized as follows. Section 2 discusses the empirical strategy and specifications. 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 2). 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:

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

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 γ_1 captures for the effect of sex ratios among the crime-marriageable age population and the coefficient π_1 captures for the effect of a surplus of males in ages outside of the crime and marriage market. The equation includes district fixed-effects (λ_d) that absorb unobserved persistent determinants of violence against women that 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](#)).⁸

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

⁷There 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 \(2011\)](#).

⁸[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.

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). A *marriage squeeze* (i.e. the phenomena in which an excess supply of brides co-exists with high sex ratios) can occur when population growth is positive and there is an age gap at marriage as then there can be an excess supply of brides even if sex ratios at birth are male-biased.

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 is still an important predictor of income in India (Virmani, 2006), and 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. The literature indicates that economic conditions influence property crime but not violent crime (including rape) (Ciccarelli, 2017; Freedman and Owens, 2016; Machin and Meghir, 2004; Gould et al., 2002). In India, Iyer and Topalova (2014); Blakeslee and Fishman (2014); Fetzner (2014) have similarly found evidence of a positive relationship between income (as measured by rainfall) and property and violent crime, but not VAW. We nevertheless account for economic conditions by controlling for district level rainfall and state income. To allow that we nevertheless do not capture all potential district-time level variation that may confound the relationship of interest, we include district-specific linear trends (g_{dt}) 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 correlations in the error within district over time, for instance due to measurement error in crime reporting (Burgess et al., 2013; Sekhri and Storeygard, 2014; Blakeslee and Fishman, 2014).

To account for possible measurement errors that may arise in transforming decadal data into yearly data, our preferred specification, we also estimate (1) using only decadal data. In this specification we model the decadal change in crime rates as a function of the sex ratio at the beginning of the decade.

Reverse causality, omitted variable bias and measurement error

A potential threat to our approach is the possible existence of an omitted variable bias that co-determines both VAW and sex ratios. Our model relies on the fact that the drivers of contemporaneous sex ratio are largely determined with a 20-year lag from contemporaneous crime. Thus, for unobservables to affect our estimates these would have to be time-invariant. To take this into account we always include in our estimates district-linear trends. Second, it is possible that the sex ratio of 20-24 year olds changes in response to crime. For instance, violence against women may lead to a more male-biased youth sex ratio if it results in the death of women age 20-24. The inclusion of other older and younger age-groups that are equally affected by crime but unequally affected by the marriage market present a test to whether we are correctly addressing this endogeneity problem. The NCRB does not provide victimization data by age-group, nonetheless, it is plausible that it is the active population that bears the costs of crime and not just the 20-24.

Sex ratios at other ages

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-prefering behaviour enacted twenty years before, making it unlikely that the 20-24 ratio is correlated with contemporary determinants of crime.⁹ What we wish to capture is the excess of young men in the crime and marriage markets conditional upon cultures of son preference.

We therefore include the sex ratio among children age 0-4. An excess of males at age 0-4 does not affect crime rates, as this is a population outside of the crime market, yet an excess of males at 0-4 captures the contemporaneous views about gender that could affect our estimates of the role of excess of males at crime-prone and marriage ages. The inclusion of this age-group addresses the potential concern that the maleness of the youth ratio captures possibly trended local cultures of son preference that are not adequately captured by district linear trends and state-year fixed effects. There is a vast literature documenting that the sex ratio age 0-4

⁹The 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. 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 makes the district the marriage market for both women and men. This will only matter for our identification if sex-differences in migration are correlated with crime. We expect our controls, especially the economic controls to adjust for this. As for mortality, this 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.

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).¹⁰ It also 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 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, in Appendix we provide evidence where we instrument 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 still positive.

Triple difference

Another potential concern relates to under-reporting of crime, and especially crime against women. The implicit assumption is that reporting behaviour is orthogonal to the sex ratio of 20-24 year olds. In view of the controls we include (as mentioned above), including district and year fixed effects, district trends, state-year fixed effects and police coverage per capita, this seems plausible. However, if it is nevertheless the case that women are less likely to report VAW in district-years in which the youth sex ratio is more male-biased, which seems the more likely direction of any violation, then we would under-estimate the relationship of interest. So, 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 order to better address this problem we also estimate a triple-difference type specification, exploiting variation across district, year and also 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

¹⁰(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.

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. The specification is as follows:

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

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 we expect that $\gamma_2 > 0$, given that estimates of the double-difference model produce a small coefficient that is not significantly different from zero. This is expected to show up again in the coefficient β_1 being statistically insignificant. 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). 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 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. This allows to take into account for possible within district-decade correlation in crime categories.

This specification does not provide any new information relative to the crime-specific estimates of equation (1) except that it allows to better take into account differences in reporting behaviour that could contaminate our estimates. 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. Moreover, 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.¹¹ Equation (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

¹¹For instance, any district-year omitted variables that lead to gender differences in migration or mortality that

property crime. It is worth mentioning that since our results are not being driven by the way we use the decadal feature of the sex ratios data, the triple difference estimates are shown using yearly data.

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.¹² The NCRB provides data for 18 categories of crime, using which we construct four major categories. These are violence against women, non-gender based violence, property and economic crime; see Table 2 for definitions.¹³ In order to remove the possibility of spurious results due to the presence of extreme outliers, 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.

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

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.

¹²The Appendix provides a description of how India gathers crime statistics of India and how we aggregated to the district-level across years. After accounting for changes in administrative boundaries over time, we are able to track 84% of the 306 districts contained in the 1971 Census. 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 Uttarakhand are merged with their pre-2001 state boundary definitions. Table ?? presents mean differences tests for the main variables. The sample of districts we are able to track consistently over time is not representative in terms of observable characteristics.

¹³VAW 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).

the incorporation of the new crime categories in 1988 and 1995 respectively). Over the period the rape rate (which is consistently measured through the period) grew from 0.6 to 1.8 crimes per 100,000 population and currently gender-based violence is the fastest growing crime rate in the country. 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 3).

In addition to administrative changes in coding crime, there are trends in reporting crime, especially in women reporting crimes against them. It is inherently difficult to disentangle reporting from incidence. However, using a variety of strategies, [Iyer et al. \(2012\)](#) present compelling evidence that the widespread (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 discussed this, and how we address it, in the empirical strategy section above.

The NCRB does not publish district or state level crime data disaggregated by age and gender of 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 gender-based crime (see Table 5). Second, there is evidence that in India, in line with other countries, crime rates peak among men under the age of 25. The sociology, criminology and economics literature generally points to the crime-prone age being between 16-25 ([Sampson and Laub, 2005](#); [Nagin and Land, 1993](#)).

In India, less than 2% of total crime is committed by individuals younger than 18. Males aged 18-45 commit 77% of crime. Gender-based violence in India is largely committed by men in the age-group of 18-30 (Table 5), with a rate of 42% (ratio of VAW committed by males aged 18-30 to total VAW incidents). The average age at first marriage for males is 22.6 (s.d. 4.7) (Table ??), and the share of unmarried men drops from 63% for the age group 20-24 to 19% in

the age group 25+, and to close to zero by age 35-44 (see Table 4).¹⁴ Overall, these descriptive statistics motivate our decision to use the sex ratio age 20-24 in the analysis. However, as discussed above, we include neighbouring and other age groups in specification checks.

3.2 Sex Ratio Data

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](#)).¹⁵

We use census data on age-specific population sex ratios at the district-level from five rounds of the decadal Census 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 in order to make use of the yearly-variation in crime we use make of several interpolation techniques to estimate the yearly sex ratio. Our main specification uses sex-ratios by age-group transformed using cubic spline interpolation between census years. This approach follow the practice used in the demography literature ([Smith et al., 2004](#))- that is also increasingly used in economics ([Karlsson and Ziebarth, 2016](#)). In robustness tests we also provide estimations using various alternative interpolation methods that include a linear, spline and cubic interpolation. The use of these different methods does not affect our results as across these different methods the estimates of the yearly sex ratios track well the sex ratio over the Census years¹⁶. In a specification check, we drop the inter-censal years and instead use as our

¹⁴Note that the NCRB creates age groups such that 18-30 is the closest to the census age grouping of 20-24.

¹⁵China 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](#)).

¹⁶In the Online Appendix that accompanies this paper we show the trend across these different methods. We also show the trend using the projections methodology used by [Edlund et al. \(2013\)](#) but since this yielded sex ratios estimates that are too far off the real ratio we present results using only the estimates with the 5 different methodologies mentioned.

main dependent variable of interest the crime rate averaged over a decade modelled as a function of sex ratios at the start of the decade and show that the estimates are not sensitive to this (see Table 7.) 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 from the Census of India publications for the rounds 1991, 2001 and 2011. Table 3 presents summary statistics. For 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 uses data from 1991 onwards.

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 3). 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 the 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.

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 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)¹⁷. 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 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

¹⁷The data are available at <http://www.trfetzner.com/data/>, accessed 09.11.2016.

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 we use state per capita income (in logarithms) collected from the Reserve Bank of India.¹⁸ Table 3 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 us 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 4). 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. In addition, we are able to 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 ?? and ??.

4 Results

We present estimates of equation (1) for total crime, violence against women (VAW), non-gendered violence, property and economic crime in Table 6. Moving from columns (1) to (2) we include state-level control variables to our baseline model and from (2) to (3) we include instead state-by-year dummies (our preferred specification). The main, striking 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 8)-and the inclusion of a dummy for the

¹⁸There 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.

introduction of gender political quotas in India (columns (2) of Table 8). In terms of 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%, which is about a third of the rise. In columns (3) of Table 8 we present estimates of (1) using instead a log-log model (see Table 8), and the elasticity is close to one for violence against women.

For non-gendered violence, the elasticity is similar in the log-log model but, in the semi-log model but this result is not robust to a change in functional form. Moreover, this result is also not robust to account for differences in population growth.

For property and economic crime, we estimate elasticities of 0.4 which are 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.¹⁹ The main 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 9). The coefficient on the sex ratio 20-24 increases slightly and the coefficients on the ratios for 15-19 and 25-34 generally have a negative and positive effect on crimes. This is consistent with the fact that men are generally more crime prone (so a higher ratio at older ages increases crime rates). The coefficient on 15-19 is negative and this suggest that in areas with higher ratio of males to females, in pre-marriage years, crime rates decrease. This coefficient on its own is not statistically significant from zero, suggesting a weak relationship with crime- see Table 2 in the Online Appendix. This effect is also small (of about 2.7%) but would be consistent with the fact that this is not a crime-prone age-group. Finding that the age 20-24 ratio is most clearly predictive of VAW also coheres with our hypotheses since this is the age range in which crime peaks and, related, the adult age range in which the share of unmarried men peaks. As discussed before, crime rates are low among

¹⁹It is statistically significant in the equations 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.

men under 18, and marriage rates are high among men over the age of 25. Controlling for the additional sex ratios substantially reduces concerns about omitted trends.

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 evidence of an electoral cycle in crime but only for 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 6 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 1981.²⁰ See Table 7.

As discussed briefly in the Introduction, the deficit of women in the age-group 20-24 may not translate directly into a deficit 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. So as 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 8). This creates no significant change in the main results. Next, we also explore the sensitivity of our results to functional form by including sex ratios in logarithms (columns 2 of Table 8). Since crime is already in logarithms, the coefficients in this log-log model are elasticities. The results are robust to this.

Next, the introduction of gender quotas in villages since the 1990's has been shown to increase reporting rates. We check whether our result is mitigated when we include in our specification the main variable in ([Iyer et al., 2012](#)). We present these results in Table ???. We find that our

²⁰In an alternative specification we followed [Dorn and Hanson \(2015\)](#) and used a first-differenced model. Results remain qualitatively similar and are available upon request.

results remain unchanged when we take this into account. In addition, when we allow the youth sex ratio to vary by the quotas implementation variable we find that while the policy did not alter positively or negatively the effect of youth sex ratio on violence against women.

Another specification check we conducted was to difference the impact of youth sex ratios on crime across crime type. Motivated by the main results which show differential impacts of the male surplus at age 20-24 on these crime categories, we selected property crime as the control crime and violence against women as the treated crime. See Table 10 which displays a sequence of successively richer controls. The relationship between VAW and the youth sex ratio is fairly stable and are smaller than those presented in the non-extended model (of Table 6). This is consistent with the existence of unobservable effects related to reporting behaviour. The coefficients range from 0.524 (0.0.142) to 0.546 (0.0.042), compared to the baseline of 0.767 (0.231). Across specifications, we allowed for crime-specific coefficients on the socio-economic controls since the literature (discussed earlier) suggests that economic factors influence property crime but not violence against women. Our results remain unchanged (column 2). Next, we take into account the 0-4 sex ratio and separately gender gap in education. The results are qualitatively similar (columns 3 and 4). Finally, column 5 presents results from the preferred model where we add state-by-year effects. Results remain unchanged.

In the empirical strategy section, we explained why an instrumental variables approach was not our first choice. Still, we explored it. Although 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 5, 6 and 7). 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.

5 Mechanisms

In the Introduction we discussed alternative mechanisms that might explain why a male surplus at age 20-24 results in increased gender-based violence (VAW). Since the maleness effect- a compositional effect- can 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 previous research linking unmarried men to crime against women, 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:

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

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 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 we are restricted by data availability to two census rounds, 1991 and 2011, for this piece of the analysis. We expect increases in 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 4). Using the National Family Health 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 11. 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 is what is most relevant for the purposes of 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 4 and the social stigma associated with being unmarried is high, at least for women. So our attention is really focused upon age-specific rates and in particular, the data demonstrate that a large fraction of men 20-24 are 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). Our hypothesis is 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). 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 individual. We employ the following difference-in-difference specification:

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

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.²¹ Using these responses we create an accep-

²¹The 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?.

tance 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 1 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 12.

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 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

To investigate the effects of sex ratio on marriage quality we follow a model similar to Anukriti et al. (2017) whereby we use two unique 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).²² The second feature, is the fact that the sex ratio varies extensively along caste lines so that non-scheduled caste/tribe groups (SC/ST) have higher a sex ratio than SC/ST (Chakraborty and Kim, 2010). This has been linked to the fact that son preference is strongest among non-SC/ST groups (Jayachandran, 2017a). In Figure 4 we plot the sex ratio by age and caste groups (on the left) and marriage rates by the same domains (on the right). We see that the sex ratio for non-SC/ST is higher across the age

²²Since we use Census data we cannot identify religion as this information is not available by age-group.

distribution. The all-India average sex ratio is 104.37 males per 100 females for non- SC/ST; 103.59 among SC and 99.77 for ST. This is consistent with the marriage rate pattern in which marriage rates are higher among groups with more equal sex ratios. As a result, we expect that with a strong preference for within-caste marriage and extensive variation in the relative number of brides to grooms across castes a surplus of males is expected to differ across groups depending on the extent of the marriage market distortion.

In our specification, we therefore, exploit the variation in the sex ratio across three dimensions: state, caste and age-group exposure to ultrasound technology. In [Anukriti et al. \(2017\)](#) the authors follow a similar strategy but use instead the sex ratio of alive individuals as their measure of sex ratio to look at the effects on female education investments and marital outcomes. The authors find that a higher sex ratio at birth decreases female education, age at marriage and labour market participation. These results are consistent with parent’s substituting the relative gains from educational investments to those of marrying off girls earlier. Moreover, it is also consistent with a higher sex ratio decreasing women’s position in the marriage market. We follow this approach but focus on direct measures of partner quality. In our model, we take the male’s parents (and the husband) to be the primary and first decision-makers in terms of marriage negotiations and intra-household allocations ([Eswaran and Malhotra, 2011](#)) and thus, we model the husband’s quality 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 exogenous marriage market determinants. Our triple difference specification is as follows:

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

where, Y_{isca} are the outcomes measuring the quality of the husband of woman i . The interaction term $(SR_{sca} \times Post_a)$ measures the differential effect of sex ratio on the husband’s quality for a given woman i in state of caste group c born before-after the introduction of ultrasound. The coefficient of interest is δ_2 . In our specification we control for predetermined characteristics of the wife (vector X_i) such dummies identifying factors such as whether the respondent resides in a rural area and religion dummies. We also control for state, caste and age-group dummies (γ_s , ζ_c and λ_t). Our outcomes of interest are the education years, 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, incidence of physical, emotional and sexual intimate-partner violence (IPV) and willingness to share decisions about money ²³. We use both the woman’s and the couples modules of NFHS-3 in order to retrieve all available measures of quality.²⁴ 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 educational investments may affect matching and bargaining outcomes we nonetheless provide estimations of (5) with and without controlling for female education (see Panel B of Table 13). Results are presented in Table 13 and 14.

We find that higher sex ratios among women born after the introduction of ultrasound technology decrease women’s marriage quality in several dimensions. In areas with higher the sex ratio, women are more likely to marry a husband with less education years, be victims of domestic violence and have less decision-making power (Table 13. The coefficients on height gap (columns 4 and 10 of Table 13) suggest that higher widens the height gap between spouses but this coefficient is not statistically significant.

In terms of magnitudes, a one standard deviation increase in the sex ratio (0.064) leads to an increase in 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. Moreover, women are also 5.9 percentages points more likely to face a controlling partner and 8.5 percentages points less likely of being involved in financial decision-making. Overall, this suggests that women’s position in the marriage market has worsened and as a result, their well-being within the household is diminished. As expected, when controlling for female education the magnitudes of the effect sizes decrease for all measures but they remain significant for measures of domestic violence and decision-making. Importantly, when conditioning on women’s education,

²³Our 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 certain very specific type of behaviours over a specific period (Table 2 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.

²⁴The 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’s files. Our results do not change whether we use the couples or the women’s files but in order to improve the precision of our estimations we report results using only the married women’s sample. Results from the couples’s files are available upon request.

the remaining effect of sex ratios on the spouse’s education is largely diminished.

Finally, we have also disaggregated our measure of intimate-partner violence by types of violence (emotional, physical and sexual) to pin point the sources of spousal increased violent behaviour. 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 14 and ??.

In a separate specification, we have also isolated the effects of non-SC/ST and SC/ST groups by repeating the estimation of model (5) with the additional interaction ($SR_{sca} \times Post_a \times Others_i$). This allow us to estimate the differential effect of a surplus of males across non SC/ST and the remaining caste groups with less unequal sex ratio. We find that among those with higher sex ratios, women married men with less education and are less likely to engage in joint decision-making.

This results are in line with Hussam (2016) who uses NFHS-3 data, working off cluster-level variation in the sex ratio at birth pre-post ultrasound to look at changes in the quality of marriages for women. Hussam (2016) sets the cluster as the village (i.e. the primary sampling unit in the NFHS) to reflect the local marriage market. However, the marriage market is unlikely to be the village as women tend to marry outside of their natal village (see for example Rosenzweig and Stark (1989)). This pattern seems to remain unchanged since the study of Rosenzweig and Stark (1989). In fact, in the Indian Human Development Survey-2005, the rate of women marrying a man outside of her natal village was 87.05%.²⁵ Nonetheless, despite the fact that marriage location could be endogenous to marriage quality, it is reassuring that our results are consistent with ours.

6 Discussion

Our findings are consistent with a mechanical effect of increasing 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 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

²⁵This information is not available in the NFHS.

alone is 0.5, and the elasticity we estimate is 1.

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 contributes to VAW, 56% of which is on account of rape, molestation, harrassment 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 (?).

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). While it seemed obvious to attribute our results to this third mechanism on its own, we do not think this is the case given that our main findings for crime cannot be attributed solely to changes in domestic violence reports. Instead, we show that while it seems that one the hand, marriage quality is deteriorating for women and this explains part of the effects on crime; one the other hand, in general, attitudes among men are worsenning as marriage markets become more competitive for men. This result while counterintuitive is not new and is consistent with findings from other settings ([Grosjean and Khattar, 2015](#)).

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 11 percent increase in VAW. Our

results suggest that changes in the youth sex ratio account for as much 35% of the change in VAW. This result is robust to a number of specification checks.

As far as we are aware, 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 evidence 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. It is plausible that our findings are due dowry payments. While there is some evidence that dowry payments are rising [Anderson \(2003\)](#) and are correlated with violence against women [Bloch and Rao \(2002\)](#). This would be consistent with theoretical models that predict that rising women's economic empowerment can represent a backlash towards women ([Eswaran and Malhotra, 2011](#)). However, this hypothesis goes beyond the scope for what this paper can deliver.

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 they are marrying later and they have less progressive attitudes towards violence against women. This is important as recent evidence suggests 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](#)).

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.
- 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.
- 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., 2003. Why dowry payments declined with modernization in europe but are rising in india. *Journal of Political Economy* 111 (2), 269–310.

- 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., 2016. On the quantity and quality of girls: New evidence on abortion, fertility, and parental investments.
- Anukriti, S., Pop-eleches, C., Salanié, B., Turner, L., 2017. Sex ratios, schooling, and the marriage market in India.
- 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.
- 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., Rawlings, S., Siddique, S., Kumbathi, U., 2016b. 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.
- Bongaarts, J., Guilimoto, 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’économie* 44 (4), 1222–1263.
- Fetzer, T., 2014. Can workfare programs moderate violence? evidence from India. Working Paper.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.

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

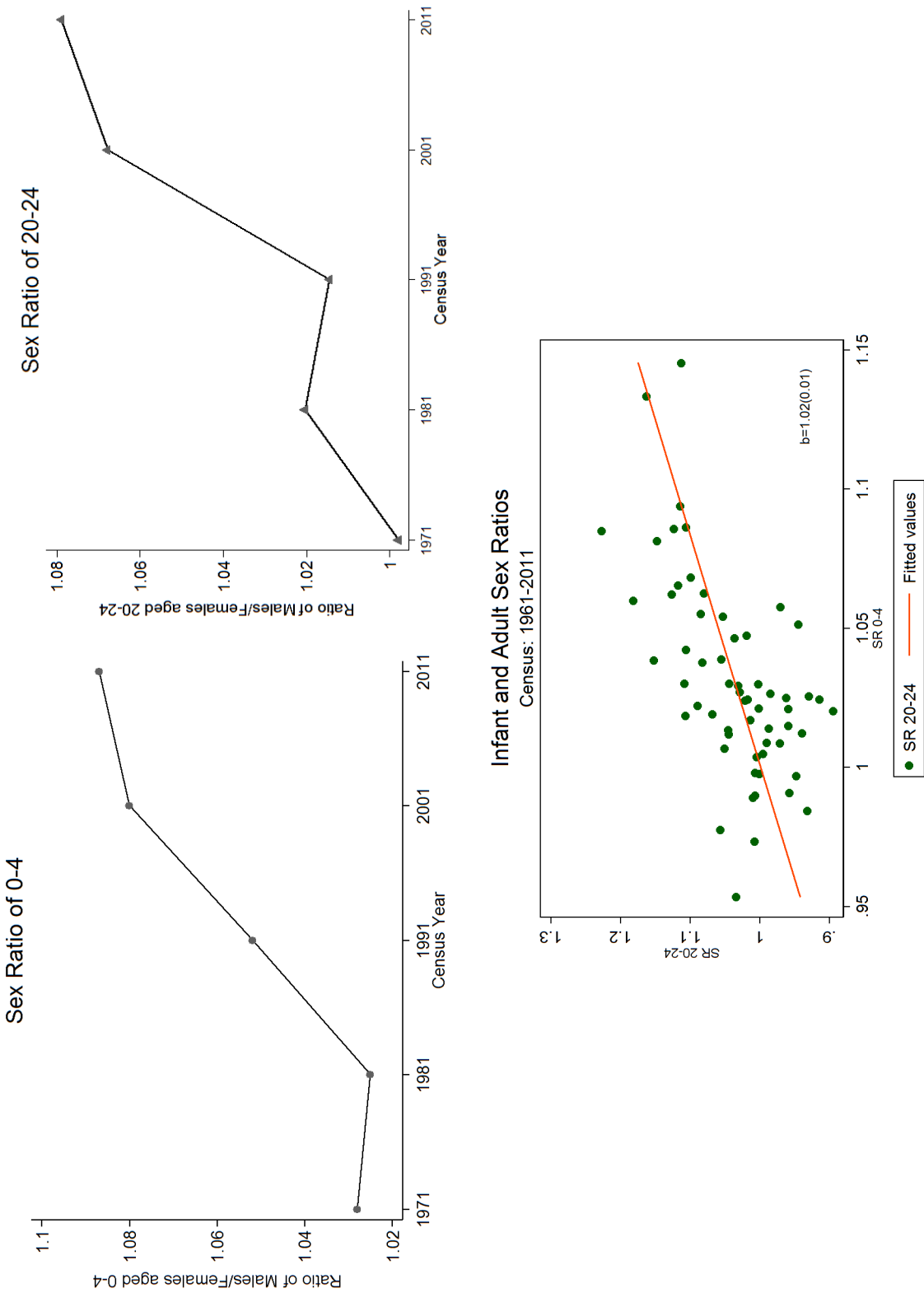
Notes: This table presents a replication of the Missing Brides Index developed by d'Albis and De La Croix (2012) using individual level data from the National Family Health Surveys rounds 1,2,3 for the years 1992-1993, 1998-1999 and 2005-2006, respectively. We use the measure of fertility rate as in d'Albis and De La Croix (2012) for thw years closest to each of the survey rounds. We created each of the components of the Index in d'Albis and De La Croix (2012) considering the mean variables of each survey. 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 age at childbearing subtracted by the age at marriage and; the sex ratio 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	
Women	Total incidents per 100, 000 population. Includes the incidents registered under male kidnappings, riots, arson and hurt, dacoity and murder.
Non-gender based Violence	Total incidents per 100, 000 population. Includes the incidents registered under male kidnappings, riots, arson and hurt, dacoity and murder.
Property	Total incidents per 100, 000 population. Includes the incidents registered under burglary, robbery and theft.
Economic	Total incidents per 100, 000 population. Includes the incidents registered under counterfeiting, criminal breach of trust and Cheating
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 pull her hair
Forced sex	Spouse ever physically forced sex when not wanted
Forced sexual Acts	Spouse ever forced other sexual acts when not wanted

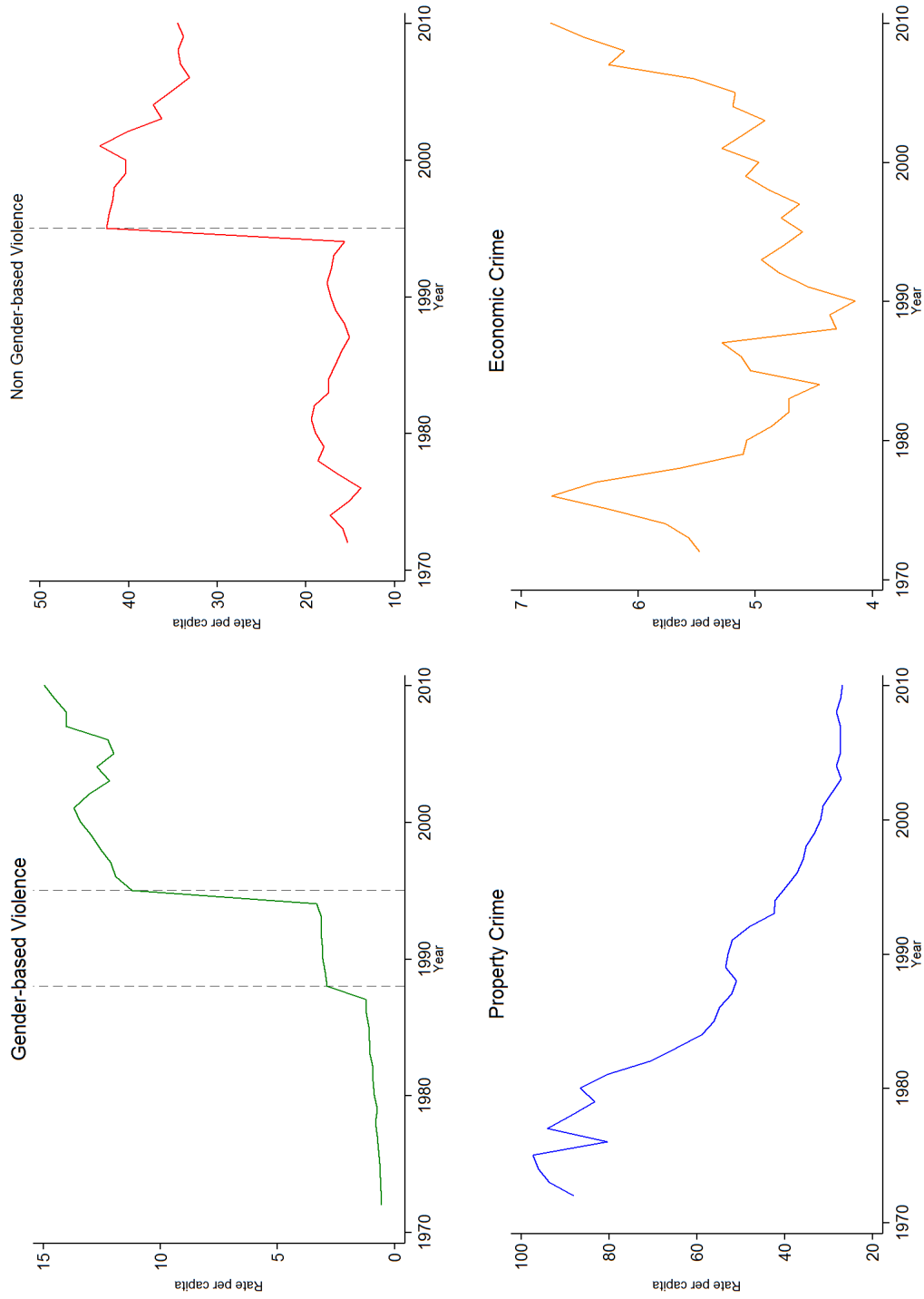
Notes: Panel A presents the definitions of crime rates variables from the National Crime Records Bureau (NCRN). Crime definitions are from the Indian Penal Code. Panel B present definitions of self-reported measures of intimate-partner violence from the National Family Health Survey, wave III (NFHS). 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..

Figure 1: Trends in Age-specific Sex Ratios



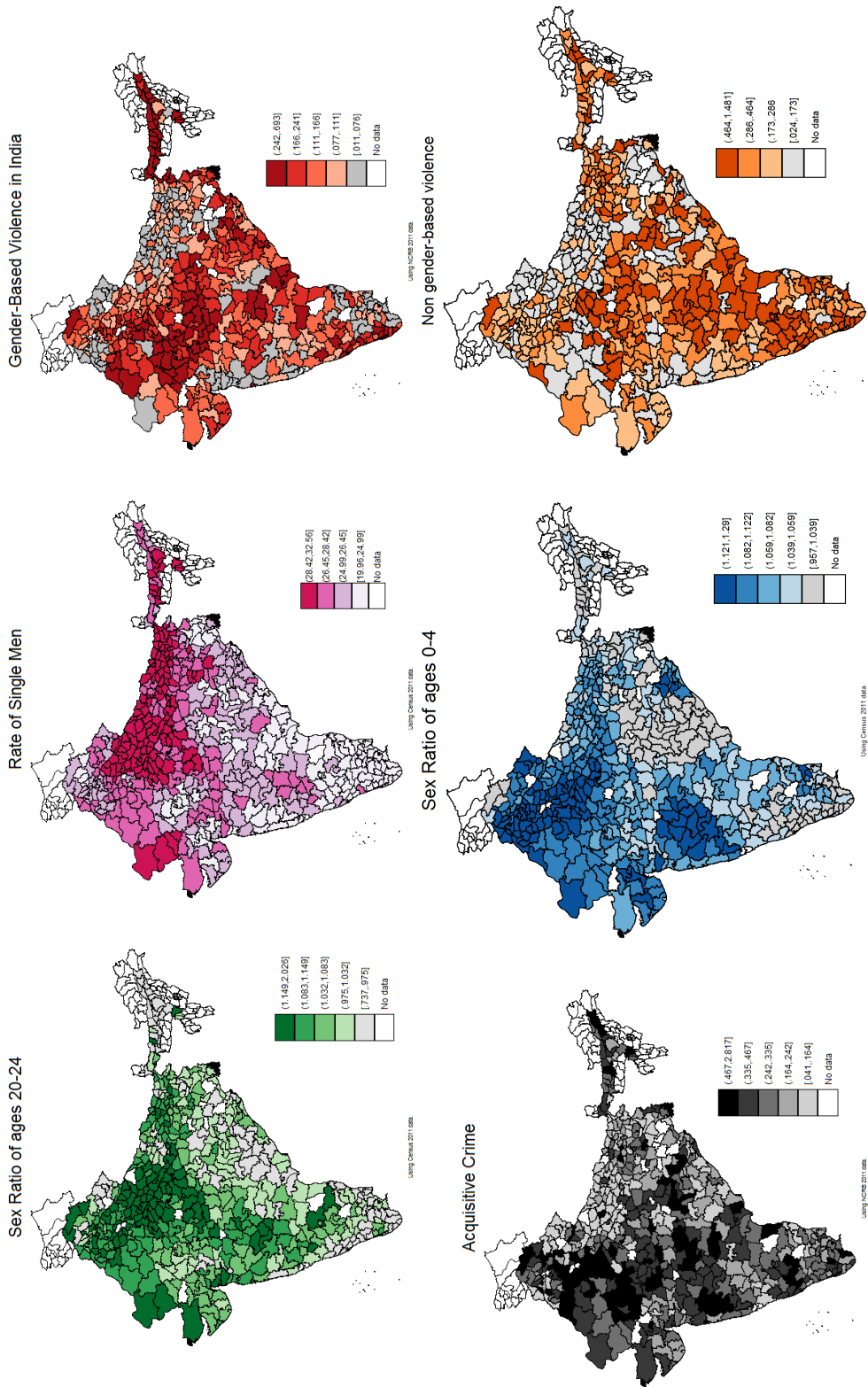
Notes: The top panels show the trend in the average ratio of male to female population of age 0-4 on the left and 20-24 on the right. These trends are obtained using pooled district-level data from Census 1971-2011. The bottom panel shows the linear fit and scatter plot of the ratio of males to females 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 2.

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 2.

Table 3: Summary Statistics- Part I

Panel A: Crime Rates (per 100,000)				
	Mean	SD	Δ 1972-2010	Δ 1995-2010
Total Crime	90.430	58.586	-0.241	-0.152
Total against Women				
Rape	6.233	8.325	26.15	0.336
Domestic Violence*	1.361	1.424	2.007	0.0382
Molestation*	5.149	4.875		0.845
Sexual Harassment*	3.926	3.818		-0.0230
Kidnapping of Females*	0.784	1.465		0.575
	1.497	1.616		0.0973
Total non-gender based				
Murder	25.61	23.83	1.259	-0.189
Dacoity	3.319	2.278	-0.251	-0.372
Riots	1.016	1.685	-0.865	-0.557
Kidnapping of Males*	9.689	10.33	-0.525	-0.532
Arson*	1.908	1.957		-0.228
Hurt*	1.059	1.112		-0.486
	26.79	26.25		-0.0234
Property				
Robbery	53.37	48.10	-0.695	-0.323
Thefts	2.610	3.064	-0.561	-0.291
Burglary	33.41	33.27	-0.657	-0.247
	17.35	15.48	-0.776	-0.470
	5.225	4.641	0.232	0.469
Economic				
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.

Table 4: Summary Statistics- Single Rates

		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 - 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.

Table 6: Crime as a function of the youth sex ratio: Yearly changes estimation

	(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*** (0.221)	0.785*** (0.220)	0.450** (0.228)	0.863*** (0.241)	0.877*** (0.241)	0.767*** (0.231)	1.770*** (0.278)	1.871*** (0.279)	0.557* (0.301)	0.248 (0.229)	0.249 (0.228)	0.341 (0.241)
SR 0-4	-0.216 (0.406)	-0.100 (0.404)	-0.940** (0.460)	-0.373 (0.360)	-0.333 (0.362)	0.091 (0.351)	-0.460 (0.512)	-0.062 (0.514)	-1.097* (0.566)	-0.002 (0.445)	-0.057 (0.445)	-0.647 (0.465)
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
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
District FE	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
District Controls	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
State Controls	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

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 7: Crime as a function of sex ratio - Decadal changes estimation

	(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 (0.841)	-0.572 (0.911)	1.997** (0.825)	1.859** (0.888)	1.213 (1.132)	-0.548 (1.251)	-0.246 (0.926)	-0.256 (0.992)	-0.784 (0.969)	-0.821 (1.013)
SR 0-4	1.704 (1.726)	0.617 (2.111)	1.400 (1.775)	0.073 (2.141)	0.650 (2.369)	0.431 (2.644)	1.875 (1.963)	1.448 (2.289)	3.736 (2.560)	1.121 (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, gender literacy gap, 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: 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 A: Total Crime Rate							
SR 20-24	0.384	0.762***	0.533	2.176**	0.449*	0.450**	0.462**
	(0.242)	(0.221)	(0.475)	(1.101)	(0.255)	(0.228)	(0.222)
SR 0-4	-1.079**	-0.171	-0.968	-5.141**	-1.300***	-0.940**	-0.940**
	(0.498)	(0.406)	(0.734)	(2.013)	(0.498)	(0.460)	(0.460)
Adj. R-sq.	0.856	0.819	0.855	0.853	0.855	0.855	0.855
Panel B: Against Women Crime Rate							
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 C: Non-Gender-Based Crime Rate							
SR 20-24	0.508	1.827***	0.564	2.583*	0.545	0.557*	0.568*
	(0.313)	(0.280)	(0.607)	(2.359)	(0.336)	(0.301)	(0.298)
SR 0-4	-1.326**	-0.198	-1.117	-5.128**	-1.425**	-1.097*	-1.097*
	(0.619)	(0.518)	(1.007)	(2.471)	(0.612)	(0.566)	(0.566)
Adj. R-sq.	0.854	0.798	0.854	0.852	0.854	0.854	0.854
Panel D: Property Crime Rate							
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
Panel E: Economic Crime Rate							
SR 20-24	0.286	0.289	0.418	2.063**	0.417*	0.342	0.297
	(0.234)	(0.214)	(0.326)	(1.027)	(0.249)	(0.226)	(0.225)
SR 0-4	-0.613	0.924**	-0.554	-3.001	-0.706	-0.504	-0.506
	(0.450)	(0.388)	(0.621)	(2.034)	(0.470)	(0.430)	(0.430)
Adj. R-sq.	0.787	0.727	0.783	0.761	0.783	0.783	0.783

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) our sex ratio measure is the log of the sex ratio by age-group; 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, gender literacy gap, 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-decade level. Significance values at 1,5 and 10% are shown by ***,** and *.

Table 9: Crime as a function of the youth sex ratio: Robustness exercise with neighbouring age-group

	Total	Against Women	Non-Gender Violence	Property	Economics
SR 20-24	0.324 (0.228)	0.866*** (0.239)	0.417 (0.305)	0.195 (0.246)	0.309 (0.222)
SR 0-4	-0.970** (0.475)	-0.016 (0.366)	-1.118* (0.586)	-0.683 (0.476)	-0.536 (0.437)
SR 15-19	-0.394** (0.196)	-0.283* (0.149)	-0.380 (0.274)	-0.464** (0.198)	-0.216 (0.170)
SR 25-34	0.492* (0.266)	-0.149 (0.205)	0.521 (0.333)	0.572** (0.269)	0.175 (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 10: Crime as a function of sex ratio: Differencing across crime types

Dep var: Crime Rate	(1)	(2)	(3)	(4)	(5)
SR 20-24*VAW	0.524*** (0.142)	0.546** (0.042)	0.542* (0.043)	0.543* (0.044)	0.545** (0.029)
SR 20-24	-0.102 (0.227)	-0.159 (0.138)	-0.378 (0.125)	-0.346 (0.119)	0.088 (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	Yes	Yes

Notes: Table reports coefficients from estimating specification (2). The dependent variable is the log of crime rates per capita by 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 2. District-level controls include rural population, literacy rate, gender literacy gap, 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 11: Sex Ratio and Never-Married Rates

	Males Single Rate			Females Single Rate		
	(1) 15-44	(2) 15-44	(3) 20-24	(1) 15-44	(2) 15-44	(3) 20-24
SR 15-44	0.0786 (0.0604)			-0.0413 (0.0621)		
SR 20-24		0.129*** (0.0328)	0.192*** (0.0601)		0.119*** (0.0274)	-0.137*** (0.0484)
SR 25-34		0.158*** (0.0406)	0.0165 (0.0698)		0.124*** (0.0318)	0.0878* (0.0506)
SR 35-44		-0.108** (0.0427)	0.116 (0.0765)		0.0333 (0.0374)	0.281*** (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

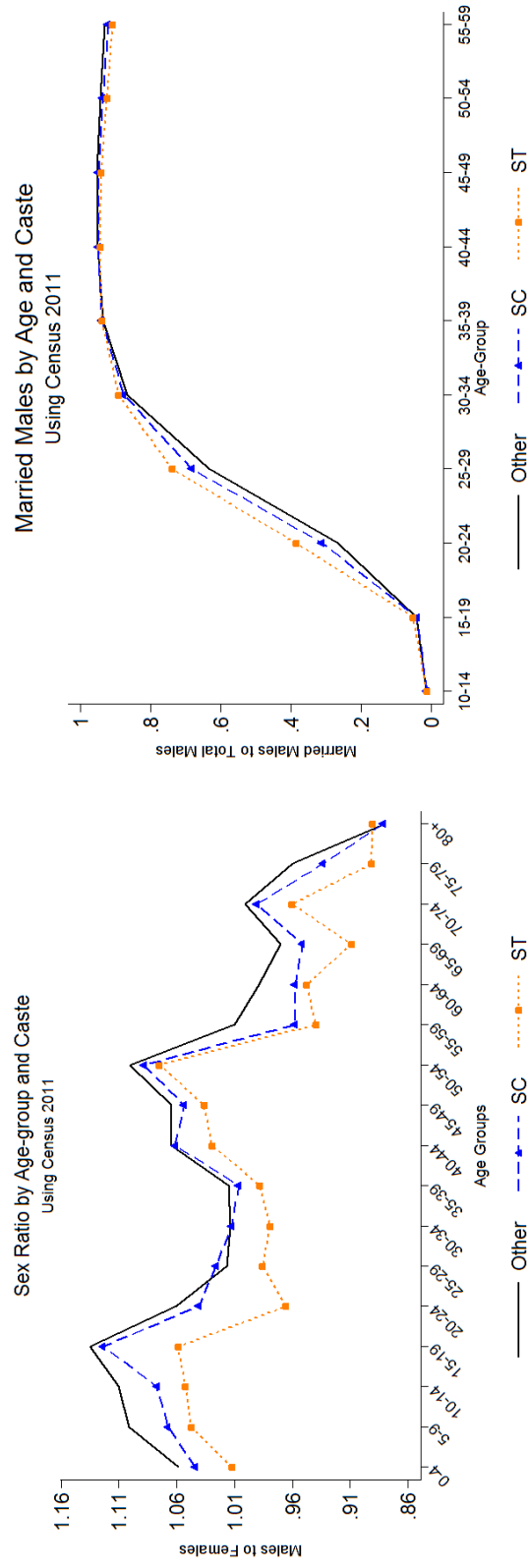
Notes: The dependent variable is the ratio of never married individuals by gender and age over total population of the same age-gender. SR is the (age-specific) sex ratio, males to females. Columns (1) and (2) use the 15-44 age-group and columns (3) use the 20-24 age-group. We control for district-level 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 12: Attitudes towards Intimate Partner Violence

	Men				Women			
	(1) Index	(2) Index	(3) Binary	(4) Binary	(1) Index	(2) Index	(3) Binary	(4) Binary
Post * Child Sex Ratio	1.883** (0.739)	2.047** (0.713)	0.533** (0.216)	0.562** (0.210)	-0.891 (0.521)	-0.907* (0.477)	-0.416*** (0.118)	-0.431*** (0.119)
Post	-1.641* (0.816)	-1.792** (0.813)	-0.465* (0.249)	-0.551** (0.226)	0.901 (0.529)	1.001** (0.459)	0.429*** (0.130)	0.488*** (0.137)
Child Sex Ratio	-1.972 (1.726)	-1.908 (1.671)	-0.590 (0.485)	-0.599 (0.476)	-0.746 (1.210)	-1.041 (1.086)	-0.165 (0.255)	-0.259 (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 13: Marriage quality as a function of sex ratio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Education Years	Education Gap	Age	Age Gap	Joint Decision	Controlling	Any IPV Index	Any IPV	Height	Height Gap
Panel A: Baseline										
Post*SR	-4.129*** (1.390)	-0.275 (1.038)	-0.461 (1.188)	0.004 (1.212)	-1.332*** (0.313)	0.916*** (0.258)	0.924*** (0.271)	0.710** (0.311)	-0.571 (1.169)	0.404 (1.369)
Post	4.386*** (1.401)	0.316 (1.089)	-1.898 (1.207)	0.089 (1.224)	1.287*** (0.315)	-0.907*** (0.268)	-0.957*** (0.268)	-0.692** (0.313)	0.343 (1.217)	-0.827 (1.383)
SR	3.049* (1.651)	0.389 (0.872)	-0.086 (1.316)	-0.436 (1.123)	-0.231 (0.309)	-0.846*** (0.276)	-0.834*** (0.229)	-0.840*** (0.270)	0.154 (1.870)	1.040 (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** (0.691)	-1.597** (0.691)	-0.693 (1.196)	-0.248 (1.198)	-1.257*** (0.293)	0.796*** (0.250)	0.737*** (0.227)	0.527* (0.300)	0.302 (1.083)	0.544 (1.420)
Post	1.713** (0.733)	1.713** (0.733)	-1.652 (1.215)	0.356 (1.210)	1.208*** (0.294)	-0.780*** (0.260)	-0.762*** (0.223)	-0.499 (0.308)	-0.524 (1.070)	-0.967 (1.434)
SR	1.301 (1.014)	1.301 (1.014)	0.075 (1.323)	-0.261 (1.134)	-0.284 (0.325)	-0.796*** (0.288)	-0.724*** (0.199)	-0.754*** (0.272)	-0.337 (1.908)	0.983 (2.163)
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 14: Intimate-partner violence as a function of sex ratio

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